
**Forensic Examination of Australian
Papers Using Isotope Ratio Mass
Spectrometry**

Kylie Jones

PhD Thesis: Science

University of Technology, Sydney

2016

CERTIFICATE OF ORIGINAL AUTHORSHIP

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Student:

Date:

ACKNOWLEDGEMENTS

I would firstly like to express my appreciation and thanks to my role model, mentor and supervisor Dr Sarah Benson, without whom I would not have been able to do this. Your constant support, patience and knowledge has been pivotal in providing the framework for me to become the forensic scientist and leader I am. Thank you!

To my UTS supervisor Claude Roux, thank you for supporting this project and providing a consistent source of encouragement and guidance in all things academic.

To the AFP Document Sciences team – in particular Rochelle Epple and Katrina Stokes – thank you for supporting me over the course of this project. I know it has put more pressure on you all in a number of ways (including having an absent, distracted or cranky team leader!) and I appreciate the support more than you know. Thank you also to Timothy Simpson who put up with my endless questions, complaints, blonde moments and general isotopic chatter.

Most importantly - to my family! Words cannot describe how lucky I feel to have you all. You taught me from an early age that I could achieve anything I put my mind to which is an amazing place to start. To my Darr, you were right – one small step at a time – you just forgot to mention how many steps there were in finishing this PhD! To Mum, Rachael and Tanya – thank you for always being proud of me and inspiring me to keep pushing. To my fur-baby Atticus – thank you for keeping my lap warm and for distracting me endlessly over the last three years. And finally, the biggest thank you is to my husband Ben, who lived with me from the first to the final day of my PhD. You never once doubted my ability to complete this research, even through my most emotional moments!

A big thank you also needs to be given to colleagues and students that have assisted with preparation of standards and samples or physical measurements over the life of this project. Timothy Simpson, Felicity Koens, Rochelle Epple, Alexandra Cameron, and Szymon Dzubinski – thank you!

Forensic Examination of Australian Papers Using Isotope Ratio Mass Spectrometry

- Table of Contents -

ABSTRACT	X
LIST OF PUBLICATIONS AND PRESENTATIONS	XI
LIST OF TABLES	XIII
LIST OF FIGURES	XVII
CHAPTER 1	1
1. LITERATURE REVIEW AND AIMS	1
1.1. INTRODUCTION	1
1.2. AIMS OF THIS RESEARCH	2
1.3. PULPING AND PAPER MAKING	3
1.3.1. PULPING	6
1.3.2. PULP BLEACHING AND WHITENING.....	7
1.3.3. PAPER MAKING.....	8
1.4. FORENSIC EXAMINATION OF PAPER	10
1.4.1. PHYSICAL EXAMINATION METHODS	10
1.4.2. CHEMICAL EXAMINATION METHODS.....	13
1.5. ISOTOPE RATIO MASS SPECTROMETRY	17
1.5.1. STABLE ISOTOPES.....	17
1.5.2. ISOTOPE RATIO NOTATION AND INTERNATIONAL STANDARDS	18
1.6. ISOTOPIC SCALE CALIBRATION AND CORRECTION.....	20
1.6.1 METHOD FOR CORRECTING RAW INSTRUMENT DATA	20
1.6.2 DRIFT CORRECTION.....	21
1.6.3 OUTLIER EXCLUSION USING GRUBBS TEST.....	22
1.6.4 ISOTOPIC SCALE CALIBRATION.....	23
1.7. ISOTOPIC FRACTIONATION.....	25
1.7.1. EQUILIBRIUM ISOTOPE EFFECTS.....	26
1.7.2. KINETIC ISOTOPE EFFECTS	28
1.7.3. CALCULATION OF ISOTOPIC FRACTIONATION.....	29
1.8. STABLE ISOTOPE INSTRUMENTATION.....	30
1.8.1. ISOTOPE RATIO MASS SPECTROMETER	30

1.8.2.	CONTINUOUS FLOW IRMS.....	32
1.8.3.	LASER SPECTROSCOPIC TECHNIQUES	36
1.9.	OVERVIEW OF ISOTOPE ECOLOGY AND ISOTOPIC DISCRIMINATION	36
1.9.1.	ISOTOPE CYCLING AND WHOLE EARTH ISOTOPE VALUES.....	36
1.10.	PHOTOSYNTHESIS AND PLANT ISOTOPIC VALUES	41
1.10.1.	CARBON ISOTOPES.....	41
1.10.2.	OXYGEN AND HYDROGEN ISOTOPES.....	45
1.10.3.	ISOSCAPES.....	47
1.11.	USE OF IRMS IN FORENSIC SCIENCE	47
1.11.1.	ANALYSIS OF ORGANIC EVIDENCE USING IRMS	48
1.11.2.	IRMS USED FOR THE COMPARISON OF DOCUMENT PAPERS	50
1.12.	RESEARCH STRATEGY	50
CHAPTER 2	52
2.	CARBON METHOD DEVELOPMENT AND VALIDATION	52
2.1.	INTRODUCTION AND AIMS.....	52
2.2.	MATERIALS AND METHODS	53
2.2.1.	STANDARDS AND SAMPLES.....	53
2.2.2.	INSTRUMENTATION AND EQUIPMENT.....	55
2.2.3.	CORRECTION OF VALUES.....	56
2.2.4.	STATISTICAL TESTS	56
2.2.5.	METHOD DEVELOPMENT	57
2.2.6.	METHOD VALIDATION	59
2.3.	RESULTS AND DISCUSSION – METHOD DEVELOPMENT	61
2.3.1.	LINEARITY.....	62
2.3.2.	WATER ABSORPTION.....	67
2.3.3.	SAMPLE HANDLING, STORAGE AND PAPER EQUILIBRATION	69
2.3.4.	DESIGN OF SAMPLE RUN.....	71
2.4.	RESULTS AND DISCUSSION – METHOD VALIDATION	76
2.4.1.	PRECISION AND STABILITY	76
2.4.2.	REPEATABILITY	78
2.4.3.	ACCURACY.....	78
2.4.4.	ROBUSTNESS	80
2.4.5.	MEASUREMENT UNCERTAINTY.....	82
2.4.6.	INTER-LABORATORY TRIAL.....	83

2.5.	CONCLUSIONS.....	83
CHAPTER 3.....		84
3.	CARBON ISOTOPES – BACKGROUND POPULATION AND HOMOGENEITY.....	84
3.1.	INTRODUCTION	84
3.2.	MATERIALS AND METHODS	85
3.2.1	SAMPLES AND STANDARDS.....	86
3.3.	RESULTS AND DISCUSSION.....	88
3.3.1	BACKGROUND POPULATION STUDY – CARBON ISOTOPES	88
3.3.2	PAPER HOMOGENEITY WITHIN ONE REAM	93
3.3.3	HOMOGENEITY BETWEEN REAMS OF THE SAME BRAND.....	100
3.3.4	DISCRIMINATION OF BACKGROUND POPULATION.....	105
3.4.	CONCLUSIONS.....	105
CHAPTER 4.....		108
4.	CARBON ISOTOPES – SOURCE, HANDLING AND FORENSIC SAMPLING.....	108
4.1.	INTRODUCTION	108
4.2.	MATERIALS AND METHODS	110
4.2.1	SOURCE AND PAPER PRODUCTION SAMPLES	111
4.2.2	USAGE OF PAPERS – PRINTING PROCESSES	113
4.2.3	USAGE OF PAPERS - ENVIRONMENTAL EFFECTS AND FORENSIC TESTING	114
4.3	RESULTS AND DISCUSSION	115
4.3.1	SOURCE AND PAPER PRODUCTION SAMPLES	115
4.3.2	USAGE OF PAPERS – PRINTING PROCESSES	122
4.3.3	USAGE OF PAPER – ENVIRONMENTAL EFFECTS AND FORENSIC TESTING	126
4.4.	CONCLUSIONS.....	131
CHAPTER 5.....		134
5.	OXYGEN ISOTOPES – METHOD DEVELOPMENT AND VALIDATION.....	134
5.1.	INTRODUCTION	134
5.2.	MATERIALS AND METHODS	134
5.2.1	STANDARDS AND SAMPLES.....	134
5.2.2	INSTRUMENTATION AND EQUIPMENT	135
5.2.3	CORRECTION OF VALUES AND STATISTICAL TESTS	137
5.2.4	METHOD DEVELOPMENT	137
5.2.5	METHOD VALIDATION	138
5.3.	RESULTS AND DISCUSSION.....	139

5.3.1	OXYGEN METHOD DEVELOPMENT.....	139
5.3.2	METHOD VALIDATION.....	145
5.4.	CONCLUSIONS.....	156
CHAPTER 6.....	157	
6.	OXYGEN ISOTOPES – BACKGROUND POPULATION AND HOMOGENEITY.....	157
6.1.	INTRODUCTION.....	157
6.2.	MATERIALS AND METHODS.....	157
6.3.	RESULTS AND DISCUSSION.....	158
6.3.1.	BACKGROUND POPULATION STUDY – OXYGEN ISOTOPES.....	158
6.3.2.	PAPER HOMOGENEITY WITHIN ONE REAM.....	163
6.3.3.	HOMOGENEITY BETWEEN REAMS OF THE SAME BRAND.....	169
6.3.4.	DISCRIMINATION OF BACKGROUND POPULATION.....	172
6.4.	CONCLUSIONS.....	173
CHAPTER 7.....	174	
7.	OXYGEN ISOTOPES – SOURCE, HANDLING AND FORENSIC SAMPLING.....	174
7.1.	INTRODUCTION.....	174
7.2.	MATERIALS AND METHODS.....	174
7.2.1.	SOURCE AND PAPER PRODUCTION SAMPLES.....	175
7.2.2.	USAGE OF PAPER – PRINTING PROCESSES.....	176
7.2.3.	USAGE OF PAPERS – ENVIRONMENTAL EFFECTS AND FORENSIC TESTING.....	176
7.3.	RESULTS AND DISCUSSION.....	177
7.3.1.	SOURCE AND PAPER PRODUCTION SAMPLES.....	177
7.3.2.	USAGE OF PAPERS – PRINTING PROCESSES.....	184
7.3.3.	USAGE OF PAPERS – ENVIRONMENTAL EFFECTS AND FORENSIC TESTING.....	187
7.4.	CONCLUSIONS.....	190
CHAPTER 8.....	191	
8.	HYDROGEN ISOTOPES – PRELIMINARY STUDY.....	191
8.1.	INTRODUCTION.....	191
8.2.	MATERIALS AND METHODS.....	193
8.2.1.	STANDARDS AND SAMPLES.....	193
8.2.2.	INSTRUMENTATION AND EQUIPMENT.....	194
8.3.	RESULTS AND DISCUSSION.....	196
8.3.1.	DISCRIMINATION WITHIN THE SAMPLE SET.....	196
8.4.	CONCLUSION.....	199

CHAPTER 9	200
9. OTHER PAPER EXAMINATION TECHNIQUES	200
9.1. INTRODUCTION	200
9.2. MATERIALS AND METHODS	202
9.2.1. STANDARDS AND SAMPLES	202
9.2.2. COMPARISON USING UV, TRANSMITTED AND AMBIENT LIGHT	202
9.2.3. MEASUREMENT OF GRAMMAGE AND APPARENT SHEET DENSITY	203
9.2.4. THIN LAYER CHROMATOGRAPHY	204
9.2.5. X-RAY DIFFRACTION ANALYSIS	206
9.3. RESULTS AND DISCUSSION	206
9.3.1. AMBIENT, UV AND TRANSMITTED LIGHT	206
9.3.2. GRAMMAGE AND APPARENT SHEET DENSITY	209
9.3.3. THIN LAYER CHROMATOGRAPHY	214
9.3.4. COLLATION AND COMPARISON OF TECHNIQUES	214
9.3.5. X-RAY DIFFRACTION	218
9.4. CONCLUSIONS	219
CHAPTER 10	221
10. EXAMINATION PROTOCOL, INTERPRETATION OF RESULTS AND REPORTING	221
10.1. INTRODUCTION	221
10.2. COMBINING TECHNIQUES	222
10.2.1. USING CARBON AND OXYGEN ISOTOPES FOR DISCRIMINATION	222
10.2.2. USING IRMS RESULTS WITH OTHER EXAMINATION TECHNIQUES	228
10.2.3. PROPOSED EXAMINATION PROTOCOL	233
10.3. INTERPRETATION OF RESULTS	235
10.3.1. FACTORS AFFECTING INTERPRETATION	237
10.3.2. PROPOSED INTERPRETATION WORKFLOW	239
10.4. REPORTING OF RESULTS	242
10.4.1. PROPOSED REPORTING SCALE AND REPORT CONTENT	243
10.5. CONCLUSIONS	244
CHAPTER 11	245
11. BLIND TRIAL OF EXAMINATION PROTOCOL	245
11.1. INTRODUCTION	245
11.2. METHOD AND MATERIALS	246
11.2.1. ORGANISATION OF BLIND TRIALS	246

11.2.2.	CHANGE TO CARBON METHOD	248
11.3.	RESULTS AND DISCUSSION	249
11.3.1.	TRIAL A	249
11.3.2.	TRIAL B	252
11.3.3.	TRIAL C.....	254
11.3.4.	TRIAL D	256
11.3.5.	TRIAL E.....	258
11.3.6.	TRIAL F.....	260
11.3.7.	TRIAL G	262
11.3.8.	TRIAL H	265
11.4.	COMPARISON OF TRIAL FINDINGS TO KNOWN RESULTS.....	269
11.5.	CONCLUSIONS.....	271
CHAPTER 12	272
12.	CONCLUSIONS AND FUTURE WORK	272
13.	REFERENCES	275
14.	APPENDIX 1 – DATA TABLES	286
15.	APPENDIX 2 - ANOVA TABLES FOR CARBON HOMOGENEITY EXPERIMENTS	312
15.1	SINGLE REAM HOMOGENEITY STUDY.....	312
15.2	BETWEEN REAM HOMOGENEITY STUDY	319
16.	APPENDIX 3 - ANOVA TABLES FOR OXYGEN HOMOGENEITY EXPERIMENTS	326
16.1	SINGLE REAM HOMOGENEITY STUDY.....	326
16.2	BETWEEN REAM HOMOGENEITY STUDY	333
17.	APPENDIX 4 – PROPOSED TECHNICAL REPORT APPENDIX	339
18.	APPENDIX 5 – ELECTRONIC FILES ON INCLUDED DISC.....	341
18.1	COMPARISON TABLES	341
18.1.1	CARBON ISOTOPES	341
18.1.2	OXYGEN ISOTOPES	341
18.1.3	HYDROGEN ISOTOPES	341
18.1.4	APPARENT SHEET DENSITY AND GRAMMAGE	341
18.2	THIN LAYER CHROMATOGRAPHY PLATE IMAGES	341
18.3	AMBIENT, UV AND TRANSMITTED LIGHT IMAGES	341
18.4	COPY OF PUBLISHED WORK	341

ABSTRACT

Isotope Ratio Mass Spectrometry (IRMS) has been shown to be a useful tool in the comparison of materials that are chemically identical or have been naturally produced. Based on this, and noting the capability gaps within the Forensic Document Examination field, the measurement of carbon and oxygen isotopic abundance values using IRMS has been developed as a technique for the examination of document papers. Through validation of the inter- and intra-ream variability of papers, appropriate guidelines for comparison and discrimination have been constructed, to ensure the technique is robust and accurate. Through the measurement of a background population of 125 papers, 89% of samples collected from within Australia and New Zealand were discriminated using pair-wise comparisons.

The IRMS results were placed in a broader context, through the use of a range of light, physical and chemical techniques. Based on these results, an examination and interpretation protocol was defined and tested through the use of a range of scenarios in a blind trial. All results within the blind trials were correct, demonstrating that the examination, comparison and reporting structure defined was accurate, robust and fit for purpose. As a result of this project, a paper examination protocol which is operationally relevant to Australian law enforcement has been developed and validated and is ready for use in forensic casework examinations.

LIST OF PUBLICATIONS AND PRESENTATIONS

The following is a list of publications relating to this research. A copy of each of these articles is included electronically within Appendix 5.

K. Jones, S. Benson, C.Roux, *The forensic analysis of office paper using carbon isotope ratio mass spectrometry, Part 1: Understanding the background population and homogeneity of paper for the comparison and discrimination of samples*, Forensic Science International, 2013, Vol 231(1), 354-363.

K. Jones, S. Benson, C.Roux, *The forensic analysis of office paper using carbon isotope ratio mass spectrometry, Part 2: Method Development, validation and sample handling*, Forensic Science International, 2013, Vol 231(1), 364-374.

Winner – National Institute of Forensic Science, Best Technical Article or Note, 2014

K. Jones, S. Benson, C.Roux, *The forensic analysis of office paper using carbon isotope ratio mass spectrometry, Part 3: Characterising the source materials and the effect of production and usage on the $\delta^{13}\text{C}$ values of paper*, Forensic Science International, 2013, Vol 233(1-3), 355-364.

K. Jones, S. Benson, and C. Roux, *The forensic analysis of office paper using oxygen isotope ratio mass spectrometry. Part 1: Understanding the background population and homogeneity of paper for the comparison and discrimination of samples*, Forensic Science International, 2016, Vol 262, pp.97-107.

The following is a list of conferences that this research has been presented at:

Forensic Isotope Ratio Mass Spectrometry Conference, 2010, Washington DC –
Poster

Australian and New Zealand Forensic Science Society Conference, 2010, Sydney –
Oral Presentation

European Association of Forensic Science, 2010, The Hague – Oral Presentation

Australian and New Zealand Forensic Science Society Conference, 2012, Hobart – Oral Presentation, *Winner – Best Oral Presentation in Document Examination*

Forensic Isotope Ratio Mass Spectrometry Conference, 2013, Montreal – Oral Presentation

American Society of Questioned Document Examination/ Australasian Society of Questioned Document Examination Joint Meeting, 2014, Honolulu – Oral Presentation

Australian and New Zealand Forensic Science Society Conference, 2014, Adelaide – Oral Presentation

European Association of Forensic Science, 2015, Prague – Poster Presentation

LIST OF TABLES

Table 1.1: Common steps involved in pulp whitening and de-lignification	8
Table 1.2: Common additives in the paper making process.	9
Table 1.3: The relative natural abundances of commonly measured stable isotopes	18
Table 1.4: International Isotopic Reference Materials	19
Table 1.5: Isotopes measured by BSIA and the actual gas species measured	33
Table 2.1: International standard materials used in the method validation.....	53
Table 2.2: Materials used in the carbon method validation experiments and for evaluation as laboratory standards	54
Table 2.3: $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ results for varying sample handling/storage conditions.....	70
Table 2.4: $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ results for varying international standard number and replicate	73
Table 2.5: $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ values of samples corrected using two or three sets of international standards.	74
Table 2.6: Summary of the analytical results obtained for three international standard materials.	77
Table 2.7: Summarized results from accuracy experiments of cellulose acetate, medium cellulose and alpha glucose.	80
Table 2.8: Mean $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ and standard deviation for samples prepared by 3 different operators	81
Table 3.1: Brand and origin of paper samples used for homogeneity studies.	87
Table 3.2: Recycled papers, mean $\delta^{13}\text{C}_{\text{VPDB}}$ and standard deviation for 7 recycled papers collected as samples for comparison to virgin paper background samples	92
Table 3.3: Measured $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ abundances from single ream homogeneity testing	94
Table 3.4: Mean $\delta^{13}\text{C}_{\text{VPDB}}$ and standard deviation (‰) results for between ream homogeneity.....	101
Table 4.1: Cellulose extracted and bulk measured $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰}) \pm 1$ standard deviation for paper source and production samples.....	116
Table 4.2: Cellulose extracted and bulk measured $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰}) \pm 1$ standard deviation, for additives used in paper production	117
Table 4.3: Papers measured for $\delta^{13}\text{C}_{\text{VPDB}}$ before and after acidification using hydrochloric acid.....	120

Table 4.4: Brands of ream used and pooled mean values for environmental and forensic test evaluation.....	126
Table 4.5: $\delta^{13}\text{C}_{\text{VPDB}}$ measurement values for environmental testing experiment	127
Table 4.6: Results of statistical t-tests for three conditions, on five papers.....	129
Table 4.7: $\delta^{13}\text{C}_{\text{VPDB}}$ measurement values for forensic testing experiment.....	131
Table 5.1: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) \pm 1 Std Dev (‰) results for replicates of coumarin, PEG and international cellulose run over 12 weeks.....	147
Table 5.2: Summarized data for sample drying experiments.....	152
Table 5.3: Summary of the analytical results obtained for three international standard materials, measured over time for their $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values	155
Table 6.1: Recycled papers, mean $\delta^{18}\text{O}_{\text{VSMOW}}$ and standard deviation for 7 recycled papers collected as samples for comparison to virgin paper background samples	161
Table 6.2: Measured $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) abundances from single ream homogeneity testing.	164
Table 6.3: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) and standard deviation (‰) results for between ream homogeneity.....	170
Table 7.1: Cellulose extracted and bulk measured $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) \pm 1 standard deviation for paper source and production samples.....	178
Table 7.2: Cellulose extracted and bulk measured $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) \pm 1 standard deviation, for additives used in paper production	178
Table 7.3: Paper $\delta^{18}\text{O}_{\text{VSMOW}}$ values before and after acidification	182
Table 7.4: Brands of ream used and pooled mean values for environmental and forensic test evaluation.....	188
Table 7.5: $\delta^{18}\text{O}_{\text{VSMOW}}$ measurement values for environmental testing experiment.....	188
Table 7.6: $\delta^{18}\text{O}_{\text{VSMOW}}$ measurement values for forensic testing experiment	189
Table 8.1: Discrimination powers obtained from pair comparisons of 25 samples measured for $\delta^2\text{H}_{\text{VSMOW}}$ values.....	198
Table 9.1: Discrimination of samples against paper sample 71, using a range of examination techniques, with the x denoting discrimination using that technique.	216
Table 9.2: Discrimination of samples against paper sample 71.....	217
Table 9.3: Filler content for 14 papers analysed by XRD.....	218
Table 10.1: Discrimination of Australian samples using oxygen and carbon isotopes	227

Table 10.2: Comparison of techniques used to discriminate 60 samples of 80gsm paper.	230
Table 10.3: Interpretation factors with a discussion of their impacts on confidence....	241
Table 11.1: Results of grammage, apparent sheet density and	250
Table 11.2: Interpretation considerations for blind trial A	251
Table 11.3: Results of grammage, apparent sheet density and isotopic measurements for blind trial B	252
Table 11.4: Interpretation considerations for blind trial B.....	253
Table 11.5: Results of grammage, apparent sheet density and	254
Table 11.6: Interpretation considerations for blind trial C.....	255
Table 11.7: Results of grammage, apparent sheet density and isotopic measurements for blind trial D	256
Table 11.8: Interpretation considerations for blind trial D	257
Table 11.9: Results of grammage, apparent sheet density and isotopic measurements for blind trial E.....	258
Table 11.10: Interpretation considerations for blind trial E.....	259
Table 11.11: Results of grammage, apparent sheet density and isotopic measurements for blind trial F	261
Table 11.12: Interpretation considerations for blind trial F	261
Table 11.13: Results of grammage, apparent sheet density and isotopic measurements for trial G.....	263
Table 11.14: Interpretation considerations for blind trial G	264
Table 11.15: Results of grammage, apparent sheet density and isotopic measurements for trial H.....	267
Table 11.16: Interpretation considerations for blind trial H	268
Table 11.17: Collated results of blind trials A-H.....	269
Table 13.1: Summarized data for linear range experiments	286
Table 14.1: Summarized data for linear range experiments for $\delta^{13}\text{C}$ method validation	286
Table 14.2: Day 1 vs. Day 16 $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ results for samples exposed to the laboratory atmosphere.	287
Table 14.3: Paper samples collected as background population samples, and their measured $\delta^{13}\text{C}$ and $\delta^{18}\text{O}_{\text{VSMOW}}(\text{‰})$ abundance values.....	294

Table 14.4: Printer manufacturer and paper information for toner printers tested for their effect on the $\delta^{13}\text{C}$ abundance values of paper.....	296
Table 14.5: Mean $\delta^{13}\text{C}_{\text{VPDB}}$ and 1 x standard deviation results for 46 toner printed documents	298
Table 14.6: Printer manufacturer and paper information for inkjet printers tested for their effect on the $\delta^{13}\text{C}$ abundance values of paper.....	299
Table 14.7: Mean $\delta^{13}\text{C}_{\text{VPDB}}$ and 1 x standard deviation results for 14 inkjet printed documents	299
Table 14.8: Summarized data for linear range experiments for $\delta^{18}\text{O}$ method validation	300
Table 14.9: Printer manufacturer and paper information for toner printers tested for $\delta^{18}\text{O}_{\text{VSMOW}}$	301
Table 14.10: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ and 1 x standard deviation results for 22 toner printed documents	302
Table 14.11: Printer manufacturer and paper information for inkjet printers tested for $\delta^{18}\text{O}_{\text{VSMOW}}$	302
Table 14.12: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ and 1 x standard deviation results for 8 inkjet printed documents	303
Table 14.13: Paper samples selected and measured for their $\delta^2\text{H}_{\text{VSMOW}}$ (‰) values....	304
Table 14.14: Mean grammage and apparent sheet density measured for papers.....	306
Table 14.15: Grammage and Apparent Sheet Density results for background papers, with green cells denoting difference values larger than the defined benchmark values.....	308
Table 14.16: Results of analyses undertaken using Thin Layer Chromatography	311

LIST OF FIGURES

Figure 1.1: Schematic showing pulping and paper production process.....	5
Figure 1.2: Chemical constituents of wood.....	6
Figure 1.3: Schematic of typical paper making machine, showing key production stages	8
Figure 1.4: Schematic demonstrating grain direction in papers.....	12
Figure 1.5: Two-point calibration plot for $\delta^{13}\text{C}_{\text{VPDB}}$ (‰) using polyethylene and sucrose standards.....	24
Figure 1.6: Bond energy differences in molecules containing ^2H	27
Figure 1.7: Carbon equilibrium fractionation during C_3 photosynthesis, top showing general pathway of movement of CO_2 , bottom showing fractionation factors (Δ) based on current atmospheric $\delta^{13}\text{C}$ value of -8 ‰	28
Figure 1.8: Schematic of an IRMS instrument showing different measurement options – either Continuous Flow or Dual Inlet.....	31
Figure 1.9: Schematic of an open split inlet in a continuous flow IRMS	33
Figure 1.10: Carbon isotopic abundance ranges for naturally occurring materials	37
Figure 1.11: Oxygen isotopic abundance ranges for naturally occurring materials	38
Figure 1.12: Effect of a range of factors on $\delta^2\text{H}$ and $\delta^{18}\text{O}$ values of water	40
Figure 1.13: Typical sources of variation in the $\delta^{13}\text{C}$ distribution. Double arrows denote an equilibrium fractionation event	41
Figure 1.14: Calvin cycle with direct input of CO_2 , as utilised by C_3 plants in photosynthesis	42
Figure 1.15: Calvin cycle with indirect input of CO_2 via an organic acid intermediate, as utilised by C_4 plants in photosynthesis.....	43
Figure 2.1: Internal packing materials used in the Flash EA 1112 for the measurement of carbon.....	56
Figure 2.2: Typical chromatogram for measurement of $\delta^{13}\text{C}$	61
Figure 2.3: Linear response for $\delta^{13}\text{C}$ of medium fibre cellulose.....	63
Figure 2.4: Linear response for $\delta^{13}\text{C}$ of cellulose acetate	63
Figure 2.5: Linear response for $\delta^{13}\text{C}$ of alpha glucose.....	63
Figure 2.6: Linear range for $\delta^{13}\text{C}$ of medium fibre cellulose.....	64
Figure 2.7: Linear range for $\delta^{13}\text{C}$ of cellulose acetate	64

Figure 2.8: Linear range for $\delta^{13}\text{C}$ of alpha glucose.....	64
Figure 2.9: Linearity of CO+ production (measured on mass 28) for paper, cellulose acetate and alpha glucose samples.	66
Figure 2.10: Day 1 (blue/diamond) vs. Day 16 (red/square) $\delta^{13}\text{C}_{\text{VPDB}}$ values for samples left exposed to laboratory atmosphere	67
Figure 2.11: Mean $\delta^{13}\text{C}_{\text{V-PDB}}(\text{‰})$ paper sample values for different treatment and storage procedures.....	70
Figure 2.12: Plot of corrected measurements of international standard polyethylene over a 24-month period.	76
Figure 2.13: Plot of corrected measurements of international standard sucrose over a 24-month period.	77
Figure 2.14: Plot of corrected measurements of international standard cellulose over a 24-month period.	78
Figure 2.15: Accuracy results for Medium Cellulose	79
Figure 2.16: Accuracy results for Cellulose Acetate	79
Figure 2.17: Accuracy results for Alpha Glucose	80
Figure 2.18: Mean $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ values obtained for samples prepared and run by three different operators.	81
Figure 3.1: Mean $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ values of 125 virgin pulp papers plotted by region of origin	89
Figure 3.2: Summary of mean $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ for region of origin of papers, with error bars denoting 1 x standard deviation.....	90
Figure 3.3: Australian papers plotted over time, according to ream packaging date (earliest to latest date).	92
Figure 3.4: Mean $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ values for seven different brands measured during homogeneity and sampling tests	95
Figure 3.5: Boxplot of Fuji Xerox $\delta^{13}\text{C}_{\text{VPDB}}$ results taken from a single ream. Error bars shown represent 95% range for discrimination.....	97
Figure 3.6: Standard deviation values (‰) for homogeneity and sampling test.....	98
Figure 3.7 (a-g): Boxplots denoting the mean $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ and standard deviation (‰) for seven reams from seven different brands	103
Figure 4.1: Mean $\delta^{13}\text{C}$ values of cellulose extracted from samples collected throughout the paper production process.....	118

Figure 4.2: Comparison of mean $\delta^{13}\text{C}$ values of extracted cellulose (diamond markers) with bulk values (square markers) collected from throughout the paper production process.....	118
Figure 4.3: Mean $\delta^{13}\text{C}$ values of bulk paper (diamond markers) compared with the same paper samples after undergoing acidification (square markers). Sample number is taken from Table 4.3.	121
Figure 4.4: Comparison of the mean $\delta^{13}\text{C}_{\text{VPDB}}$ values of papers before and after printing using toner.....	123
Figure 4.5: Comparison of the $\delta^{13}\text{C}_{\text{VPDB}}$ values of papers before and after printing using inkjet ink.....	124
Figure 4.6 a-c: Box plots of five papers measured before and after environmental testing	130
Figure 5.1: Configuration and materials used in the furnace of the TC/EA	136
Figure 5.2: Linear response for $\delta^{18}\text{O}$ of paper	140
Figure 5.3: Linear response for $\delta^{18}\text{O}$ of IAEA-CH-3.....	140
Figure 5.4: Linear response for $\delta^{18}\text{O}$ of IAEA-601	140
Figure 5.5: Linear response for $\delta^{18}\text{O}$ of IAEA-602	141
Figure 5.6: Linear range for $\delta^{18}\text{O}$ of paper.....	141
Figure 5.7: Linear range for $\delta^{18}\text{O}$ of IAEA-CH-3.....	141
Figure 5.8: Linear range for $\delta^{18}\text{O}$ of IAEA-601.....	142
Figure 5.9: Linear range for $\delta^{18}\text{O}$ of IAEA-602.....	142
Figure 5.10: Typical chromatogram for measurement of oxygen in IAEA-601.....	144
Figure 5.11: Typical chromatogram for measurement of oxygen in document paper samples.....	144
Figure 5.12: Calibration of VSMOW and SLAP2 International Standards.....	145
Figure 5.13: IAEA-CH-3 measured as a quality control material over the course of 18 months of experiments.	154
Figure 5.14: Corrected values for IAEA-601, used as a standard material over the course of 18 months of experiments.....	154
Figure 5.15: Corrected values for IAEA-602, used as a standard material over the course of 18 months of experiments.....	155
Figure 6.1: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values of 123 virgin pulp papers collected from Australia and New Zealand.....	158

Figure 6.2: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values of 123 virgin pulp papers plotted by region of origin.....	159
Figure 6.3: Summary of mean $\delta^{13}\text{C}_{\text{VPBD}}$ (‰) for region of origin of papers, with error bars denoting 1 x standard deviation.....	160
Figure 6.4: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values of Australian paper samples collected over time, plotted in order of their packing date.....	162
Figure 6.5: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values of Australian paper samples averaged per season, from 2009 until 2012.....	162
Figure 6.6 (a-g): Boxplots of $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values for single ream homogeneity....	167
Figure 6.7: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values for seven different brands measured during homogeneity and sampling tests.....	168
Figure 6.8: Standard deviation values (‰) for homogeneity and sampling test.....	168
Figure 6.9 a-g: Boxplots denoting the mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) and standard deviation (‰) for seven reams from seven different brands.....	171
Figure 7.1: Mean $\delta^{18}\text{O}$ values of cellulose extracted from samples collected throughout the paper production process.....	180
Figure 7.2: Comparison of mean $\delta^{18}\text{O}$ values of extracted cellulose (diamond markers) with bulk values (square markers) collected from throughout the paper production process.....	180
Figure 7.3: Mean $\delta^{18}\text{O}$ values of bulk paper (diamond markers) compared with the same paper samples after undergoing acidification (square markers).	182
Figure 7.4: Comparison of the mean $\delta^{18}\text{O}_{\text{VSMOW}}$ values of papers before and after printing using toner.....	185
Figure 7.5: Comparison of the $\delta^{18}\text{O}_{\text{VSMOW}}$ values of papers before and after printing using inkjet ink.....	186
Figure 8.1: Typical chromatogram for H_2 measurement of document papers.....	195
Figure 8.2: Mean $\delta^2\text{H}_{\text{VSMOW}}$ (‰) values of 25 papers collected from Australia and New Zealand.....	196
Figure 8.3: Mean $\delta^2\text{H}_{\text{VSMOW}}$ (‰) values of 25 papers plotted by region of origin.....	197
Figure 9.1: Image of developed TLC plate, showing edge effect on the sample run in the far right position.....	205
Figure 9.2: Example of a simple discrimination task using transmitted light between sample 71 (left) and sample 94 (right).	207

Figure 9.3: Example of a more complex discrimination task using transmitted light between sample 71 (left) and sample 85 (right).....	207
Figure 9.4: Grammage vs. Apparent Sheet Density for seven paper brands	210
Figure 9.5: Grammage vs. Apparent Sheet Density for seven paper brands, with brands highlighted in coloured circles.	210
Figure 9.6: Grammage and Apparent Sheet Density for seven replicates of papers measured from six reams of the same brand (Reflex).....	211
Figure 9.7: Grammage and mean measured thickness for seven replicates of papers measured from six reams of the same brand (Reflex).....	212
Figure 9.8: Comparison of grammage and apparent sheet density values between sample 71 and the selected background population.	213
Figure 10.1: Background population samples measured for carbon and oxygen isotopic values. Error bars denote the 0.5 and 0.7‰ benchmark discrimination powers previously defined.....	223
Figure 10.2: Background population samples measured for carbon and oxygen isotopic values, plotted by region of origin	224
Figure 10.3: Background population samples measured for carbon and oxygen isotopic values, with regions highlighted	224
Figure 10.4: Australian produced samples measured for carbon and oxygen isotopic abundance values.	226
Figure 10.5: Summary of discrimination powers for range of techniques for comparison of 60-sample paper set with paper sample 71	231
Figure 10.6: Proposed examination flow chart using a range of examination techniques	234
Figure 10.7: Interpretation factors for the reporting of paper examination techniques	239
Figure 11.1: Image of paper sample 11 from test F showing toner printing on page. ...	247
Figure 11.2: Light examination including ambient, transmitted and UV for trial A	250
Figure 11.3: Light examination including ambient, transmitted and UV for trial B	252
Figure 11.4: Light examination including ambient, transmitted and UV for trial C	254
Figure 11.5: Light examination including ambient, transmitted and UV for trial D	256
Figure 11.6: Light examination including ambient, transmitted and UV for trial E.....	258
Figure 11.7: Light examination including ambient, transmitted and UV for trial F....	260
Figure 11.8: Light examination including ambient, transmitted and UV for trial G	262
Figure 11.9: Light examination including ambient, transmitted and UV for trial H	266

Figure 11.10: Carbon and oxygen isotopic results for seven sheets in blind trial H.....267

Chapter 1

1. Literature Review and Aims

1.1. Introduction

Traditionally, Forensic Science is the application of scientific principles to the law, by means of solving questions for the courts and interpreting physical evidence left as a by-product of incidences of crime. More modern definitions of forensic science focus on the observation, collection and interpretation of traces. Material transferred during an event that can be used to decipher a source (who, what) or activity (what, how, when and why) (Margot, 2011).

The role of a forensic scientist is a challenging one, requiring the interpretation of traces left behind or transferred during spontaneous or abnormal events. The remnants left behind from these events require identification as being outside of the norm, before being classified as evidence and submitted for expert examination and interpretation. In contrast to this, the emergence of forensic intelligence has seen forensic information play a fast turn-around and proactive role in guiding investigations (Baechler et al., 2012, Ribaux et al., 2013).

As a discipline, Forensic Document Examination seeks to answer questions related to the production, authenticity and authorship of documents produced in support of crime. It is a broad discipline that covers: the examination of handwriting and signatures; the physical and chemical examination of papers and inks; printing processes and security documents; the decipherment of obliterated or altered entries on documents and; the visualisation of indentations.

The examination of paper has been an under utilised examination type within the document examination field, with writing and printing inks taking the majority of research focus particularly for chemical examinations. In casework, papers are only compared in certain circumstances and other examinations are typically completed

before paper comparisons are considered. The examination of paper however can result in significant outcomes for investigation or intelligence gathering in determining the source or association of documents. In terms of traces, this means that paper comparisons have the ability to deliver information based on who (e.g. questioned paper vs. suspect ream), what (e.g. page substitutions in multiple page documents) or where (e.g. the origin of paper used to create a document).

A number of analytical techniques have demonstrated the potential to discriminate papers from different sources however no technique has progressed to full forensic validation (Adams, 2011, Causin et al., 2011, Causin et al., 2010, Häkkänen et al., 2001, Kumar, 2011, McGaw et al., 2009b, Spence et al., 2000, Trejos et al., 2010, van Es et al., 2009). The most comprehensive paper comparison techniques are used for quality assurance purposes in paper mills (Murphy, 2009). Of these techniques, paper fibre microscopy offers the most information by visually identifying the species types and per cent fibre compositions used to create the sheet. Given the high technical demand of this technique however, it is not generally feasible for forensic laboratories to invest and maintain this capability.

Isotope Ratio Mass Spectrometry (IRMS) is a powerful analytical technique that has been used in a number of scientific disciplines, in particular geochemistry, ecology and climate studies. With respect to materials of forensic interest, it is a novel technique and its application is only just starting to be realised, with the number of research papers containing various forensic applications increasing (Gentile et al., 2015). IRMS is additive to other analytical methods in use in criminalistics laboratories and it has the ability to compare and individualise materials that have been identified as chemically identical using traditional analytical techniques. For papers, the technique has potential not just for comparison work but also for use as a tool that may indicate a papers' origin.

1.2. Aims of this Research

This research aims to introduce a new technique for the forensic comparison of office paper available within the Australian and New Zealand market. IRMS was selected as a technique due to the large amount of research that has been conducted in other

scientific disciplines such as ecology and the environmental sciences. Given that paper is organic in origin, this is a significant base of research on which to build. The technique also meets a number of casework requirements – it is semi-destructive, requiring only a small sample size to be removed, precise and generally observed to be discriminating when used for comparison.

The primary aim of this research project is to determine whether IRMS aids in the comparison and discrimination of paper. This will be done with reference to the future forensic casework applications of the technique, if shown to be discriminating. A number of scenarios, including questioned to questioned paper comparisons (where questioned is defined as a sample of unknown origin), questioned to specimen paper comparisons (where specimen is a suspect paper of known origin) or examination of a multiple page document (to determine whether a page substitution has occurred) will be considered.

To complement this, a number of other techniques will be tested to define the most appropriate examination methodology for document papers within the AFP laboratory. Including techniques such as the physical measurement and calculation of sheet grammage and apparent sheet density, alongside optical techniques and the chemical comparisons of paper dyes using Thin Layer Chromatography (TLC) will provide the expert with more information for interpretation of the results in addition to ensuring that non-destructive techniques are available for use in cases where destructive sampling is not permitted.

For all new techniques that result in evidence that will be presented in court, quality assurance standards must be met thus a significant portion of this work will be undertaken to demonstrate that the results produced are fit for purpose within internationally defined quality assurance guidelines; in addition to the gathering of background data which is imperative in the interpretation phase of court reporting.

1.3. Pulping and Paper Making

The invention of modern paper making is credited to the Chinese, who first produced a sheet from the bark of a mulberry tree that was treated with lime, bamboo and cloth

(Biermann, 1996). Modern paper is made from cellulose pulp derived from any number of species of trees, fillers such as clay or colourants, and additives that are used to impart any number of physical properties to the sheet. A general overview of the pulping and production process is shown in **Error! Reference source not found.**

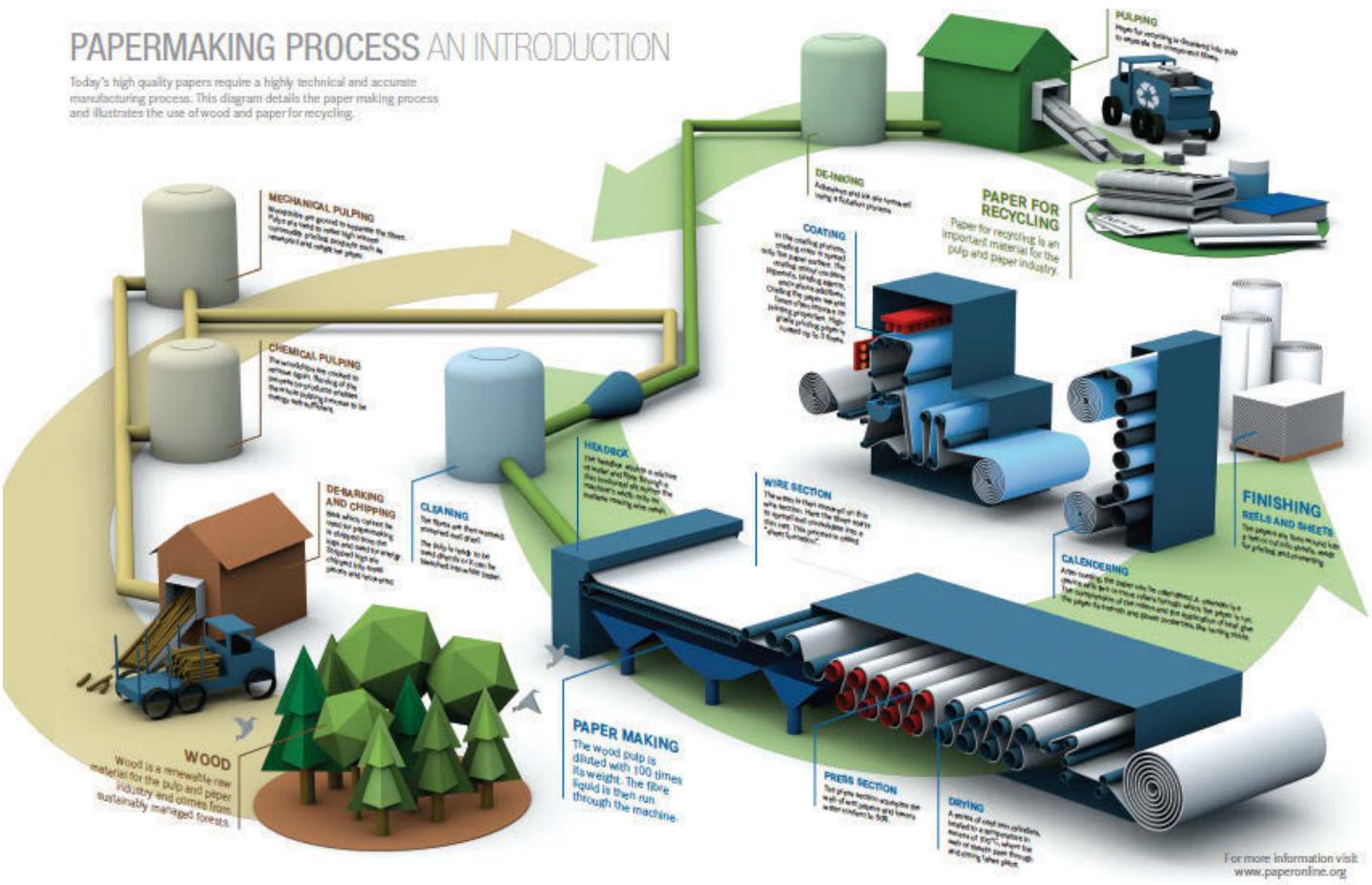


Figure 1.1: Schematic showing pulping and paper production process
(<http://www.paperonline.org/paper-making/paper-production/papermaking>)

1.3.1. Pulping

The overall aim of pulping is to reduce wood into its constituent fibres so that it may be reformed into a flat sheet. Chemically, wood is not a homogenous material however. Figure 1.2 shows the main chemical constituents of wood and their general component percentages. It should be noted here that for the purposes of papermaking and throughout the remainder of this thesis, hemicelluloses are included with discussion of celluloses.

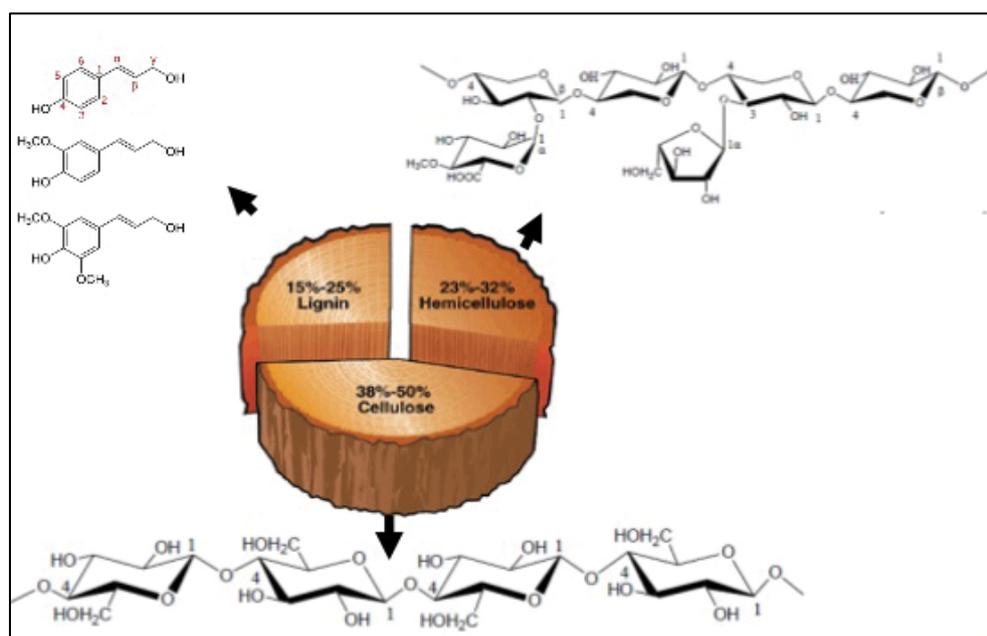


Figure 1.2: Chemical constituents of wood (<http://www.slideshare.net/SappiHouston/basics-of-wood-pulp-and-paper-november-2012>)

Generally, there are two pulping processes used to prepare raw wood, straw, cotton and other fibrous material for papermaking. The method used depends on the end use of the paper and the lifespan required for the product. Pulp for short use papers such as newspaper and brown paper packaging require less durability and are produced using a mechanical process. Mechanical pulping involves physically grinding the wood into fibres in the presence of water. The mechanical abrasion is achieved using stones with silicon carbide or aluminium embedded into them (Biermann, 1996).

Where longevity is important such as in office, printing and book papers, chemical processing is used to remove lignin – a natural glue like substance that causes a dark colour, brittleness and yellowing over time (Roberts, 1996a). Removal of lignin also

increases the bonding strength between fibres (Bierman, 1996). Chemical pulping uses one of two chemical processes to break down fibres and remove lignin:

Sulphite Processing

Salts of sulphuric acid (sulphite or bisulphates) are added to wood chips in large digesters and sealed under pressure. The pH of the mixture is adjusted using basic salts to achieve a pH between 1 and 1.5 (Roberts, 1996a). Sulphite processing is acidic and produces papers that are prone to degradation. It is also an expensive process that has largely been replaced due to environmental concerns.

Kraft Pulping

Sodium sulphide and sodium hydroxide are added to wood chips to produce fibres that are almost completely de-lignified. The process takes place in closed digesters that produce steam – which is used to power the process. The lignin and other contaminants are dissolved in the strongly basic mixture and the solid pulp is removed, bleached and washed (Roberts, 1996). Kraft pulping is now the most common pulping method for its efficiency, low cost and high quality product.

1.3.2. Pulp Bleaching and Whitening

Historically, pulp bleaching used chlorine or hypochlorite to whiten refined fibres. In response to environmental concerns, bleaching is now a multi-stage process that uses a number of chemicals to reduce harmful residues and pollution.

Depending on the grade and quality of paper to be used, bleaching (or more accurately, whitening) aims to remove all of the residual lignin (up to 5%) from the cooked chips, while creating a clean white fibrous mixture. The steps that can be included in the process are abbreviated for ease of description of methods. Table 1.1 shows some of the chemicals used, what the step achieves and the common abbreviation for that step (Suess, 2010).

Letter Designator	Chemical	Use
O	Oxygen	Molecular oxygen is used under alkaline conditions to oxidise lignin
D	Chlorine dioxide	Acidic solution of 5% ClO ₂ in water for lignin oxidation
E	Caustic soda	Extraction of oxidised lignin
P	Hydrogen peroxide	Alkaline bleaching stage at high temperature
OP	Hydrogen peroxide and oxygen	Alkaline bleaching with the addition of oxygen in a high pressure environment
Z	Gaseous oxygen	Delignification
Wash	De-ionised water	Wash step to remove residues

Table 1.1: Common steps involved in pulp whitening and de-lignification

1.3.3. Paper Making

Prior to entering the paper making machine, processed fibres are diluted with water down to 1% fibre content and combined with a number of chemical fillers and additives. Common types of additives, some common chemical forms and their purposes are shown in Table 1.2. A schematic of a typical paper machine is shown in Figure 1.3 (Kipphan, 2001).

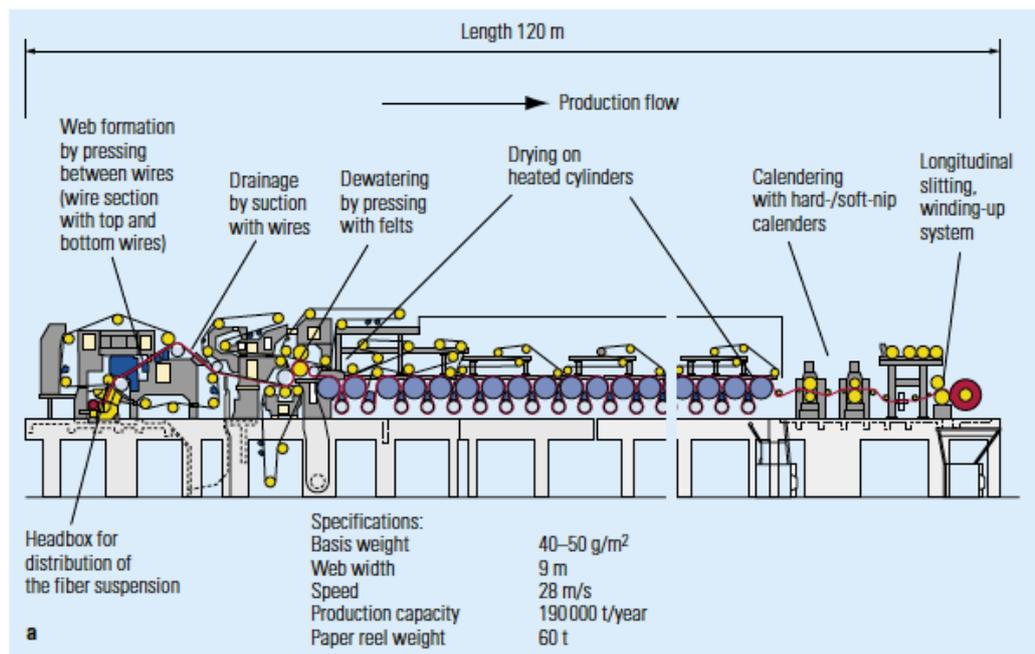


Figure 1.3: Schematic of typical paper making machine, showing key production stages (Kipphan, 2001).

Type of Additive	Chemical Forms	Purpose in Papermaking
Sizing agents	Alkenyl succinic anhydride, alkyl ketene dimers, starches, soda ash.	Surface agents that are applied once the paper is approx. 80% dry. These introduce hydrophobic groups onto the surface of cellulose, which impart water resistance and help to smooth the surface.
Bulking additives	Calcium carbonate, kaolin clay, chalk, limestone, talc, titanium dioxide, silica.	Improve brightness, opacity and smoothness. Reduce cost by lowering fibre content (can be up to 20 or 30% of paper weight).
Stabilisation and retention	Aluminium sulphate, polyacrylamides, cationic starches, polyethylamine.	Controls the stability of the suspension during process, retain fibres and fillers.
Fluorescent brighteners	Stilbenes – two benzene rings bonded by a carbon double bonded bridge.	Increase emission of violet light, producing an increase in whiteness and brightness.
Dyes	Red, yellow, blue and violet coloured dyes.	Used to adjust the opacity, hue and tint of papers.

Table 1.2: Common additives in the paper making process (Roberts, 1996b).

As shown in Figure 1.3, at the start of the paper making process a mixture containing the diluted refined fibres and additives is sprayed onto a moving wire (denoted the “fibre suspension”. The wire allows water to gravity feed away from the forming paper sheet. At the same time, vacuum units and a large flat roller remove more water before the sheet is picked up and fed onto a belt that is made of a fabric similar to felt. Felt covered rollers are pressed against the belt to further reduce water content in the sheet. After this process, the sheet is sufficiently internally bonded to hold its own weight and is fed through a long series of hot rollers that dry and press the paper flat.

After the paper is dried it is fed into the sizing press. Here, a mixture of starches, water resistance agents and surface smoothing agents are sprayed onto the surface of the sheet to control its external properties. The paper then moves through another set of drying rollers, followed by a calendaring section which smooths and polishes the sheet before it leaves the press (Biermann, 1996).

Following paper production, the formed sheet is spooled on to large rolls that typically hold 60 tonnes of paper (Kipphan, 2001). These reels are stored until the paper is cut and packaged into reams (packages of paper that typically contain 500 sheets, of a specified size). In large scale processing, cutting of sheets can occur from multiple reels at the same time. For example, while at the Australian Paper manufacturing facility in Maryvale, VIC, the cutting/packaging machine was observed to be cutting and stacking six reels of paper into a single ream simultaneously.

1.4. Forensic Examination of Paper

The examination of paper has long been overlooked by the document examination community in preference for other components of documents such as inks and toners. There are a number of potential reasons for this. Firstly, paper is assumed to be homogenous and lacking in the typical broad variation that may allow for discrimination. Secondly, paper examination is seen as a difficult undertaking (when done comprehensively) that requires a large amount of additional training and resources.

The most extensive paper characterisation techniques are undertaken by paper quality assurance laboratories that exist within manufacturing facilities. The purpose of these examinations however is characterisation of the paper for the purpose of maintaining product consistency and specifications. These examinations can include measuring the fluorescence, colour, tint, strength and density of the paper (Roberts, 1996b, Murphy, 2009). These laboratories often have experts who also specialise in identifying plant species microscopically from the fibrous content within the sheet.

1.4.1. Physical Examination Methods

Paper can be examined a number of ways non-destructively. The physical attributes of paper can be unique to the production process that created it however the true “uniqueness” of these features is limited due to mass production. The most simple paper examination techniques are to measure the physical dimensions of the page including the mass, length, width and weight. Calculation of the grammage and

apparent sheet density from these measurements can give an indication of whether two pieces of paper are of the same grade. These examinations are prescribed in International Standards ISO 536:1995 and ISO 534:1988 and Australian/New Zealand Standards 1301.405s:2004 and 1301.426s:1994 respectively (published online and accessible through Sai Global (2016)).

Forensically, a number of techniques have been trialled, generally on small sample sizes, for the comparison of questioned document papers. One of these investigations, by Kumar (2011), aimed to utilise tensile strength testing. Four samples of paper were compared by removing strips from the paper and measuring the load they could hold and their strain peak, defined by the author as “the strain withstood by the sample just before rupture”. Although only a small sample was tested, the author concluded that with the addition of spectral techniques these tests were quite effective in discriminating papers.

Spectral techniques, such as the examination of paper with ultraviolet (UV) light have been used by a number of researchers. Originally, UV light was used to observe differences in fluorescence between processed cellulose tree and cotton fibres (Hilton, 1949). In more modern paper the test gives less meaningful results given the widespread inclusion of fluorescent optical brighteners and pigments in paper (Grant, 1973). An in-depth study published by Green (2012) showed that up to 30% of samples tested had differences in the fluorescent properties of sheets within the same ream, which would be due to cutting of multiple reels into the same ream of paper.

Causin et al. (2011) examined 20 paper samples using diffuse-reflectance ultraviolet-visible-near infrared spectrophotometry that could not be visually discriminated. Using this technique in several different configurations, including examining the paper using reflectivity mode between 680 and 900nm, paper samples were found to be both homogenous when measured across the same side and discriminable between samples.

Grain direction, wire marks and edge striations are other characteristics that can be used to compare and discriminate samples. These marks are created on a sheet during the papermaking and sheet cutting processes.

Grain direction in the fibres of paper is formed due to vibration of the papermaking machine, with the fibres aligning in the direction of movement of the web. This is shown pictorially in Figure 1.4 (Kipphan, 2001). This becomes a point of differentiation in cases where sheets are cut in either the machine direction (with fibres aligning to the long edge of the sheet, shown in Figure 1.4 as long grain) or the cross direction (fibres aligning to the short edge of the sheet, shown in Figure 1.4 as short grain). The grain direction is determined by visual examination, sometimes with the aid of shallow angle (oblique) light. Marchand (1989) found that this type of lighting was the most effective, producing correct results in all 100 tests by a skilled examiner.

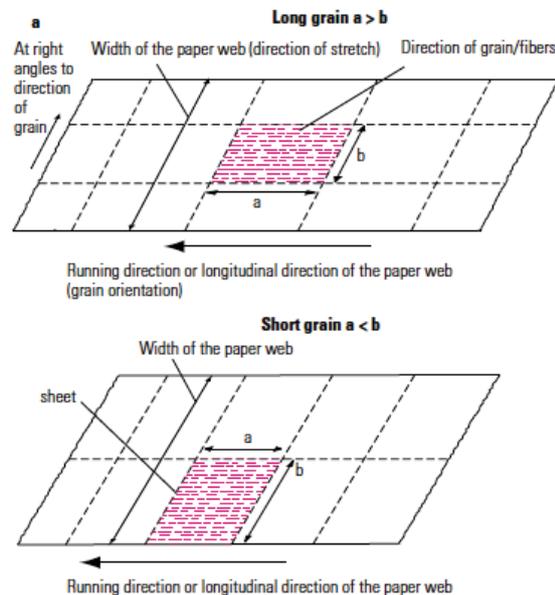


Figure 1.4: Schematic demonstrating grain direction in papers (Kipphan, 2001)

The wire of the web of the paper machine also imparts an impression onto the sheet as it forms. Wire marks will vary depending on the mill and the frequency of replacement of the wire within the mill. The mark is not visible with the naked eye and is most commonly visualised using Fourier Transform processing. Holland (2004) used a high resolution digital camera, while Miyata et al (2002) and Berger (2009) both used transmitted light and a flatbed scanner to record the wire pattern. The results achieved are abstract representations of the frequency of repetitive features. While Miyata et al completed a larger study on the homogeneity of paper, all three studies showed the usefulness of using a Fourier Transform method for visualisation of this repetitive feature. Berger and Ramos (2012) have more recently published a new algorithm to

compute the Fourier Transformed image, which will search against a larger database of known samples.

Edge striations are created when stacks of sheets of paper are guillotined to size. The tool mark created can be unique to the guillotine or blade that created it and is a method of comparing papers that may have come from the same source (Bodziak, 1998). In contrast to wire marks, the images of the striations produced need no data processing and can be compared similarly to the manner in which a bullet casing would be compared (Cain, 1984).

Special Features

Security features are intentionally incorporated into specialist papers during production as protection against counterfeiting and alteration. Generally the paper used will be of a different composition to standard office papers and will contain a large proportion or 100% cotton fibres. Some of the security features that can be incorporated into these papers include watermarks (images impressed into the forming sheet), planchettes and fibres (extraneous materials added into the paper mixture) and security threads (long strips of plastic embedded into the sheet). Security papers are generally free from optical brighteners and will not fluoresce during UV light examination.

1.4.2. Chemical Examination Methods

A number of articles discussed in this section show promise in terms of their comparative abilities. However, prior to this research, none of these studies can be considered as validation studies for forensic casework purposes. This is especially true in Australia, where examination of paper (if conducted at all) is primarily of a physical nature.

Measurement of the elemental composition of paper utilising various techniques, comprises the bulk of paper research conducted for forensic comparison. Techniques such as Neutron Activation Analysis (NAA), Inductively Coupled Plasma Mass Spectrometry (ICP-MS), X-Ray Fluorescence (XRF), Laser Induced Plasma Spectroscopy (LIPS) and Laser Induced Breakdown Spectroscopy (LIBS) have all

been used. The target elements measured vary, although all are expected to have been introduced into paper as contamination of the components used in the papermaking process (Manso et al., 2008).

Schlesinger and Settle (1971) discuss the use of Neutron Activation Analysis in the determination of the elemental compositions of papers. This study tested a large number of bond papers and found that there was inhomogeneity between brands that was traceable to the source. Blanchard (1978) also used Neutron Activation Analysis for higher quality papers. The samples used in this study were different to the samples used by Schlesinger and Settle however as only the clay component of the papers was analysed. The clays used were procured from paper mills and after analysing the bulk clays, papers were prepared by hand for further measurement. This research does not present a modern scenario however due to the samples measured and is hard to reconcile with casework given the high level of automation and engineering on modern papermaking.

Inductively Coupled Plasma Mass Spectrometry (ICP-MS) has been successfully used to compare and discriminate trace elements in paper. Spence et al. (2000) used microwave digestion prior to ICP-MS to look at 17 different papers from 10 countries. All 17 samples could be discriminated and further to this, four samples from one mill produced over 4 consecutive months showed differences when 3-4 elements were compared. This technique was used in a casework situation (Spence et al., 2002) where a threatening letter was differentiated from paper seized from a suspects premises. McGaw et al. (2009b) also successfully used the technique on a small number of 100% recycled papers with some success. The recycled paper in this study would be expected to be much more inhomogeneous (and therefore better able to be discriminated using such a technique) than other virgin pulp papers given the varied fibre sources utilised in production. Additionally, no homogeneity experiments have been undertaken within these studies so there is no experimentally defined view of the natural range of variation expected within one source.

Laser Induced Plasma Spectrometry (LIPS) was used by Hakkanen (2001) to measure the elemental composition of papers that had multiple coatings. Using the laser to ablate thickness sections, the two coats on the surface of the papers could only be

detected if there was an elemental difference between them. Similarly, Laser Induced Breakdown Spectroscopy (LIBS) was used by Trejos et al. (2010) and coupled ICP-MS using Laser Ablation (LA-ICP-MS) rather than the digestion methods used by McGraw and Spence. Using both techniques, greater than 98% discrimination was achieved for 17 papers. Using these two methods, weekly and monthly batches were also discriminated. Further work by Lennard et al. (2015) utilised LIBS on a sample set of 33 Australian papers taken from this project. Discrimination was observed on single sheets from each sample, with 98% discrimination observed within a 2 standard deviation comparison.

Throughout these elemental studies, and particularly given some of the results presented throughout this project, homogeneity and variability for the trace chemical composition of document papers has not yet adequately been addressed. There is further work to be done undertaken with laser ablation techniques to elucidate how much environmental handling, printing and usage would affect the trace elements measured on the surface of the sheets. Given that both in Trejos et al. (2010) and in Lennard et al. (2015) no laser depth experiments were published, it would be useful to investigate whether the surface of the sheet could be ablated and measurements taken from the centre of the sheet to negate the possibility of environmental contamination. Both variability and environmental contamination should be a consideration during interpretation and discrimination, as even in the clean sheets removed straight from the ream, contamination could be present from the cutting and packaging processes.

Energy Dispersive X-Ray Fluorescence (ED-XRF) was used to measure a range of papers (from ancient parchments to modern papyrus and newspaper) by Manso et al. (2007) finding that modern papers contained a lower elemental composition than ancient papers. This was also seen in newspapers and the author suggests that the observation is due to greater quality control of modern papermaking. Rozic et al. (2005) used ED-XRF to measure a range of paper samples. The samples were combusted at 900°C prior to measurement to remove the bulk organic material. Differences were seen between the raw paper samples and samples that had been printed on using inkjet or toner processes, indicating that usage effects, so often overlooked, need to be considered when proving the usefulness of any technique.

McGaw et al. (2009a) used the undigested portion of the paper from the previous ICP-MS study (McGaw et al., 2009b) to characterise the type and composition of the fillers and clays using Fourier Transform Infrared Spectroscopy (FTIR) and X-Ray Diffraction (XRD). Bulk paper has also been characterised using FTIR by Andrasko (1996), with all 10 office papers tested able to be discriminated.

Similarly, XRD was used by Foner and Adan (1983) to characterise the cellulose crystallinity and filler content of a range of papers. The technique was found to be useful in identifying the type of filler and in giving a rough indication of the filler concentration. Likewise, some papers were able to be discriminated based on the different crystalline structures of the cellulose used to make the paper. Causin et al. (2010) used the spectra from FTIR to calculate a peak area ratio for comparison of 19 paper samples. In combination with XRD, all 19 samples were discriminated however peak area determination in FTIR is not sufficient to conclusively include or exclude common source.

Pyrolysis Gas Chromatography has been used in three main studies which aim to compare specialist photographic paper (Washall and Wampler, 1989), polymers in coated papers (Crockett et al., 1987) and in plain and printed officer papers (Wampler and Levy, 1986). All three studies show some potential for the technique however are studies that have only included small sample sizes or been conducted on very specific materials.

Direct Analysis in Real Time Mass Spectrometry (DART-MS) has been used in a conservation study to identify the type of pulping process used by detecting different ions of interest that are present in the sample due to the high-temperature thermal processes used to break down formed wood into cellulose fibres. This interesting study has potential for forensic applications in comparison of papers that have been produced in different mills. DART-MS as a technique is also able to analyse the samples without pre-processing, with samples used in this study taken directly from the paper sheet measuring 6.4mm in diameter.

The composition of papers is also able to be determined using a series of simple spot tests, including tests for the presence of calcium carbonate using dilute hydrochloric

acid, for rosin sizing using sulfuric acid and sugar, and using an iodine solution for detecting the presence of starch (Brunelle and Reed, 1984). Ziderman (1981) used the paper sample itself as a chromatographic sorbent for characterisation and comparison.

While all of these techniques have shown some positive results in terms of their abilities to compare and discriminate papers, none of the published studies have moved beyond proof of concept to the level necessary for use in casework in an accredited laboratory. For the elemental techniques and those utilising laser ablation in particular, there is also the issue of environmental contamination, which could change or mask the true elemental composition of the paper through general use and handling. The limitation here is only assessing these techniques using clean papers sourced directly from purchased reams. Overall, given these limitations in addition to no ream homogeneity data, a lack of blind trials and method validation, it is difficult to assess the results of these publications in terms of their potential for use in forensic casework comparisons.

1.5. Isotope Ratio Mass Spectrometry

1.5.1. Stable Isotopes

Isotopes are defined as nuclides of the same element, differing only in the number of neutrons present in the nucleus. This produces a mass number difference between atoms of the same species and molecules of the same configuration. There are two types of isotopic species – radioactive isotopes (those that undergo decay) and stable isotopes (those that have not been observed to undergo decay) (Criss, 1999).

For those elements with stable isotopes, there are at least two types of stable isotopic species – the light, dominant or major isotope; and the heavy, rare or tracer isotope. Some elements are monoisotopic, with only one stable isotopic species while some, such as lead, have up to 10 stable isotopes. The most commonly measured light stable isotopes are the isotopes of Hydrogen, Carbon, Nitrogen, Oxygen and Sulfur. The natural relative abundances of these isotopes are given in Table 1.3 (De Groot, 2008).

Element	Isotope	Average Relative Abundance (%)
Hydrogen (H)	^1H	99.985
	^2H (also known as Deuterium or D)	0.015
Carbon	^{12}C	98.89
	^{13}C	1.11
Nitrogen	^{14}N	99.64
	^{15}N	0.36
Oxygen	^{16}O	99.763
	^{17}O	0.0375
	^{18}O	0.1995
Sulfur	^{32}S	95.02
	^{33}S	0.75
	^{34}S	4.21
	^{36}S	0.02

Table 1.3: The relative natural abundances of commonly measured stable isotopes (De Groot, 2008)

1.5.2. Isotope Ratio Notation and International Standards

Due to the low abundance of most heavy isotopes, measuring the absolute isotopic concentration in any material becomes difficult with respect to achieving precision. Instead, stable isotope measurements calculated as the relative difference between a standard material with known value and the unknown samples measured within the same experimental run. The isotope ratio, an expression of the isotopic abundance of a sample, is given by (Coplen, 2011):

$$\delta^h E_{std} \text{‰} = \left(\frac{R_{Sample}}{R_{Std}} \right) - 1$$

Where ‘h’ refers to the heavy isotope, ‘E’ refers to the element being measured, ‘std’ is the international standard scale being referenced against and ‘R’ is the ratio of the light and heavy isotope. An example of this notation, as utilised throughout this text, for carbon is $\delta^{13}\text{C}_{\text{VPDB}}$.

Delta values (δ) are expressed in parts per thousand or per mil (‰), which is a unitless number. The correct notation for expression of delta values is δ . For ease of comparison, international standards are arbitrarily set to 0 ‰ and thus experimental results are expressed as the ratio of heavy to light isotopes relative to the standard used. In practical terms, this means that a negative delta value represents depletion in the heavy isotope, while positive values are said to be enriched in the heavy isotope, relative to the standard that the sample is being scaled against.

Using delta values is an effective method of expressing isotope ratio results where by very small differences are more apparent and easier to interpret with respect to their significance. A paper published by Coplen (2011) outlines a range of description and calculation standards for working with isotopic results, recommended by the Commission on the Isotopic Abundance and Atomic Weights of the International Union of Pure and Applied Chemistry. This paper is now the standard for all results published containing isotope ratio measurement results.

The International Atomic Energy Agency (IAEA) and the National Institute of Standards and Technology (NIST) both supply a range of natural and non-natural isotopic standards. The international standards produced by these agencies have values that are traceable back to a master scale or material (those that are set to 0 ‰). For comparability, all reported isotopic abundance values must be traceable back to these scales. Commonly referenced international standard scales and the stable isotopes they reference are shown in Table 1.4 (Werner and Brand, 2001, De Groot, 2004).

Element	Standard Name
O, H	Vienna Standard Mean Ocean Water (VSMOW)
C	Vienna Pee Dee Belemnite (VPDB)
N	Atmospheric Nitrogen (AIR)

Table 1.4: International Isotopic Reference Materials

International standard materials with published values are in limited supply and are only able to be purchased by a laboratory from the IAEA once every three years. For high-throughput laboratories, this means that working standards should be developed

for routine measurement. Guidelines for selection of materials for use as laboratory standards have been discussed by Kipphardt (De Groot, 2004). Additionally, forensic laboratories that are accredited to international standards must ensure that any reference materials used meet international standard guidelines for their laboratory (e.g. ISO/IEC 17025:2005 (ISO, 2005)). Guidelines have been produced to assist in this process such as those set out in ISO Guide 34: 2009 (ISO, 2009).

Werner and Brand (2001) discuss the criteria for selection and calibration of materials as laboratory standards. These criteria include:

- Similarity of chemical composition to the sample material
- Bracketing of the expected isotope ratios of the sample material
- Homogeneity
- Stability (no fractionation during storage, preparation or measurement)
- Availability
- Similarity of decomposition during pyrolysis and/or combustion
- Ease of use

While this list was intended for calibration of new materials, it is equally appropriate as a general guide to the selection of isotopic reference materials, particularly this respect to the first two criteria.

1.6. Isotopic Scale Calibration and Correction

1.6.1 Method for Correcting Raw Instrument Data

Prior to being exported to Microsoft Excel[®], the isotopic ratio values obtained from the IRMS software have undergone a series of calibrations and corrections which vary depending on the isotope being measured. Most fundamental to this is “single point anchoring” of the measured sample peak heights against a defined value for a pure reference gas.

In continuous flow measurements, there are two options for single point anchoring of the measured values using reference gas – the isotopic ratio of the working gas tank

can be calibrated or, as is the practice utilized within the AFP laboratory, the working gas value is arbitrarily set to 0 ‰. The primary advantage of using a 0 ‰ reference gas value avoids any mass dependent effects from the progressive emptying of the tank, with the values expected to shift over the life of the tank as the heavy isotope is more likely to be retained in the tank longer, shifting the isotopic ratio of the reference gas over time.

Other corrections that take place within the IRMS software include:

- ^{17}O correction for oxygen and carbon measurements - ^{17}O correction is used to account for the contribution of ^{17}O to the isotopologues of CO_2 and CO by assuming that their contribution (species containing $^{17}\text{O}/^{16}\text{O}$) is half that of the heavier isotope (i.e. that $^{17}\text{O}/^{16}\text{O}$ contributions are half that of $^{18}\text{O}/^{16}\text{O}$). The correction is applied through scaling of the three masses measured (m/z 45, 46 and 47) using a correction calculation developed by (Craig, 1957). The contribution of ^{17}O is much smaller in the measurement of CO but is corrected within the software in a similar fashion (Carter, 2011).
- H_3^+ factor for hydrogen measurements - H_3^+ is created within the ion source at increasing gas pressures through the ionization of H_2^+ with H_2 (creating H_3^+ and H). The number of H_3^+ ions formed is proportional to the gas concentration in the ion source. The factor is determined using an on/off reference gas sequence, increasing the gas pressure, and scaling the m/z 2 intensity against the m/z 3 intensity across the different pressures. The H_3^+ factor is then applied to each H peak measured (Carter, 2011).

1.6.2 Drift Correction

During long analytical sequences it is possible for the values produced by the instrument to shift over time due to changes in the isotopic content of the reference gases or small changes in the ion source. To correct for this, standard materials should be placed at the start and the end of each analytical sequence so that their values can be monitored and compared. Using the mean value for each replicate group (noting that any outlying values should be identified and removed prior) an absolute difference should be calculated between the two replicate sets of standards at the start and the end of the sequence. Where the difference is larger than the published analytical error of

the instrument (published in FIRMS Good Practice Guideline, adopted from Thermo Fisher as 3 ‰ for $\delta^2\text{H}$, 0.5 ‰ for $\delta^{18}\text{O}$, and 0.3 ‰ for $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$) drift correction should be applied to the entire analytical sequence (Carter, 2011).

If drift correction is required, the difference between the two sets of standards (*mean drift*) is used to correct each δ -value (*corrected value*) in the sequence using the instrument derived value (*raw value*) and both the analytical sequence length (*# samples total*) and the position of the sample δ -value being corrected within that sequence (*position #*), using the following equation:

$$\text{Corrected Value} = \text{Raw Value} \pm \left(\left(\frac{\text{mean drift}}{\text{\# samples total}} \right) \times \text{position \#} \right)$$

This correction should be applied prior to isotopic scale calibration.

1.6.3 Outlier Exclusion using Grubbs Test

The internal materials and reactors of both the Flash EA and TC/EA can become contaminated with previous samples, leading to the introduction of memory effects. Likewise, when analysing materials with widely different δ -values, the detectors of the IRMS can take a number of samples to accurately record the change and stabilise. This is most prominently seen in measurements of hydrogen, particularly when running water standards that vary widely in their $\delta^2\text{H}$ values (e.g. SLAP2 and VSMOW2 have a difference of -428 ‰).

Grubbs test is a common test for outlier identification within a replicate group. The formula is given as (Hibbert and Gooding, 2006):

$$\text{Outlier Statistic (G)} = \frac{|\text{measured value} - \text{group mean}|}{\text{standard deviation}}$$

The G value obtained for each sample can then compared to the G_{critical} value used in Grubbs test, which varies depending on replicate number and the confidence required (see page 78 of Hibbert and Gooding).

Grubbs test is ideally utilised to remove a single outlier only. As a normal distribution is assumed, which can confidently be transposed on to standard materials but cannot be assumed for document papers, care was taken not to exclude paper values unless the G value was large and the difference could reasonably be attributed to instrument shift or random error.

1.6.4 Isotopic Scale Calibration

Isotopic scale calibration (also termed “normalisation”) is used to correct raw/measured isotopic abundance values of a sample to the “true” value that is reported against an internationally traceable standard material. While some correction to the raw δ -value occurs within the instrument through referencing against the reference gases run in each sample analysis, manual processing of the raw values is required, generally in Microsoft Excel[®] or similar.

A number of conditions must have been met in the experimental design and results obtained for scale calibration to be successful:

1. Good chromatography of the analytical sequence including efficient peak elution and separation, and low background noise.
2. Using the two sets of standards, placed at the start and the end of the analytical sequence for assessment of drift, with drift correction applied to the raw results if required
3. The standards selected bracket the unknown sample values
4. A quality control (QC) material, with a published or laboratory defined value is included in the sequence to provide a check of the scale calibration.

For continuous flow instruments, isotopic scale calibration using two-point linear regression normalisation is undertaken for each analytical sequence produced. This process can also be termed “linear shift normalization”, “multipoint normalization” or “stretch-shift correction”.

A calibration plot is obtained, where the measured value for the standards are plotted against the published (or laboratory calibrated) value for the material. An example of a calibration plot obtained for standard materials IAEA-CH-7 (Polyethylene) and IAEA-CH-6 (Sucrose) is shown for $\delta^{13}\text{C}_{\text{VPDB}}$ in Figure 1.5.

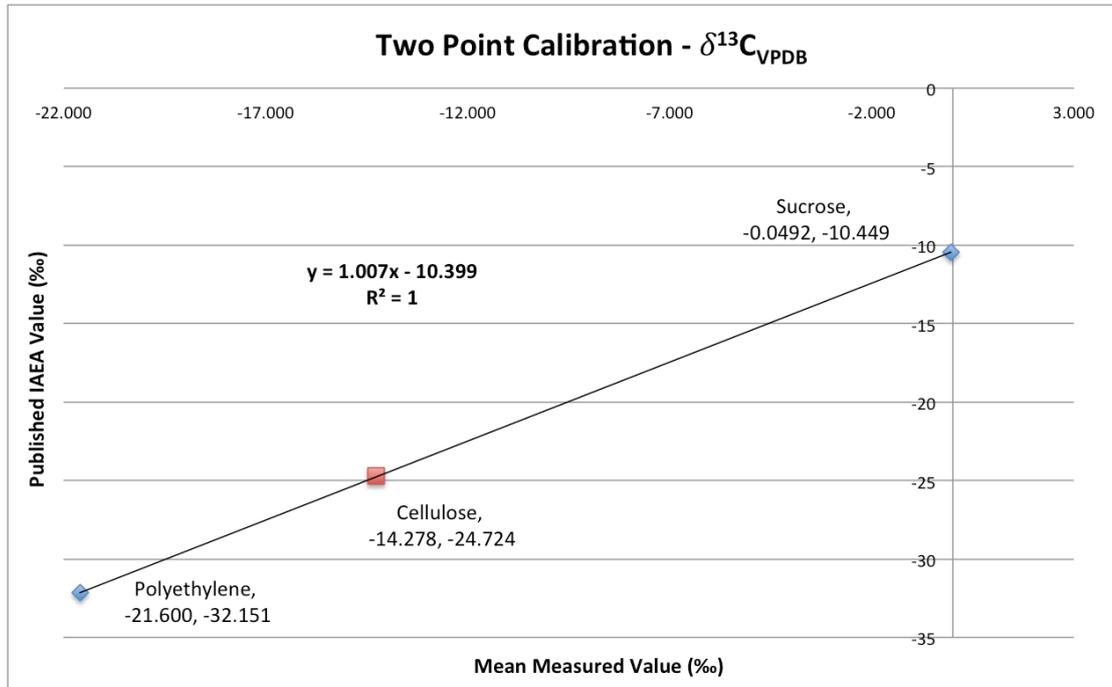


Figure 1.5: Two-point calibration plot for $\delta^{13}\text{C}_{\text{VPDB}}$ (%) using polyethylene and sucrose standards

There are two factors being determined in the calibration – the stretch factor and the shift factor. These two factors are combined using a linear regression ($y=mx + b$) in the equation of the line of best fit between the two standards such that:

$$\text{Corrected} = \text{stretch factor} * \text{raw value} + \text{shift factor}$$

The stretch factor is used to ensure that the difference between the two standards used is retained. So, using the example in Figure 1.5, the difference between the published values of Polyethylene and Sucrose is 21.702 ‰. The measured difference between the same standards is 21.551 ‰, slightly smaller than the true/published difference between the two materials. To account for this, a ratio is taken through the use of linear regression, to stretch the measured scale to the required difference. In this case, the ratio of true difference/measured difference (given by the m value of the line) is 1.007.

The shift factor is used to move the position of the line to bring it into agreement with the scale being used. For carbon this is the VPDB scale, for oxygen it is VSMOW. This positional difference is given by the b value in the linear regression equation.

When three or more standard materials are used in a calibration, the R^2 value is used to understand how well the data fit the regression model. This can be used to monitor the behaviour of the instrument and identify errors. Conversely, when a two-point calibration is utilised, it is essential to include a quality assurance material to monitor the performance of the instrument and the calibration. This is shown in Figure 1.5 with international standard cellulose (IAEA-CH-3) being used as a calibration check. While a standard material with a published value has been used here, it is equally valid to utilize an alternative material whose laboratory value has been defined and its homogeneity demonstrated. Ideally, this material should be matrix matched to the unknown samples being measured.

Once the two-point normalization chart has been constructed and the QC material shown to sit on the calibration line, correction of unknown samples can proceed utilizing the regression equation.

1.7. Isotopic Fractionation

Fractionation is defined as any chemical or physical reaction or process that shows a bias between the heavy or light isotopes of a compound (Meier-Augenstein, 2010). Fractionation occurs due to differences in the vibrational frequencies of the bonds between light and heavy isotope containing molecules. Molecules containing heavy isotopes are more tightly bonded and have lower vibrational frequencies than lighter isotopes, creating a difference in reaction rates during chemical reactions. Overall, heavier isotopes have less free energy available for reaction and hence require more energy to break a bond, making them react slower than light isotopes (Fry, 2006). Selective enrichment and/or depletion of materials can occur through a variety of equilibrium and kinetic fractionation processes.

1.7.1. Equilibrium Isotope Effects

Equilibrium isotope effects are those that occur due to differences in the physicochemical properties of molecules. Some examples of these kinds of properties include infrared absorption, molar volume and boiling and melting points (Meier-Augenstein, 2010). Equilibrium effects also include equilibrium exchanges - when the exchange between two different phases or product pools is equal (Fry, 2006). Equilibrium exchanges occur in reactions such as the transformation of water vapour to liquid precipitation (Meier-Augenstein, 2010).

Equilibrium fractionation is based on the bond strength differences (due to mass differences) when two phases are exchanging isotopic molecules (Yinon, 2004, Sharp, 2007). This is demonstrated in Figure 1.6, which shows the different potential energy states for diatomic hydrogen. In this diagram, the mass differences observed between $^1\text{H}-^1\text{H}$, $^1\text{H}-^2\text{H}$ (H-D) and $^2\text{H}-^2\text{H}$ (D-D), result in an increase in potential energy and hence an increase in bond strength (Sharp, 2007). The consequence of this is that in any equilibrium reaction, the heavier isotope will form a stronger bond, be slower to change phases and overall partition (or remain) in the phase containing the higher bond energy. This follows three general rules (Schauble, 2004 in (Sharp, 2007):

1. Equilibrium fractionation between two phases generally decreases with increasing temperature
2. The degree of fractionation is generally larger for isotopes with a larger mass ratio difference
3. The heavy isotope is preferentially portioned into the phase containing the highest bond energy.

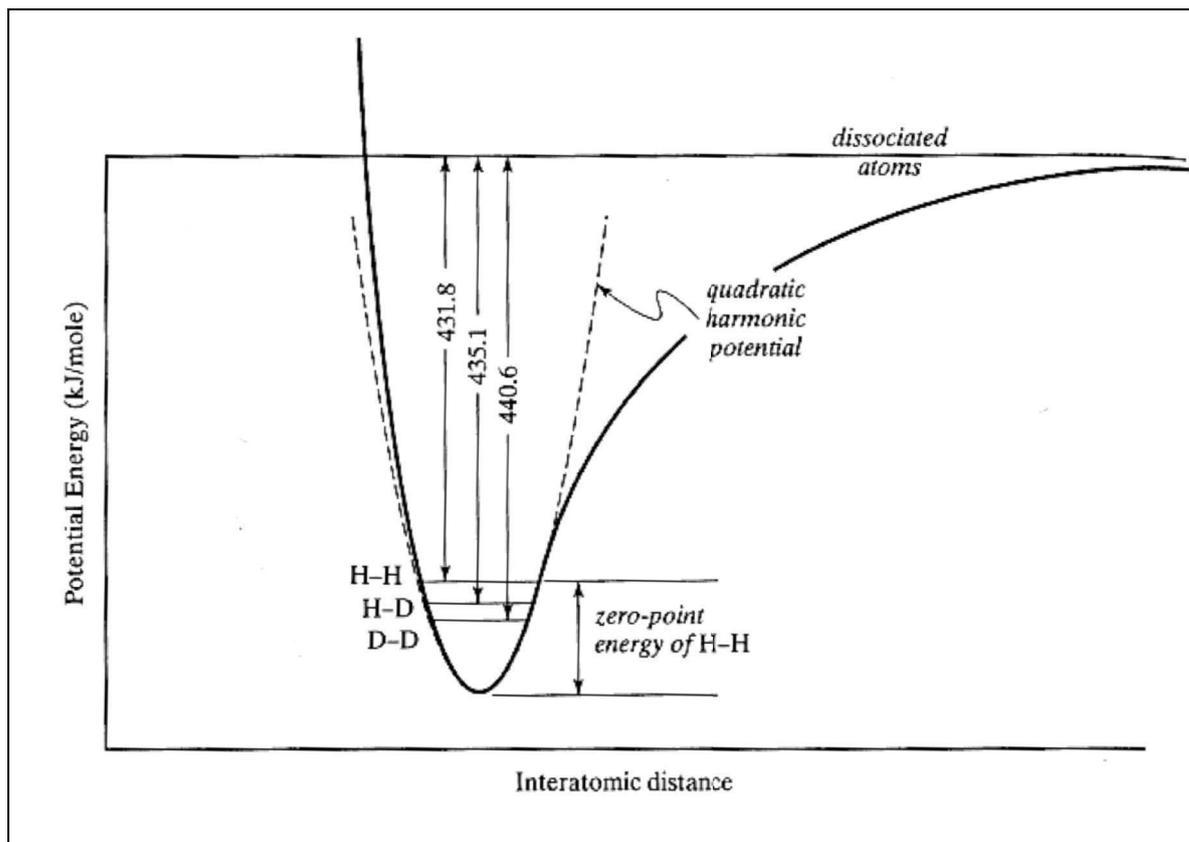


Figure 1.6: Bond energy differences in molecules containing ^2H (annotated here as D) (Sharp, 2007)

An example of equilibrium fractionation is observed in the uptake of carbon dioxide from the atmosphere by plants utilising C_3 photosynthetic pathways (see further discussion in section 1.10). Figure 1.7 (Fry, 2006) demonstrates carbon fractionation during uptake from the atmosphere (-8‰), across a diffusion membrane with a fractionation of 4.4‰ and secondary (kinetic) fractionation inside the leaf stomata to fixed carbon with a further fractionation of 29‰ . Calculation of fractionation factors is discussed in section 1.7.3.

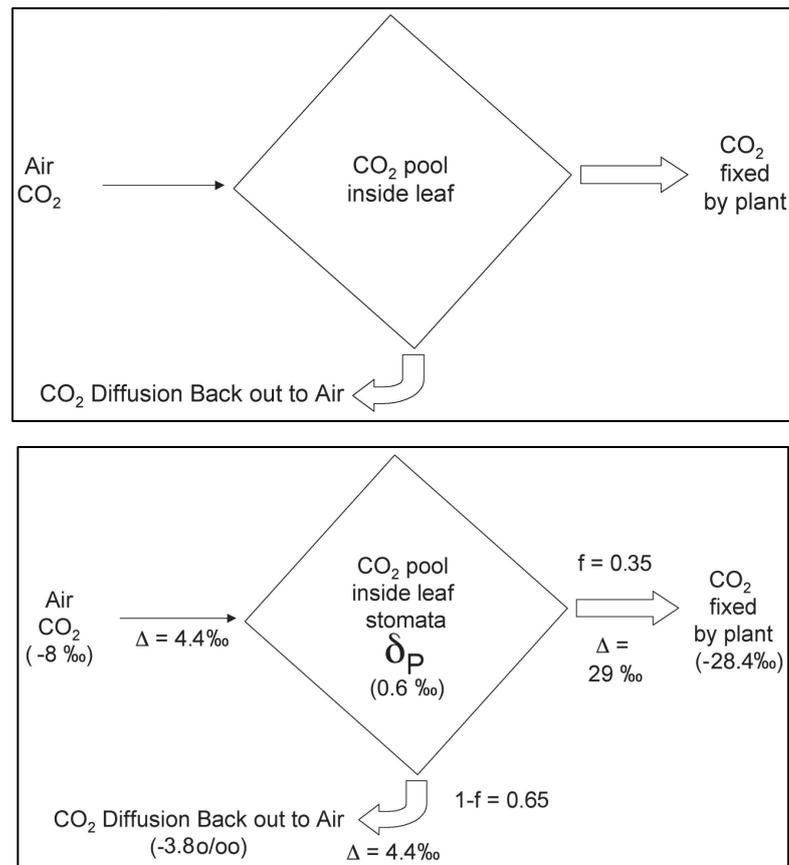


Figure 1.7: Carbon equilibrium fractionation during C_3 photosynthesis, top showing general pathway of movement of CO_2 , bottom showing fractionation factors (Δ) based on current atmospheric $\delta^{13}C$ value of -8 ‰ (Fry, 2006).

1.7.2. Kinetic Isotope Effects

Kinetic isotope effects are those that take place in reactions that are irreversible. They again are based on differences in the reaction rates created by tighter bond energies in heavy isotope containing molecules. This causes permanent fractionation and a net depletion of the heavy isotope within the product of the reaction (Dawson and Siegwolf, 2007). Kinetic effects occur because of differences in the diffusion of molecules across a surface, diffusion of molecules through a substance and due to enzyme catalysed reactions (Hoefs, 2008). The magnitude of the fractionation is affected by the pressure, temperature and the availability of reactants (i.e. whether the system is open or closed) in any reaction. An example of a kinetic isotope effect is fixation of CO_2 in photosynthesis, which in an open system with continuous input of

atmospheric CO₂ causes depletion of ¹³C in the products of photosynthesis due to differences in the reaction rates with photosynthetic enzymes (Ehleringer, 1993).

1.7.3. Calculation of Isotopic Fractionation

The fractionation that occurs during a reaction can be calculated using the following equation (Fry, 2006):

$$\Delta = \delta_{\text{source}} - \delta_{\text{product}}$$

The fractionation factor for a material (α) is defined as the ratio of rate constants where (L) denotes the light isotope, (H) is the heavy isotope and (k) is the reaction rate constant (Fry, 2006). This forms the equation:

$$\alpha = \frac{k_L}{k_H}$$

The value derived from this calculation is used to describe the isotopic preference of a molecule in a given reaction. The differences are small – in the low percent values – and are commonly scaled up to the per mil notation so that the values are comparable to the scale used for measurement. This is done in one of two ways, depending on the subject matter being studied.

For chemists and biologists Δ (“big delta”) is commonly used where the following equation gives a positive δ value and expresses the amount of fractionation in a particular molecule:

$$\Delta = (\alpha - 1) * 1000$$

In geochemistry and hydrology, the equation for α is reversed, giving:

$$\alpha = \frac{k_H}{k_L}$$

and fractionation is expressed using epsilon (ϵ), which is calculated using the same equation as for Δ :

$$\varepsilon = (\alpha - 1) * 1000$$

The difference when using ε is that the value obtained is a negative per mil number. The terms are interchangeable and in most respects, result only in a change in terminology. For further information on the fractionation calculations used in isotopic studies, chapter 46 of De Groot (2004), Fry (2006) and Dawson and Siegwolf (2007) discuss in more detail.

1.8. Stable Isotope Instrumentation

1.8.1. Isotope Ratio Mass Spectrometer

A mass spectrometer ionises molecules and separates them based on their mass to charge (m/z) ratio. This separation is achieved by accelerating the charged molecules through a magnetic field. Due to differences in the mass of ions, different path lengths are formed within the flight tube, causing the molecules to finish at different points at the end of the tube (Harris, 2003).

In structural mass spectrometry, there is one detector that scans and detects a range of masses present in a sample. In an IRMS, multiple detectors simultaneously collect only one m/z ratio each. The importance of this is two-fold – firstly the high sensitivity of these detectors allows for the collection of low abundance levels within a sample i.e. accurate recording of the heavy isotope. Secondly, without collecting a full complement of both the light and the heavy isotope, a true ratio could not be calculated due to time vs. signal effects- the most predominant being later presentation of the heavy isotope due to a slightly heavier mass (Yinon, 2004).

The 4 key components of an isotope ratio mass spectrometer are (as shown in Figure 1.8):

1. The sample preparation method and inlet system – where the sample enters the mass spectrometer either using continuous flow or dual inlet methods
2. The ion source – where the sample is ionised, focussed and accelerated

3. The mass analyser – which separates ions leaving the ion source based on their m/z ratio as they pass through a magnetic field
4. Detection – in an IRMS instrument, between 3 and 7 detectors (typically five, as shown in Figure 1.8) called Faraday collectors are used. As an ion hits the detector, an electrical impulse is created which is amplified. To increase sensitivity and to reduce noise, each faraday collector is mounted in a set position and contains its own amplification and counting devices. Overall, this ensures that the very small amounts of heavy isotope are detected while in the presence of a much larger light isotopic signal.

Detailed information on the set-up and operation of an IRMS is provided in De Groot (2004), De Groot (2008), Yinon (2004) and Platzner et al. (1997).

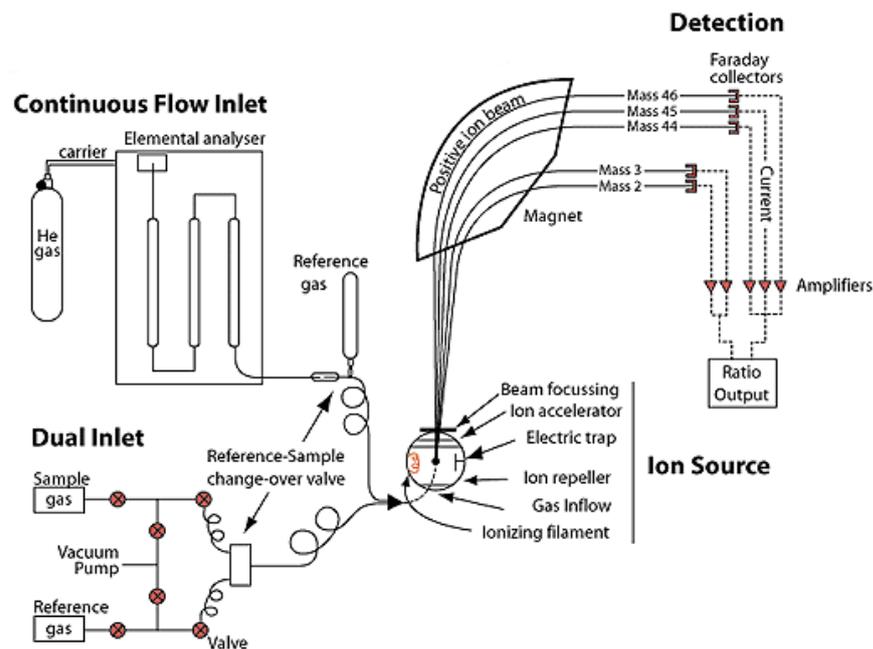


Figure 1.8: Schematic of an IRMS instrument showing different measurement options – either Continuous Flow or Dual Inlet (<http://www.gwadi.org/tools/tracers/methods>).

Dual Inlet IRMS is the traditional method used for isotopic studies and it is still considered the most accurate in terms of sensitivity of measurement. Both a reference of known isotopic composition and the sample gas to be measured are loaded into separate compartments called bellows, which function as sample delivery chambers. During a measurement, the reference and the sample gases are alternatively fed into

the ionisation chamber of the mass spectrometer throughout a sequence. Doing this allows the instrument to continually correct the unknown samples gas against the reference gas values, increasing the accuracy and precision of the results obtained and removing the need for further isotopic scale calibration (Platzner et al., 1997).

The main drawback of dual inlet measurement of gases is sample preparation. The off-line sample preparation processes required to convert solid samples into gases are time consuming and can be complex, generally requiring a set-up of vacuum lines, cryogenic traps and furnaces to be completed (SIRFER, 2010).

1.8.2. Continuous Flow IRMS

In continuous flow systems, the preparation of a sample into gas form is conducted on-line with the IRMS. This decreases the time and the cost of analysing samples and makes preparation simpler. The commercialisation of these sample systems has significantly increased the number of users of IRMS (De Groot, 2004).

The key to continuous flow methods is the split system. For precision and accuracy to be maintained during measurement, the gas pressure entering the ionisation source must be kept stable. This is undertaken using an open split interface where a small capillary is used to calibrate and sub-sample gas from the main flow exiting the sample preparation instrument. Typical gas pressures exiting an elemental analyser are anywhere from 80 to 150mL/min. This is reduced and standardised to ~0.3mL/min for the rate entering the IRMS (2007).

A schematic of the open split interface is shown in Figure 1.9 (Revesz et al., 2012). The interface allows the sample (or standard) gas to be mixed with He carrier gas (the same carrier as used in sample preparation systems) before entering the inlet capillary into the mass spectrometer (Platzner et al., 1997, Werner et al., 1999). It also allows for control of reference gas peak height, which should be scaled to the same height as sample peaks to maximise sample precision and calculation of δ value. Note, in Figure 1.9, either the sample open split or the reference open split are open at any one time (not both).

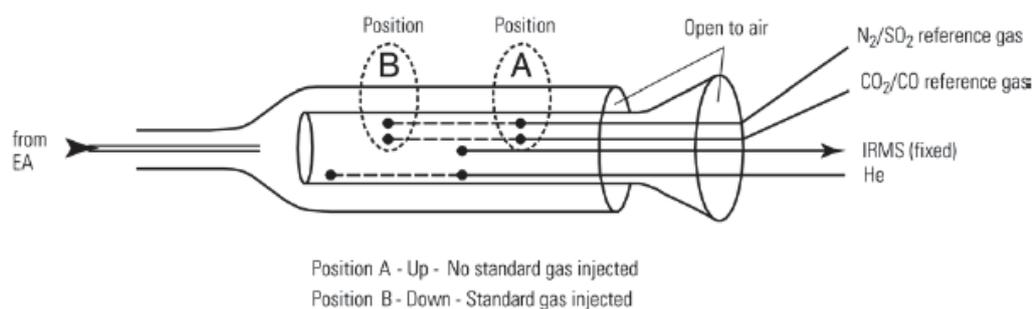


Figure 1.9: Schematic of an open split inlet in a continuous flow IRMS (Revesz et al., 2012)

Depending on the sample or element to be measured, different combustion or separation techniques can be used to convert the sample into a simple gas. These techniques include elemental analysers for bulk analysis, gas or liquid chromatography systems for compound specific studies, or gas bench set-ups for measuring the head space of vials containing water, atmospheric gas or carbonates (where acids are injected into the vial before sampling).

Bulk Stable IRMS

Bulk Stable IRMS Analysis (BSIA) utilises combustion or temperature conversion by an elemental analyser to convert a sample of interest into its gas constituents. The gases measured by the mass spectrometer are shown in Table 1.5.

Element	Isotopes of Interest	Gas Species Measured	Common Isotopologues of Interest	Masses Measured
Nitrogen	^{14}N , ^{15}N	N_2	$^{14}\text{N}^{14}\text{N}$, $^{14}\text{N}^{15}\text{N}$, $^{15}\text{N}^{15}\text{N}$	28, 29, 30
Carbon	^{12}C , ^{13}C	CO_2	$^{12}\text{C}^{16}\text{O}_2$, $^{13}\text{C}^{16}\text{O}_2$, $^{12}\text{C}^{18}\text{O}^{16}\text{O}$	44, 45, 46
Oxygen	^{16}O , ^{18}O	CO	$^{12}\text{C}^{16}\text{O}$, $^{13}\text{C}^{16}\text{O}$, $^{12}\text{C}^{18}\text{O}$	28, 29, 30
Hydrogen	^1H , ^2H	H_2	$^1\text{H}^1\text{H}$, $^1\text{H}^2\text{H}$	2, 3

Table 1.5: Isotopes measured by BSIA and the actual gas species measured

Also shown in Table 1.5 are the isotopologues (molecules that differ only in their isotopic composition) measured for each element, with the mass number measured. Section 1.6.4 details key corrections undertaken for CO_2 and H_2 measurements that are prone to interference from ^{17}O and H_3^+ .

The isotopic values obtained during measurement represent the isotopic composition of the material as a whole. Samples are weighed into either tin or silver capsules. Isotopically known materials are run alongside unknown samples in the same sequence utilising an auto sampler carousel that is purged with helium to dry and prevent atmospheric contamination. There are two types of elemental analysers that are used for bulk stable analyses.

Quantitative High Temperature Combustion

Work began to adapt elemental analysers for IRMS in the 1980's. The first successful attempt was a manual system in 1983 (Preston and Owens, 1983), followed shortly after by a fully automated system in 1984 (Barrie and Workman, 1984 as discussed in (Barrie et al., 1989)). Current instruments are highly automated and purpose built for IRMS.

High temperature combustion elemental analysers are used for the measurement of carbon, nitrogen and sulfur. For analysis, samples are weighed into tin capsules that are dropped into a combustion furnace heated between 900 and 1700°C, though typically temperatures around 1000°C are utilised. The combustion furnace is packed with oxidative materials such as chromium oxide on aluminium and silvered cobaltous oxide for carbon and nitrogen, and reduced copper or tungstic oxide for sulfur (SIRFER, 2010).

The sample is dropped into the combustion tube that contains an inert helium carrier gas. An additional pulse of oxygen is introduced to increase the temperature of the reaction and to cause combustion of the sample at over 1800°C (termed 'flash combustion'). The resulting mixture of gases is swept by the helium carrier stream through to a reduction tube packed with reduced copper to remove excess oxygen and to reduce nitrogen oxides to elemental N₂ (SIRFER, 2010). A packed 5 Å GC molecular sieve is used to separate N₂ and CO₂ before they enter the open split interface for entry into the IRMS (Benson, 2009).

Carbon and nitrogen isotopes can be measured from the same sample however sulfur must be measured separately due to the difference in internal packing material required for combustion. Precision of the measurement values obtained however, is

increased when isotopes are measured separately due to stability being maintained within the IRMS detectors.

The internal packing materials and configuration of the set-up of the Flash instrument for measurement of carbon isotopes can be found in Chapter 2.

Quantitative High Temperature Conversion

High Temperature Conversion Elemental Analysers (TC/EA) are used for the conversion of bulk samples into the simple gases H₂ and CO for the measurement of hydrogen and oxygen isotopes. The sample is weighed into a silver capsule and dropped into a hot glassy carbon furnace held at 1450°C. As for elemental analysers, a molecular sieve is used to separate the reaction gases before they enter the open split interface (SIRFER, 2010). Hydrogen and oxygen are generally measured separately due to issues that can arise with stability when changing the settings within the IRMS for detection of hydrogen isotopes mid-run (Gehre and Strauch, 2003). The internal packing materials and conditions used in the measurement of oxygen isotopes using TC/EA are detailed in Chapter 5.

Compound Specific IRMS

In compound specific IRMS analyses, the isotopic compositions of target molecules within a mixed sample are measured. The most common method of separating a sample on-line is by gas chromatography. Liquid chromatography is becoming more prevalent as the technology adapts however little work has been published using this technique in a forensic context. Chromatographic methods separate a bulk sample into its constituents and during peak windows of interest a switch diverts the separated sample into a combustion tube heated between 700 and 800°C to convert the sample into a simple gas. The combusted sample is directed through an open split to the mass spectrometer. This method is most useful for volatile substances such as plant fatty acids, petroleum, food and perfume authentication and drug studies (Lichtfouse, 2000).

1.8.3. Laser Spectroscopic Techniques

Laser spectroscopic methods are becoming prevalent for water, carbon dioxide and methane gas studies. The most common type of spectroscopic method is via cavity ring down spectroscopy (CDRS). In this method an enclosed cavity is filled with laser light by bouncing it between three mirrors. The intensity of the light is monitored and once the laser is turned off, small leakages of light lead to degradation of the signal over time. If a sample is also present in the chamber, absorption of the light takes place, decreasing the time taken for the light to degrade. The difference in degradation between an empty chamber and a sample filled chamber is characteristic for the gas species and isotopic ratio of the sample being measured (Piccaro, 2011). These systems are deployed remotely for gas measurement and climate studies (Schauer et al., 2003).

1.9. Overview of Isotope Ecology and Isotopic Discrimination

1.9.1. Isotope Cycling and Whole Earth Isotope Values

Whole earth isotope values were set when the earth was formed. The relative amounts sectioned into different materials within the earth however are able to shift through the equilibrium and kinetic fractionation events observed as part of natural cycles (Meier-Augenstein, 2010). Coplen et al (2002) presents a valuable reference of the expected δ values for a range of materials., Figure 1.10 and Figure 1.11 show the typical carbon and oxygen isotopic ranges measured for a series of naturally occurring materials (Coplen et al, 2002). Also included in these figures are references to available standard materials and their published $\delta^{13}\text{C}_{\text{VPDB}}$ and $\delta^{18}\text{O}_{\text{VSMOW}}$ values.

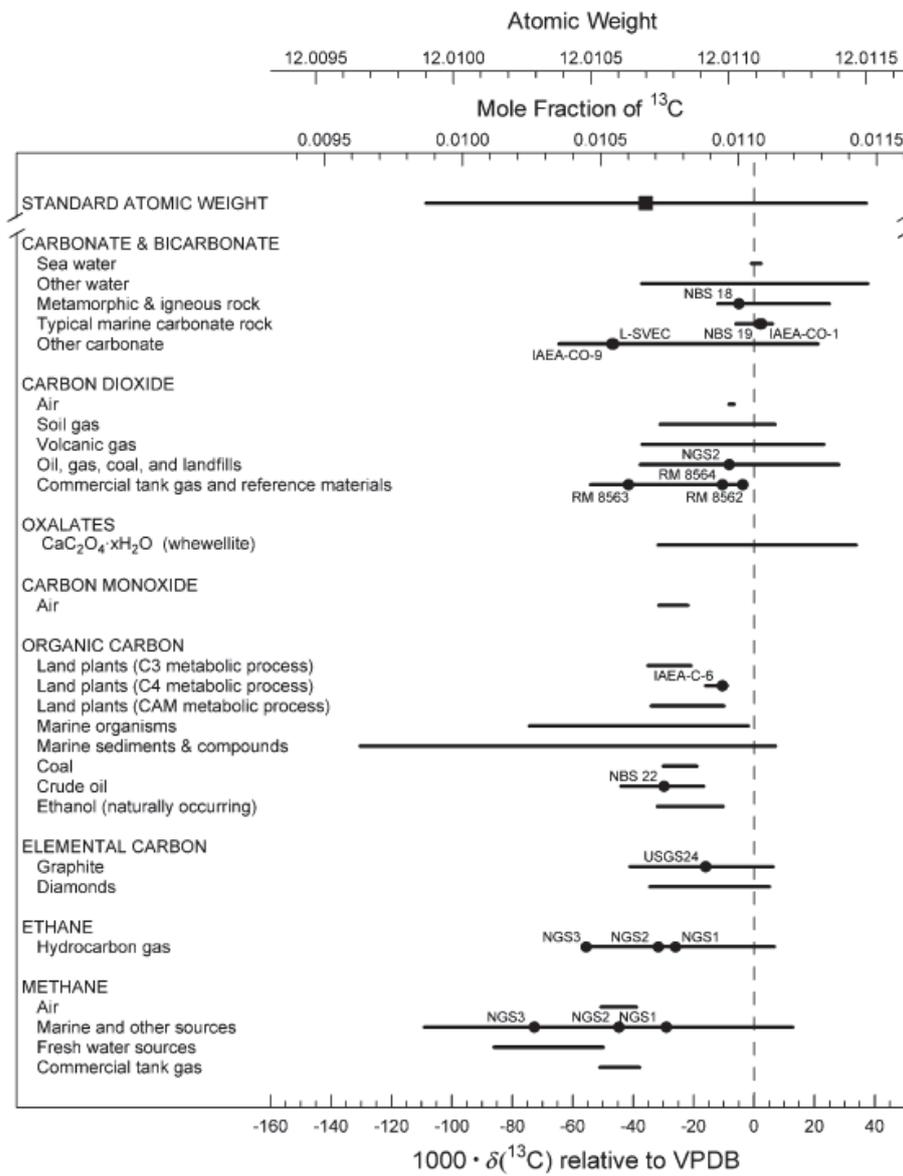


Figure 1.10: Carbon isotopic abundance ranges for naturally occurring materials (Coplen et al, 2002).

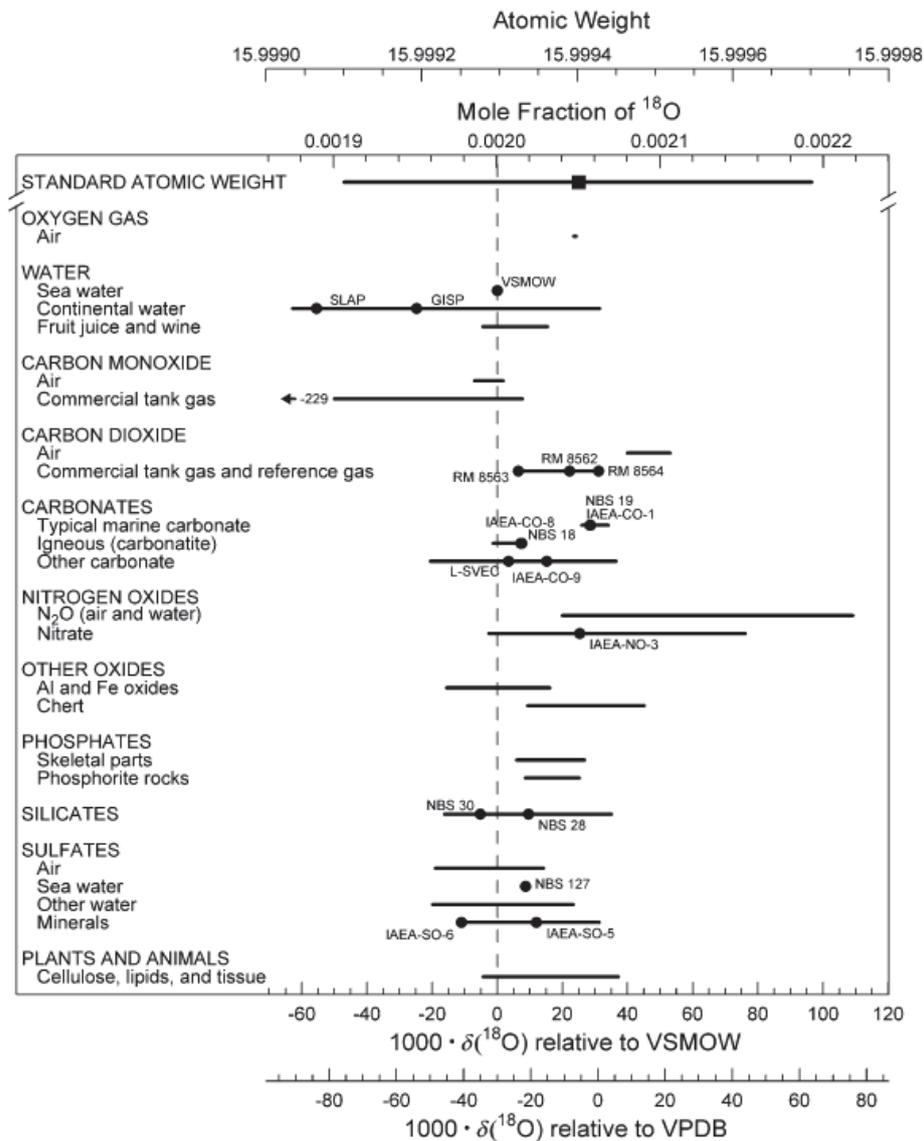


Figure 1.11: Oxygen isotopic abundance ranges for naturally occurring materials (Coplen et al, 2002).

The most fundamental natural cycle for isotopic distribution is the global hydrological cycle. This cycle encompasses how water is cycled between the oceans, air masses and ground or surface waters. The isotopic compositions of precipitation are related along the Global Meteoric Water line defined by Craig (1961), with the equation:

$$\delta^2\text{H} = 8.0 * \delta^{18}\text{O} + 10$$

Variations in the isotopic values of precipitation and land mass water occur for two main reasons (Gat et al., 2001):

- Evaporation of surface ocean water; and
- Rain out of vapour masses as they move to cooler and higher climates. Evaporation of water from the surface of the ocean is a kinetic fractionation process that varies with temperature and air humidity (Lloyd, 1966). This leads to varying levels of fractionation with the vapour always being depleted in the heavy isotope compared to the source water (Gleixner and Mugler in (Dawson and Siegwolf, 2007)). At higher temperatures fractionation is smaller than at lower temperatures due to the increased energy available for reaction by the heavy isotope (Cappa et al., 2003).

Fractionation also occurs during precipitation – with heavy isotopes being lost at a faster rate than light isotopes. This is due to (Gleixner and Mugler in (Dawson and Siegwolf, 2007)):

- The continental effect – when vapour travels over land, prior rain out depletes the remaining air mass. The longer the distance, the more distinct the effect. In Europe for example, fractionation occurs at around -2 ‰ per 1000km over inland areas (Rozanski et al., 1982). Further to this, fractionation and mixing of air masses occurs as temperatures and humidity varies.
- The amount effect – the heavier the rainfall event, the more depleted the remaining cloud vapour becomes.
- The altitude effect – when a cloud moves up a mountainside, cooling of the vapour results in depletion of the precipitation the higher the vapour travels.
- The latitude effect – the more that precipitation moves away from the equator, the more depleted it becomes due to cooling of the air mass.

The effect of these factors on the Meteoric Water Line (as defined by Craig, 1961) are summarised in Figure 1.12.

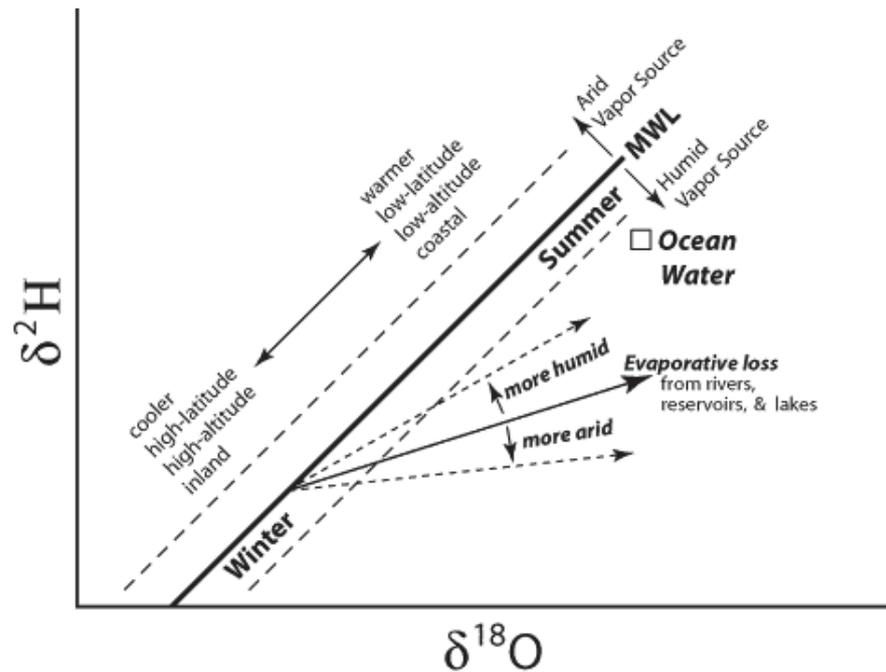


Figure 1.12: Effect of a range of factors on $\delta^2\text{H}$ and $\delta^{18}\text{O}$ values of water

(<http://web.sahra.arizona.edu/programs/isotopes/oxygen.html>)

Also observed in Figure 1.12, water sources such as ground water, rivers and basins are primarily a product of the precipitation that feeds into them, less the effect of evaporation. In particular, closed basins such as lakes can exhibit highly enriched isotopic values due to the continual evaporation of light isotopes from stored water (Gat et al., 2001).

The overall isotopic variability of water globally is monitored and measured by the Global Isotopes in Precipitation Network (GNIP, http://www-naweb.iaea.org/naweb/ih/IHS_resources_gnip.html). This monitoring program is run by the IAEA and measures $\delta^2\text{H}$ and $\delta^{18}\text{O}$ isotopes alongside traditional metrological information in precipitation collected from a large number of collection stations internationally. Further discussion on this kind of monitoring is included in section 1.10.3.

Naturally occurring carbon dioxide cycling is primarily driven by the equilibrium exchange of dissolved inorganic carbon (bicarbonate, HCO_3^-) in oceans and CO_2 in the atmosphere (Fry, 2006). This in turn informs other sources of carbon including

plant $\delta^{13}\text{C}$ values, which are derived from atmospheric uptake of CO_2 and fixation using photosynthesis. Typical inputs, fractionation and processes effecting the carbon cycle are shown in Figure 1.13 (Fry, 2006).

Active exchanges are also now occurring between the atmosphere, terrestrial ecosystems and the ocean. In the last 50 years, anthropogenic inputs from terrestrial ecosystems have caused a net depletion in atmospheric carbon to -8‰ (Keeling et al., 1979, Keeling, 2001). The main cause of this depletion is the large increase in the concentration of burnt fossil fuels entering the atmosphere (Sharp, 2007).

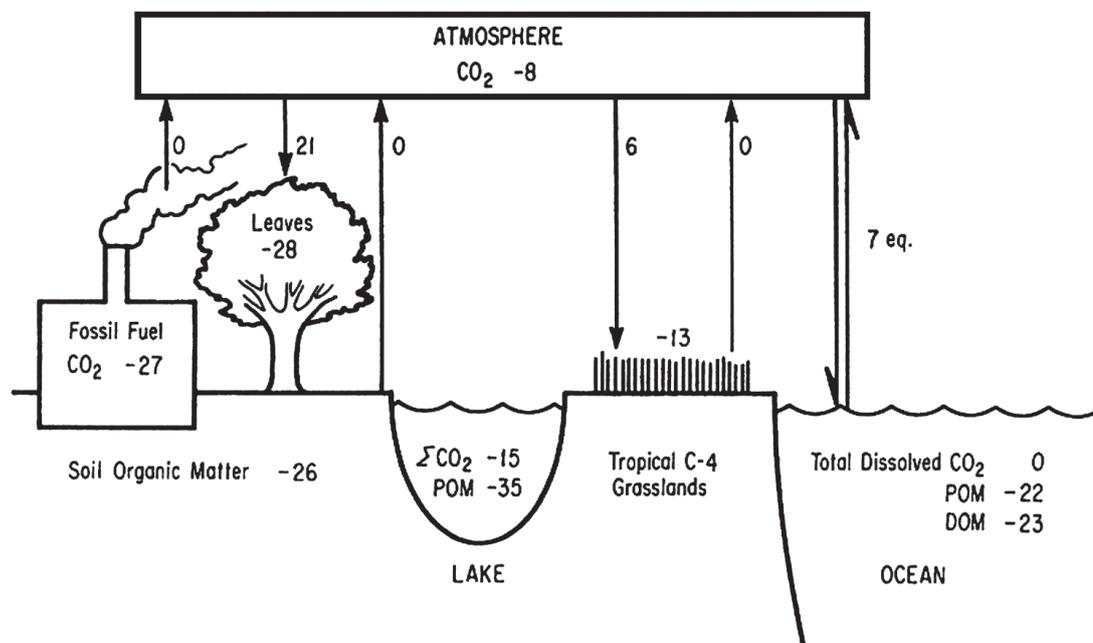


Figure 1.13: Typical sources of variation in the $\delta^{13}\text{C}$ distribution. Double arrows denote an equilibrium fractionation event (Fry, 2006).

1.10. Photosynthesis and Plant Isotopic Values

1.10.1. Carbon Isotopes

Overall, plant tissue values are a product of the isotopic value of the CO_2 absorbed, the environmental conditions CO_2 is absorbed under and the method used to assimilate the carbon into plant tissues.

The fundamental process affecting the stored isotopic values of trees lies in the type of photosynthesis used (Mathews et al., 2000). Adaptations present in leaf cells are used in times of heat stress and low water availability. The stomates (or pores) in the leaf allow diffusion of carbon dioxide and oxygen into and out of the intracellular space. When stomates are open in times of high heat, water loss due to evaporation is a key issue. Both spatial and temporal adaptations by plants exist to overcome this problem.

The most common type of photosynthesis (C_3 photosynthesis or the Calvin Cycle), uses a three carbon enzyme – ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) - to convert carbon dioxide in the mesophyll cells of the leaf, through a number of intermediates to sucrose. Rubisco has the potential to reverse the reaction however, recreating carbon dioxide when oxygen levels are high and carbon dioxide levels decrease. This occurs when stomates are forced to close to conserve water due to dry or arid conditions. The process flow for the Calvin Cycle in C_3 plants is shown in Figure 1.14 (Campbell et al., 2002).

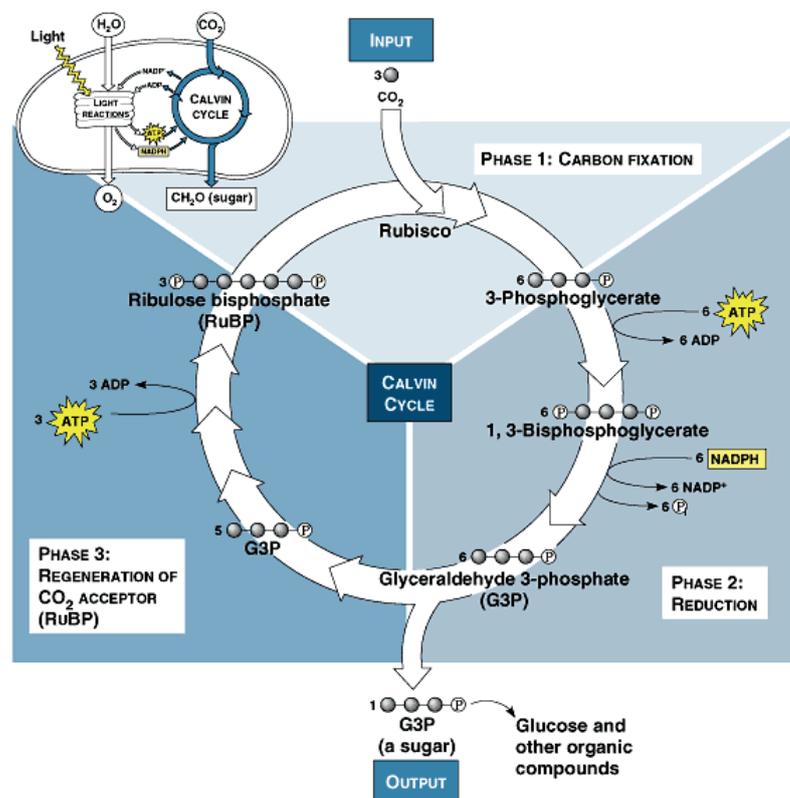


Figure 1.14: Calvin cycle with direct input of CO_2 , as utilised by C_3 plants in photosynthesis (Campbell, 2002).

In heat adapted or C_4 plants, another type of cell – the mesophyll cell – is utilised to prevent this reverse reaction from occurring. The mesophyll cells of the leaf are located closest to the stomates, catalysing a reaction using phosphoenolpyruvate (PEP) carboxylase to transport carbon dioxide to the bundle sheath cells through the conversion of CO_2 to a four carbon organic acid. This enzyme has a high affinity to carbon dioxide and prevents the backward reactions of Rubisco during stomate closure (Mathews et al., 2000, Campbell et al., 2002). As seen in Figure 1.15, this organic acid is released back as CO_2 to the Calvin Cycle.

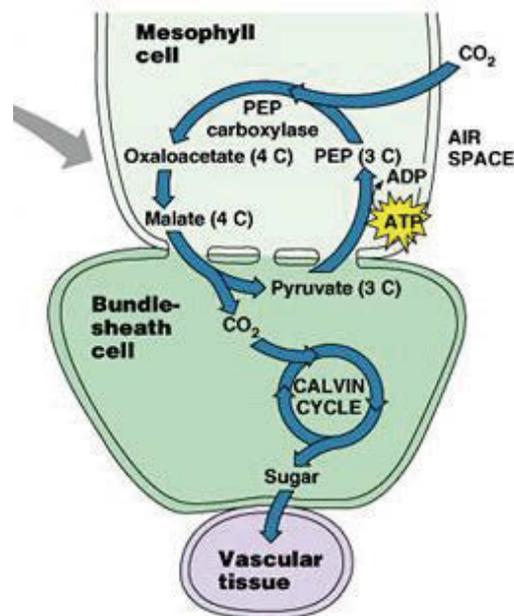


Figure 1.15: Calvin cycle with indirect input of CO_2 via an organic acid intermediate, as utilised by C_4 plants in photosynthesis (Campbell, 2002).

Further to this, Crusselean Acid Metabolism (CAM) plants have additional adaptations for highly arid conditions. These plants are unable to open their stomata during the day and are forced to capture as much carbon dioxide at night as possible. This is achieved by creating an organic acid intermediate at night that is stored and processed into a sugar during daylight hours. CAM plants are also adapted to use C_3 metabolism during milder climatic conditions (Campbell et al., 2002).

Isotopically, the differences in the photosynthetic pathways used to capture carbon dioxide play a significant role in the values for $\delta^{13}C$ of organic materials. In leaves and plants that are able to open their stomates during the day, ^{12}C can be continually

replenished and discrimination against the use of ^{13}C occurs. In leaves and plants that must close their stomates, the reactions are forced to draw from a limited pool of substrate and over time ^{13}C is utilised more than in open systems (Dawson and Siegwolf, 2007).

Farquhar et al (1982) combined these factors into an equation which predicts the carbon isotope value of cellulose:

$$\delta^{13}\text{C} = \delta^{13}\text{C}_{\text{atm}} - a - (b - a) (C_i / C_a)$$

Where $\delta^{13}\text{C}_{\text{atm}}$ is the current $\delta^{13}\text{C}$ value of the CO_2 in the atmosphere (-8 ‰), 'a' is the fractionation associated with diffusivity across the stomatal membrane ($a = 4.4$ ‰), 'b' is the net fractionation by Rubisco ($b = -27$ ‰ for C_3 plants, 5.7 ‰ for C_4 plants) and C_i/C_a describes the difference between the intercellular (C_i) and atmospheric (C_a) concentrations. The C_i/C_a ratio describes the effects observed in plants that create sucrose with their stomata closed and open (See also (Farquhar et al., 1989)). For example, if the stomates are closed, the C_i value will increase, causing the overall δ value of the fixed carbon to rise. Conversely, if the stomates are open, the C_i value will be close to equilibration with C_a , creating δ values that are close to the fractionation seen with Rubisco alone (around -20 ‰) (McCarroll and Loader, 2004).

Overall, the expected $\delta^{13}\text{C}$ values for cellulose in C_3 plants are between -20 and -35 ‰ with an average of -26.7 ‰ (Cerling et al., 1997). For C_4 plants, the average $\delta^{13}\text{C}$ value is -12.5 ‰ (Cerling et al., 1997) with a range of -9 to -19 ‰ (Deines, 1980). For CAM plants, that utilise C_3 photosynthesis when conditions permit, the measured $\delta^{13}\text{C}$ values span the full range of both C_3 and C_4 plants; from -10 to -29 ‰ (Lerman et al., 1974). These results are further confirmed in Sternberg et al. (Sternberg et al., 1984b, 1984a).

Natural variations in $\delta^{13}\text{C}$ values also occur within different tree tissues, due to differences in the chemical structures and enzyme pathways used during synthesis. Marshall in Lajtha and Michener (1994) discuss the difference between the cellulose (polymeric sucrose molecules) and lignin (complex phenolic cross-linked polymers)

molecules produced in the same tree. Differences up to 4 ‰ were measured in different tree tissues by Leavitt and Long (1986).

Francey and Farquhar (1982) and Mazany et al. (1980) found tree ring variation changes (i.e. variation over time) from 0.5 to 1.5 ‰ between rings, likely due to source input differences. In addition, a number of studies discussed by Leavitt (2010) showed that the same species of tree at the same site had a $\delta^{13}\text{C}$ variability ranging from 1 to 3 ‰ due to either microenvironmental factors (small differences in moisture, light and temperature) or genetic variation within a species.

CO_2 source input is still the major contributor to wood $\delta^{13}\text{C}$ variability however. The steadily decreasing $\delta^{13}\text{C}$ value due to anthropogenic input of depleted CO_2 from the burning of fossil fuels (Rayner et al., 1999) is key to this variability. Seasonal trends are also recorded in atmospheric gases, particularly in the northern hemisphere due to increases in photosynthetic fractionation and increased anthropogenic input due to burning of fuels during winter (Randerson et al., 1999, Peters et al., 2007, NOAA).

1.10.2. Oxygen and Hydrogen Isotopes

The oxygen and hydrogen isotopes fixed in cellulose and other molecules within plants have originated from the uptake of water from the soil (McCarroll and Loader, 2004) and reflect, with some fractionation, the δ value of precipitation. Amongst others, Dawson and Ehleringer (1991) found that there is no net fractionation of either hydrogen or oxygen during water uptake by plants into the stem. In the same study, the maturity of the plant was also found to be a factor due to the depth that could be reached by the tree roots.

After a precipitation event, the first site of oxygen and hydrogen fractionation is via soil evaporation. This affects the δ value of the water absorbed into the tree stem. Barnes and Allison (1983) found that a steady state for evaporation occurred up to a set depth. The same study also indicated that mixing from other sources such as degrading organic material, ground or stored water from previous precipitation events also contributed to differences in ground water oxygen and hydrogen values.

The water source utilised by trees is also an important consideration. Ehleringer and Dawson (1992) found that some trees living next to an easily accessible stream water source were instead using deeper ground water sources. Another study by Dawson (1993), showed that less mature plants, who were assumed to be utilising water from precipitation, were actually being helped by larger trees which drew water up from underground sources via a process called hydraulic lift. Similarly, Dawson (1998) showed that fog water run-off was used as an alternative water source to deeper water stores during the dry season in California resulting in different isotopic signals being recorded within the tissues of the trees than expected.

In the tree, oxygen assimilation prior to cellulose production takes place in the leaf. Roden et al. (2000) puts forward a model for the fractionation of oxygen and hydrogen, based on two key fractionation events. The first is fractionation due to transpiration of water in the leaf through the stomates, leading to enrichment within the leaf space. Further fractionation occurs due to evaporation on the leaf surface creating a higher proportion of heavy isotopes outside of the leaf, which in turn are available for re-absorption into the leaf. Helliker and Ehleringer (2002) showed that the $\delta^{18}\text{O}$ values of grasses grown in high humidity were different compared to the same grasses grown in lower humidity (up to 10 ‰).

Included in this first fractionation event is the fractionation associated with sucrose synthesis. Using the models developed, Roden et al. (2000) predict that the non-exchangeable H and O atoms in sucrose exported from the leaf will be -171 ‰ depleted and 27 ‰ enriched respectively compared to leaf water due to sugar synthesis.

The second fractionation event (Roden et al., 2000) is based on further exchange of hydrogen and oxygen during cellulose synthesis into the intermediates hexose phosphate and triose phosphate. Farquhar (1998) predicted that during hexose formation 2 of the 10 oxygen atoms in a cellobiose unit are able to exchange. If the triose phosphate pathway is utilised then 6 out of 10 oxygen atoms are able to exchange (Hill et al., 1995). Results from Farquhar (1998), Hill et al. (1995) and Roden et al. (2000) indicate that somewhere between 35% and 50% of sucrose follows

the triose phosphate pathway however multiple cycles through the intermediate pathways are possible, increasing total observed fractionation.

1.10.3. Isoscapes

Isoscapes (or isotopic landscapes) are visual representations of the isotopic variations seen within environmental cycles (West et al., 2010). The most predominant use of isoscapes to date has been in mapping the variations of δD and $\delta^{18}O$ in the hydrological cycle (Bowen in (West et al., 2010). Using the data collected by the Global Network for Isotopes in Precipitation (GNIP), maps have been published describing the movement and variation of precipitation on a whole earth scale (<http://wateriso.eas.purdue.edu/waterisotopes/>). This modelled data has been instrumental in understanding the changes and the isotopic behaviour of atmospheric vapour. As part of an ongoing project with the University of Purdue, www.isomap.org allows users to create isoscapes models and maps within specific spatial and temporal parameters (Bowen et al., 2014).

Forensically, isoscapes have been produced or used for comparison with data from human provenancing studies (Valenzuela et al., 2011, Podlesak et al., 2012), milk (Chesson et al., 2010b), animal products (Chesson et al., 2011, Monahan et al., 2012), beeswax and honey (Chesson et al., 2010a) and marijuana (Hurley et al., 2010a, West et al., 2009a). Although very useful in comparing measured data, the isoscapes produced are still prediction models that are governed by the parameters set by the creator (West et al., 2010). For some locations that are not as well studied, the variability described in the hydrogen and oxygen isoscapes is significantly limited to prediction only and hence the accuracy and resolution is low.

1.11. Use of IRMS in Forensic Science

The prospect of using isotopes to learn more from the investigation of materials that have forensic applications is an enticing one. While modern analytical techniques are powerful in identifying what a substance is, the major limitation is comparing beyond chemical classification. IRMS is a tool that can be used to address this limitation and

the number and variation of studies being undertaken has increased dramatically in the last 5-10 years.

In particular a large amount of work has been done into the comparison of illicit drugs, explosives and human provenancing. Comprehensive review articles by Benson et al. (2006), Carter et al. (2005), Daeid et al. (2010), Gentile et al. (2015) and the papers published by the Forensic Isotope Ratio Mass Spectrometry (FIRMS) network (e.g. Carter et al. (2009)) explore the isotope work conducted in support of a broad range of forensic evidence types. For this review, only forensic IRMS studies that include organic or plant based materials will be discussed.

1.11.1. Analysis of Organic Evidence using IRMS

The majority of studies to date have centred on provenancing and comparison of illicit drug samples – in particular marijuana, cocaine and heroin. Marijuana is the only drug that is used in its natural form and little fractionation occurs between growing and seizure. Denton et al. (2001) measured the $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ of marijuana from Australia, Papua New Guinea and Thailand. Although the samples could not be geographically discriminated by their isotopic values, the study found that hydroponically grown plants were more depleted in $\delta^{13}\text{C}$ than outdoor plants. The same depletion was observed for plants that were grown outdoors in high rainfall areas. These findings were confirmed by West et al. (2009b).

More recently, building on the work of Ehleringer et al. (1999), a large amount of work has been performed by Hurley et al. (Hurley et al., 2010a, Hurley et al., 2010b) to build and define a model for marijuana seized in the United States. Two models – based on either region of origin (δD values) or cultivation conditions (indoor, outdoor, shade) were developed and calibrated using 500 known samples. When tested using unknown samples, the model to predict region of origin was up to 67% correct while the model to predict cultivation conditions was up to 86% correct.

Muccio et al. (2012) compared the analysis of marijuana samples in bulk with a compound specific analysis of cannabiniol. Bulk measurements were observed to be more discriminating however the sample size in this experiment was small.

Derived from the Coca leaf, cocaine is a semi-synthetic drug. In investigating possible fractionation due to chemical processing, Casale et al. (2005) found that there was no fractionation of carbon and minimal fractionation (up to 1 ‰) for nitrogen isotopes. In a batch study, Galimov (2005) and Ihle and Schmidt (1996) found that processed cocaine samples has low variability. Galimov (2005) in particular measured 10 pure samples with standard deviations for $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ at 0.32 ‰ and 1.2 ‰ respectively, greatly reducing the potential of IRMS for geographically linking seized samples to coca growth locations.

Ehleringer et al. (1999) detected a more distinct range for $\delta^{15}\text{N}$ however, with four location samples varying by up to 7.1 ‰ and successfully built a model to predict the origin of samples (Ehleringer et al., 2000). This model correctly identified the source of 96% of the 200 samples measured in the study.

Heroin and morphine are derivatives of the poppy plant. Since this is a chemical process, fractionation and recombination will occur between the natural material and the chemical used for processing. Meier-Augenstein (2010) discusses that this would result in the isotopic values of the heroin becoming a combination of the geographic region of growth and of the processing chemical (in this article acetic anhydride). Casale et al. (2005) measured fractionation in both $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ while a geographic study by Desage et al. (1991) found that the values of seized heroin were more uniform than expected given the geographic distance of some of the samples, suggesting that the fractionation caused by processing of the samples masked any differences due to growing conditions.

Wooden match sticks studied by Farmer et al. (2005) showed a wide range of variation within the same brand – up to 2.5 ‰ for carbon. Of the 11 European brands of match measured, matches from a similar geographic region were indistinguishable however a sample from India could be differentiated.

Further work on burnt matches by Farmer et al. (2009a) found that the isotopic values of the wood could still be measured using the internal unburnt portions of the match stick. No fractionation was detected between burnt and unburnt matches and the δ

value of these internal portions was unaffected by the presence of ignitable liquids or fire extinguishing materials.

Daeid et al. (2011) investigated the use of IRMS to province un-dyed cotton and found that there was good discrimination and correlation to region of origin of the sample when measuring hydrogen and oxygen isotopes. For casework purposes, the use of IRMS would require larger sample sizes than those typically seen in fibre trace cases however.

1.11.2. IRMS Used for the Comparison of Document Papers

In a study published by van Es et al. (2009), European office papers were analysed by three techniques – X-Ray Fluorescence (XRF), Laser Ablation Inductively Coupled Mass Spectrometry (LA-ICP-MS) and IRMS. 25 brands of paper from 16 manufacturers were collected and measured for carbon, oxygen and hydrogen isotopes using IRMS and a range of elements using XRF and LA-ICP-MS. One sheet from each ream was tested and using the IRMS results in isolation, 21 of the 25 samples could be discriminated. The author also recommended using IRMS in combination with (LA-ICP-MS) to achieve a higher discrimination power. This was the only published study prior to this research that focussed on the use of IRMS for the comparison of document papers.

1.12. Research Strategy

The objective of this research is to develop and validate the use of IRMS for the comparison of document papers in forensic casework. Given the cellulosic nature of paper, carbon and oxygen isotopes have been selected and the following structure will be utilised for both isotopes:

- Method development, optimisation and validation to international standards;
- Determination and establishment of a database of known sample values that are representative of the office paper market within Australia and New Zealand, to inform the overall variability expected;

- Determination of the isotopic homogeneity (inter-ream, intra-ream brand variability) of document papers to aid in interpretation of unknown casework samples;
- A study to understand what is being isotopically measured in bulk paper samples through an examination of the effect of the paper production process on the source cellulose samples; and
- Determination of the effect of a number of variables on the carbon isotopic ratios of paper including moisture, storage environment, age, handling, a number of forensic sampling techniques and printing processes.

A preliminary study utilising a selection of background paper samples for hydrogen isotopes was also undertaken to inform future research into the use of IRMS for document papers.

More generally, the values obtained for paper samples were compared against the work being done internationally on isoscapes to determine whether the results obtained are indicative of the origin of growth and production more broadly. This comparison will be undertaken with particular reference to the source material and production process examination results, to determine whether the bulk measured paper samples are able to be directly compared to isotope prediction maps.

To sit alongside the results obtained using IRMS and to provide a more holistic approach to casework comparisons of paper, the following work was undertaken to ensure that IRMS sits within a structured examination protocol rather than on its own:

- Examination of a sub-set of the background population samples to construct comparison standards for other examination techniques including grammage, apparent sheet density, light examination and TLC;
- Evaluation of the discriminability of papers using the range of different techniques;
- Construction of an examination, interpretation and reporting structure; and
- Testing of the proposed structure through the use of a blind trial.

Chapter 2

2. Carbon Method Development and Validation

The key results of this chapter have been published in Jones et al. (2013c). The published article is included electronically within Appendix 5.

2.1. Introduction and Aims

A common problem faced by forensic document examiners is to compare documents in order to provide an expert opinion as to whether they could have originated from a common source. Most prior research in this discipline has centred on the chemical comparison of pen and printing inks. Where paper examination studies have taken place, they have tended to end at the proof of concept stage and have not focused on the application of the results to forensic casework. Clients of forensic science (investigators, the courts, the public) have an expectation that the results produced by laboratories are demonstrably fit for purpose and robust.

Accreditation of Australian forensic laboratories is granted by the National Association of Testing Authorities (NATA) who assess laboratories against International Organisation for Standardisation (ISO) standard 17025 (2005) and the NATA Field Application Document (FAD) for forensic science laboratories (2015). Regarding method validation, the FAD sets out a number of broad issues that may be required to be examined during a study such as homogeneity, concentration and linear range. Further information on conducting method validation activities are included in a guide published by NATA (April 2009a), with a large number of other publications on the topic, including those published by ISO (1994) and EURACHEM (1998).

For techniques that present a numerical value at their conclusion, international standards advise that an estimate of the measurement uncertainty of the method should be provided with the analytical measurement value to enable the customer or client to assess the precision of the results. The standard for producing this estimate is the

Guide to Expression of Uncertainty in Measurement (GUM), published by the Joint Committee for Guides in Metrology (2008).

This chapter presents the method development and validation experiments undertaken for the measurement of the carbon isotope abundances of document papers using IRMS. For method development, key parameters such as sample size, linear range, number and placement of standards, and sample storage and handling have been tested. Method validation experiments aim to show that performance characteristics relating to precision, accuracy and repeatability are met. Finally, a value for measurement uncertainty was calculated based on the results of the experiments. The results demonstrate that the method is fit for purpose to analyse and report carbon measurements to assist in the discrimination or association of multiple document papers.

2.2. Materials and Methods

2.2.1. Standards and Samples

Table 2.1 contains the international standards and their certified values used. These materials were purchased from the International Atomic Energy Agency (IAEA, Vienna).

Certified Reference Material Name and Material Type	$\delta^{13}\text{C}_{\text{VPDB}}$ Certified Value (‰)	
	Mean	Standard Deviation
LSVEC - Lithium Carbonate	-46.6	0.2
NBS 19 - Limestone	1.95	None published
IAEA-CH-7 - Polyethylene	-32.151	0.050
IAEA-CH-6 - Sucrose	-10.449	0.033
IAEA-CH-3 - Cellulose	-24.724	0.033

Table 2.1: International standard materials used in the method validation

Table 2.2 contains the chemically analogous materials used for the validation experiments, and for assessment as potential laboratory standard materials. These materials were purchased through Sigma Aldrich (Sydney, Australia).

Material Name	Sigma Aldrich Product Number	Batch Number
Microgranular cellulose	C6413	069K0082
Medium fibre length cellulose	C6288	039K6171
Long fibre length cellulose	C6333	038K0055
α - cellulose powder	C8002	109K0114
D + Glucose	G8270	089K00603
Mannose fibres from wood	112585	099K0054
α - D- glucose	159968	MKBB8469
Cellulose Acetate	180955	40198LJ
Starch from Corn	S4126	084K0009
Starch from Wheat	S5127	079K0309
Sigma ultra cellulose	S6790	099K1098
α - Lactose Monohydrate	L3625	099K1540

Table 2.2: Materials used in the carbon method validation experiments and for evaluation as laboratory standards

A number of plain 80gsm papers were used in these experiments. These were chosen at random from a population of paper samples that were purchased from throughout the Canberra (ACT, Australia) region.

All standards (international and laboratory) were stored in glass vials with teflon lined screw top lids (Sigma Aldrich) or the original packaging in a Perspex desiccator with self-indicating silica gel (LabServ, Biolab, Auckland, New Zealand). Paper samples were stored either in the ream or in a plastic sleeve in a laboratory environment. Unless otherwise specified, all sample measurements were performed in triplicate.

Laboratory gases were purchased from Coregas (Canberra, Australia). All were analytical grade with purity higher than 99.99%. The gases used for this method validation were helium, carbon dioxide and oxygen.

2.2.2. Instrumentation and Equipment

A Genius ME5 (Sartorius, Goettingen, Germany) analytical balance was used to weigh the international standards sucrose and cellulose. After experimentation, the standard weight for analysis on the IRMS for these standards was set at a weight of $100\mu\text{g} \pm 20\mu\text{g}$. Polyethylene and paper samples were prepared using a 1mm and a 1.2mm Harris Uni-core micro-punch (Proscitech, Queensland, Australia), respectively. Samples were placed into 3.3 x 4 mm tin capsules for solids (IVA-Analysentechnik, Dusseldorf, Germany) and stored in a Perspex desiccator with self-indicating orange silica gel until measurement.

Individual samples were dropped into a FlashEATM 1112 elemental analyser using an AS300 auto sampler (Thermo Finnigan, Bremen, Germany). Combustion was in a quartz reactor packed with silvered cobaltous/cobaltic oxide and copper oxide held at 900°C. An oxygen pulse at 250ml/min for 3 seconds was used to increase combustion efficiency and conversion of the sample to CO₂. The gases were then passed through a reduction furnace of reduced copper at 680°C, before moving through a water trap of magnesium perchlorate.

The internal packing materials for the combustion and reduction furnaces are shown in Figure 2.1. A packed GC column held at 45°C was used to separate analyte gases. A ConFlo III interface fed the analyte gas into a DELTAplus XP IRMS (Thermo Finnigan) instrument for carbon isotope ratio measurement, scaled against pulses of CO₂ reference gas with a purity of 99.999%, set to a $\delta^{13}\text{C}$ value of 0 ‰. The operating software for the Flash EA and IRMS instrument was Eager 300 Version 2.1 and Isodat 2.0 respectively which was upgraded to Isodat 3.0 during the course of the experiments.

When assessing the effects of other paper examination techniques such conditioning of the papers, a Binder KBF-115 environmental chamber (Crown Scientific, Sydney, Australia) was used to control temperature and humidity.

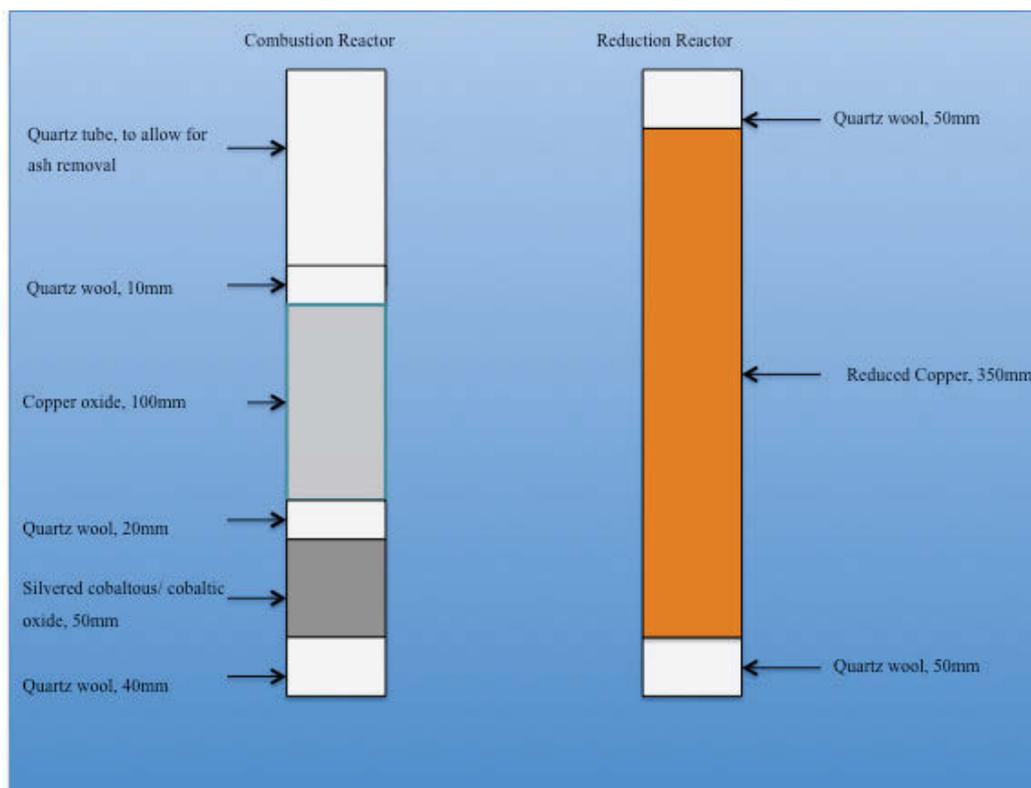


Figure 2.1: Internal packing materials used in the Flash EA 1112 for the measurement of carbon

2.2.3. Correction of Values

The delta values produced by the Isodat (IRMS control) software were exported into Microsoft® Excel. Instrument drift and scale normalization was undertaken using the procedures outlined in Chapter 1.6.

2.2.4. Statistical Tests

Basic statistics, including calculation of the mean and standard deviation of sample measurements were undertaken using Microsoft® Excel.

GraphPad Prism (version 5 for Windows or Mac, GraphPad Software, La Jolla California USA) was used for other statistical tests. This software was chosen for the inclusion of the Kruskal-Wallis non-parametric analysis of variance (ANOVA) test with Dunn's post-hoc test. This test was used instead of a parametric ANOVA test as in other experiments some paper reams used were shown not to conform to a parametric curve, even when run in high sample numbers (i.e. where $n \geq 50$ or more).

2.2.5. Method Development

The following target performance characteristics were set for method development. The method parameters were varied to ensure that these targets were met at all stages of an experimental run and in successive experimental runs. The targets included:

- Full sample peak elution with baseline separation from reference peaks (appropriate reference peak placement);
- Matching of sample peak and reference peak heights by determining the most appropriate sample size and reference peak pressure;
- Linearity of instrument response with increasing sample size for carbon (paper) sample peaks;
- Precision of repeat measurements by ensuring that the most appropriate sample size is selected; and
- Linearity of CO⁺ production (detected on mass 28) with increasing sample weight for each sample type measured (i.e. paper, cellulose and glucose).

Water Absorption

Cellulose, glucose and paper are hygroscopic materials and hence require water content to be controlled for accurate measurements to be able to be assured.

Five papers and five laboratory working standard samples were selected to determine whether water absorption affected the measured δ values of the samples. All samples were measured on day 1 and day 16 (from the time of initial exposure). All samples were left exposed to atmosphere on a laboratory bench for the duration of the test.

Sample Handling, Storage and the Effect of other Paper Examination Procedures

Following on from the absorption study, secondary tests were conducted to measure and evaluate the effect of altering the storage conditions of the samples on the carbon isotope abundances. In routine casework, it was expected that samples would be weighed and stored in desiccators until measured. To test the effect of this, three sets of samples were taken from five papers. One set was measured on the IRMS immediately (named: day one). The two other sets of samples were prepared and

placed in a Perspex desiccator for 15 days. Of these two sets, the sample capsules in the first set were crushed (named: desiccator closed), while the sample capsules of the second set were left open (named: desiccator open) until just prior to measurement to test whether closing the capsule allowed the samples to dry adequately.

Included in the secondary tests were samples that had been treated according to the standard procedure used to prepare papers for measurement of grammage and density. This standard procedure, detailed in Australian Standard 1301.405s (2004), is used to remove any effects from the source of the paper by equilibrating the paper sheets' water content. The paper sheets are placed in a humidity and temperature controlled chamber for 24 hours at 50% relative humidity and 25°C. The humidity of the chamber is then reduced to 25%, where the paper sheets are held for a further 48 hours.

To test the effect of this humidification procedure, the five sheets from which the day one and desiccator samples were taken were placed into the chamber and exposed to the humidity/temperature program. After this was completed, one set of samples was taken from the sheets and run immediately (named: humidity only). A second set of samples was taken, crushed and placed into a desiccator for 7 days (named: humidity + desiccator).

The results from these four different sample treatment and storage conditions were compared against each other and against the day one results.

Design of Sample Run

A number of experiments were undertaken to ensure that the structure of the analytical sequences maximized the accuracy of the results after correction. This included measuring the following variables:

- Number of international standards – two end member pairs (polyethylene and sucrose) or three standards comprising the two end-member pairs plus one mid range value (cellulose).
- Number of replicates of international standards – three, five or seven.

- Placement of international standards – start and end of the run or start, middle and end of the run.
- Inclusion of blank tin capsules between samples.

To determine the effects of changing these parameters, polyethylene (IAEA-CH-7) and sucrose (IAEA-CH-6) international standards were used to correct the measured values for two international standards (LSVEC and NBS 19) and three laboratory standard materials (medium cellulose, alpha glucose and cellulose acetate), run as unknown materials. Where the international standard cellulose (IAEA-CH-3) was not used as a factor in the experiment, it was included as an unknown sample.

2.2.6. Method Validation

Precision and Stability

Carbon isotope abundances for international standards polyethylene (IAEA-CH-7) and sucrose (IAEA-CH-6) measured and corrected over a 24-month period were plotted to determine the range of variation expected over a period of time.

Repeatability

International standard cellulose (IAEA-CH-3) was measured and plotted as an unknown sample in each analytical run over a 12-month period. The corrected values were plotted to determine their range and conformance to the published value and standard deviation.

Accuracy

Medium cellulose, cellulose acetate and alpha glucose samples were run over a 24-month period to assess whether the method produced accurate values over time.

Robustness

Three different laboratory staff of varying levels of experience prepared the same analytical run containing paper, cellulose, glucose and sucrose samples. Triplicates of each of the samples were prepared on different days and corrected against

international standards. All three runs were analysed on the IRMS and corrected by the author (named: Operator A, B and C).

In addition, the same analytical sequence was prepared, run on the IRMS and corrected by a person familiar with IRMS operation but external to the project (named: Operator D).

All four groups of results were compared to determine the robustness of the method to changes in sample preparation and operation of the instrument.

Measurement Uncertainty

The entire procedure, from preparation of international standards through to correction calculations was evaluated for potential sources of uncertainty. Following the method outlined in practical guides published by NATA (April 2009b) and the equations used in Benson et al (2006), the uncertainties associated with the method were separated into sources of method bias/accuracy and method precision. Overall these two types of uncertainty were combined to give an estimate of measurement uncertainty.

Inter-Laboratory Trial

The carbon method developed for paper was applied to two different materials that were distributed as part of the Forensic Isotope Ratio Mass Spectrometry (FIRMS) group trial (2010). Ten replicates of each sample were measured using the method and correction calculations described above.

The results reported back to each laboratory included both a grand mean for the collated results from the 28 participating laboratories, in addition to a corrected mean, based on a Huber's method of correction (2001), annotated to the H15 mean and H15 standard deviation. Using these values and assigning a target standard deviation of 0.2 ‰, a z-score was calculated using the following formula (Brief, 2005):

$$Z \text{ Score} = \frac{\text{Reported Mean} - \text{H15 Mean}}{\text{Target Standard Deviation}}$$

The Z-score provided is an indication of a laboratories' closeness to other participants' measurement values. A value within the range -2 to +2 is considered within a normal range of scores.

2.3. Results and Discussion – Method Development

Using the instrument parameters set out by Benson et al. (2010), full and consistent peak elution was achieved. CO₂ reference peaks were placed at 20, 60 and 320 seconds with the pressure set at approximately 0.6 bar to ensure a sample height of approximately 3000mV (compared to a baseline of between 4 and 10mV). The total run duration was 360 seconds. A typical chromatogram is shown in Figure 2.2. For the majority of sequences, drift correction was not required.

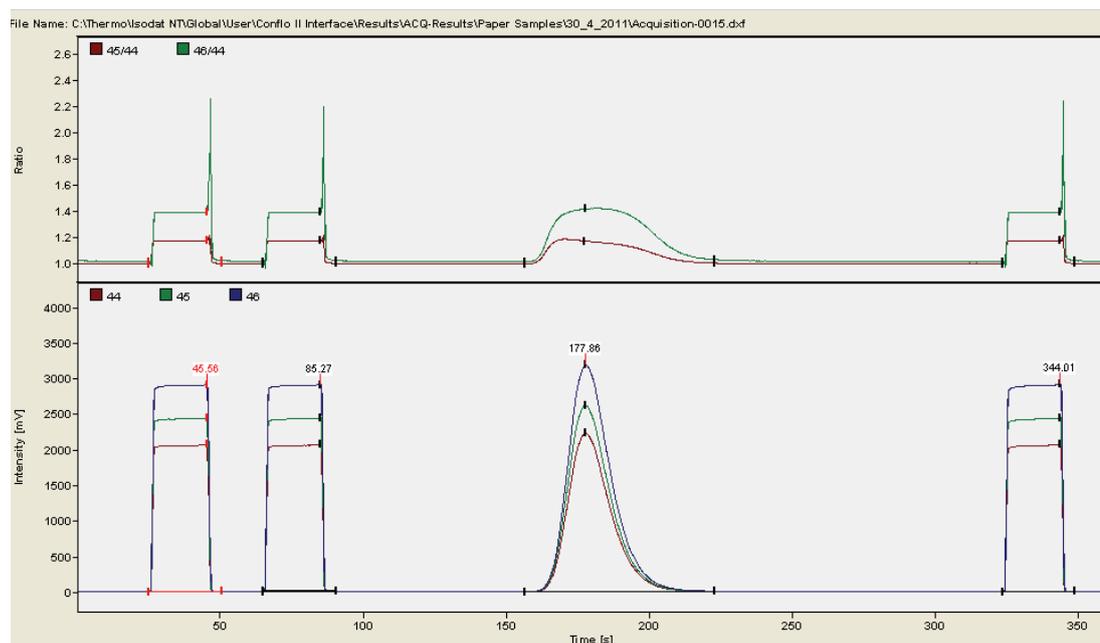


Figure 2.2: Typical chromatogram for measurement of $\delta^{13}\text{C}$

2.3.1. Linearity

A linearity experiment was conducted to determine whether a consistent instrument response was observed as sample size increased for three materials – medium cellulose, alpha glucose and cellulose acetate. This data is plotted in Figure 2.3, Figure 2.4 and Figure 2.5. Additionally, the $\delta^{13}\text{C}$ values of each sample type were measured in a single sequence, corrected using international standards and plotted against the instrument response to determine whether there were any significant variations in the measurement result. This data is shown plotted in Figure 2.6, Figure 2.7 and Figure 2.8. All linearity experiment results are included in Appendix Table 14.1

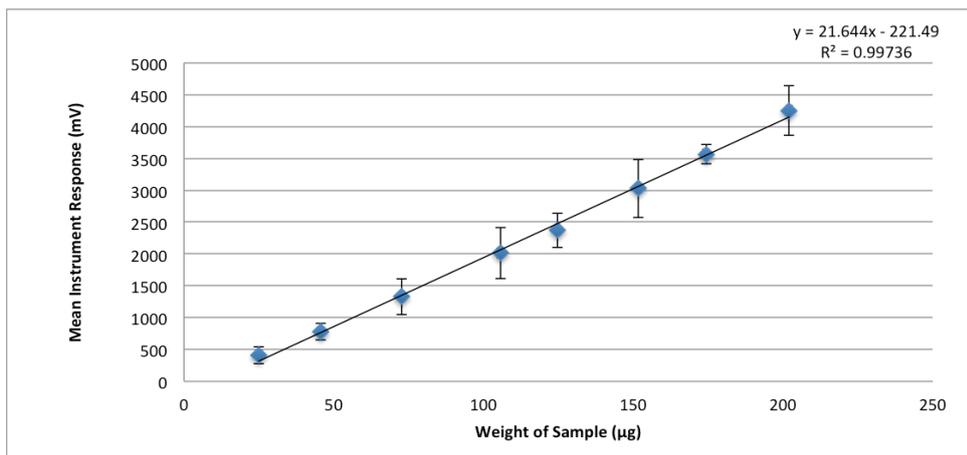


Figure 2.3: Linear response for $\delta^{13}\text{C}$ of medium fibre cellulose

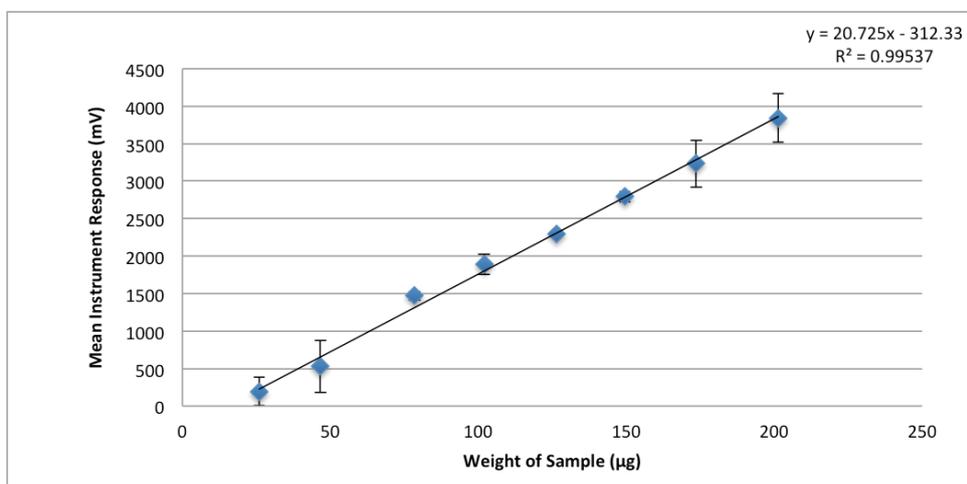


Figure 2.4: Linear response for $\delta^{13}\text{C}$ of cellulose acetate

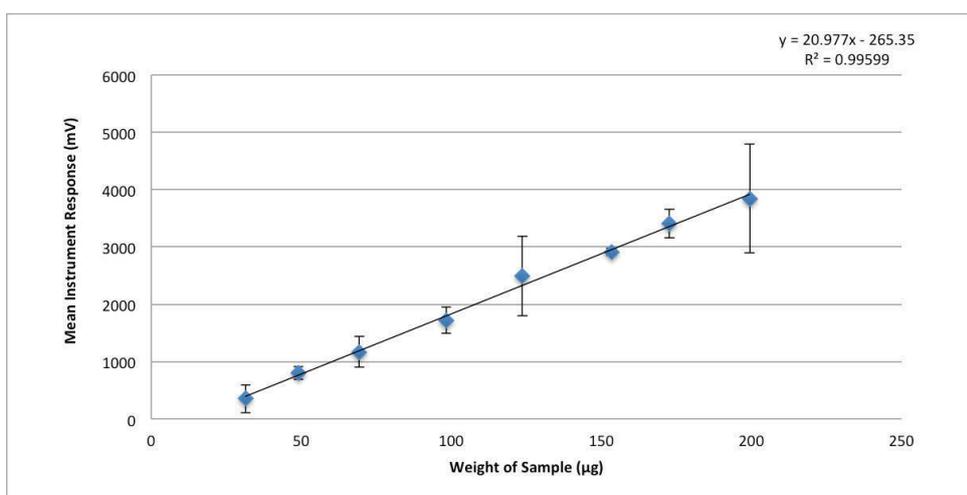


Figure 2.5: Linear response for $\delta^{13}\text{C}$ of alpha glucose

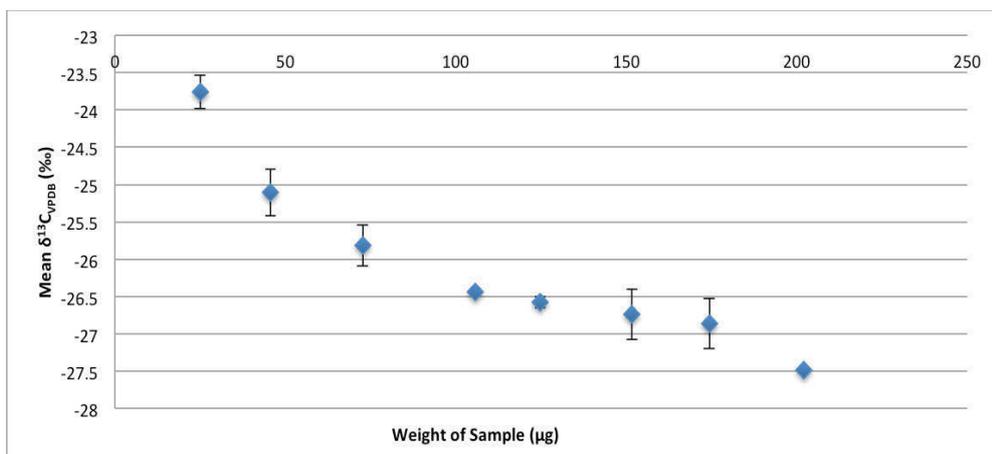


Figure 2.6: Linear range for $\delta^{13}\text{C}$ of medium fibre cellulose

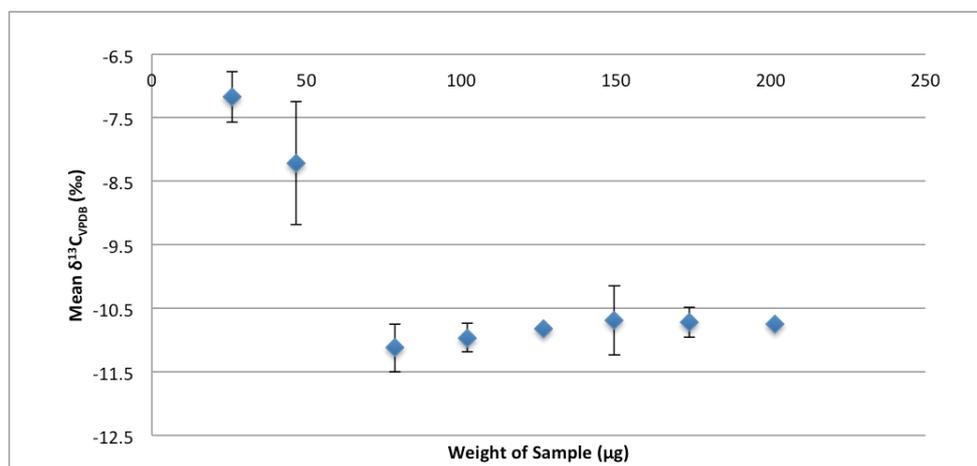


Figure 2.7: Linear range for $\delta^{13}\text{C}$ of cellulose acetate

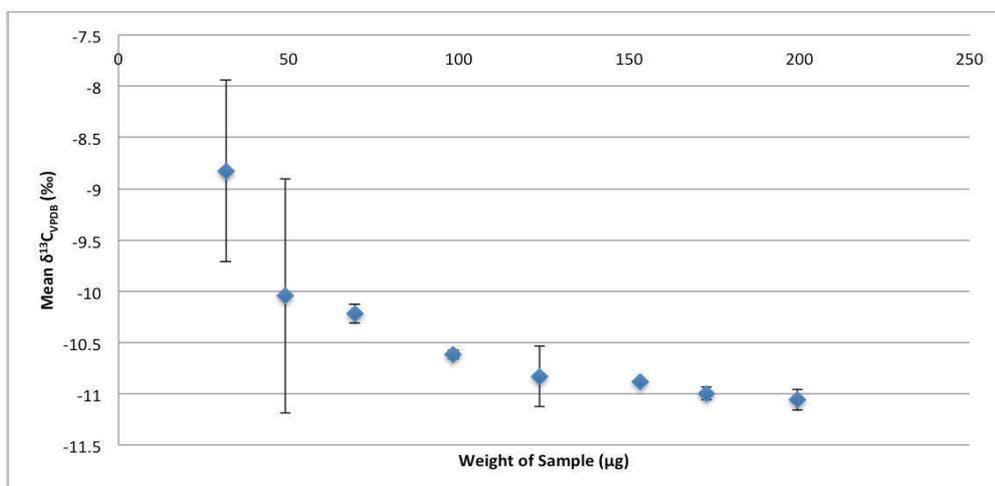


Figure 2.8: Linear range for $\delta^{13}\text{C}$ of alpha glucose

The results of the linearity experiments show that instrument response is linear for increasing sample size and this is not of concern as long as sample size is normalized and controlled within an analytical sequence when measuring $\delta^{13}\text{C}$ of cellulosic materials.

When sample size was plotted against the measured $\delta^{13}\text{C}$ value, a minimum critical value was detected at 75 μg of sample. This was consistent between the three sample types measured. Beyond 75 μg , up to the maximum size tested (200 μg), the $\delta^{13}\text{C}$ values were stable and comparable. Using the largest standard deviation for each of the measured values between 75 and 100 μg for discrimination, only one value at 75 μg would be discriminated from the other values (medium fibre cellulose sample type). This indicates that sample sizes of 100 μg and above are fit for purpose with respect to stable measurements and comparable peak heights, with the required sample size determined by scaling against the desired reference gas peak height.

One known issue with the measurement of carbon isotopes in samples with high carbon content is the production of CO in the combustion reactor, which is subsequently ionized to CO⁺ in the source. This isobaric interference produces molecules that elute a peak at the same time as CO₂, with a mass of 28 (the same mass as measured during nitrogen isotope measurement) (Platzner et al., 1997). While this is of primary concern during nitrogen measurement due to peak interference, it is still of concern with respect to CO₂ measurement due to loss of sample. It is important to monitor therefore, to ensure that the principle of identical treatment still applies to samples that may have slightly different sample weights.

Given that sample loss is of concern, evaluation focussed on determining whether CO⁺ production was linear as a proportion of increasing sample size. Figure 2.9 shows the linear instrument response for the CO⁺ ions produced from cellulose, glucose and paper. For the three sample types, the R² values of the lines of best fit were 0.997, 0.997 and 0.998 respectively. These results show that the production of CO⁺ for each sample type is linear and that the combustion conditions being used are consistent, producing isobaric CO⁺ in a proportion relative to the sample size.

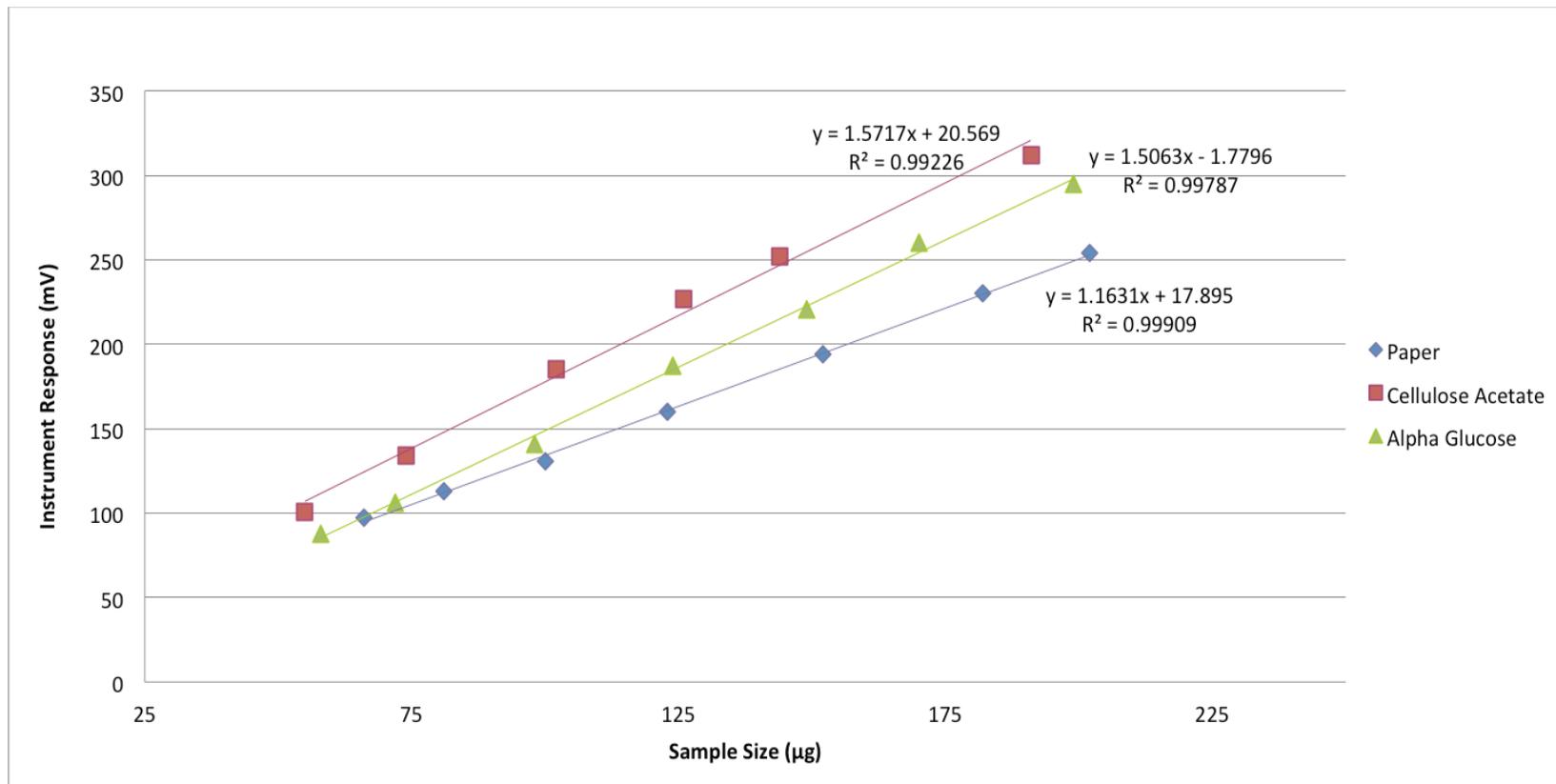


Figure 2.9: Linearity of CO⁺ production (measured on mass 28) for paper, cellulose acetate and alpha glucose samples.

Note: mass 29 was proportional for all materials measured.

2.3.2. Water Absorption

The mean $\delta^{13}\text{C}$ and standard deviation for day 1 and day 16 for each sample type have been plotted in Figure 2.10. The data used to produce this plot is detailed in Table 14.2.

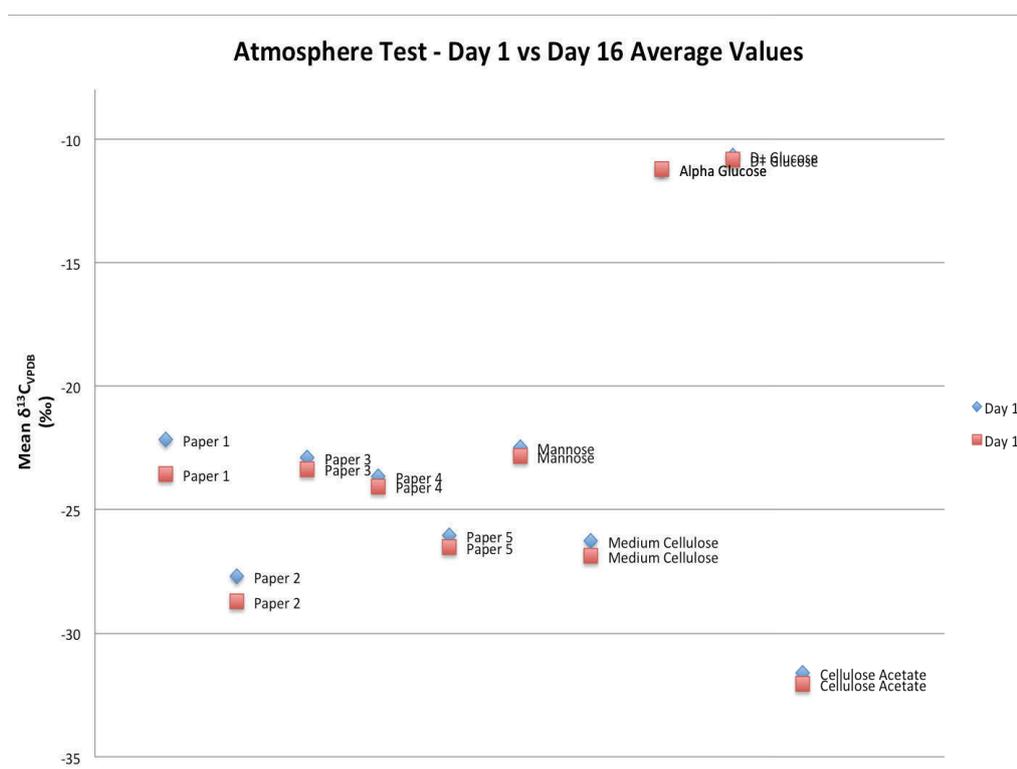


Figure 2.10: Day 1 (blue/diamond) vs. Day 16 (red/square) $\delta^{13}\text{C}_{\text{VPDB}}$ values for samples left exposed to laboratory atmosphere. Error bars are not included as they are minimal at this scale.

The day 1 standard deviation values indicate some issues with the precision of the instrument on this particular day. Due to this, the standard deviation results presented in the difference column in Table 14.2 should be used with caution. The corrected mean values are still comparable to the day 16 values however a number of measurements were excluded as outliers due to this imprecision.

Due to the exclusion of some measurement results, a number of samples contained only two usable measurements, too few to undertake the Kruskal-Wallis test. Instead, an unpaired t-test was used to give an indication of whether the mean values are

significantly different between day 1 and day 16. Whilst this is not ideal, the statistical results provide an indication of any differences, which can be weighed against the actual values and graphical results shown.

Significant differences (to 95% confidence level) were observed in the following samples:

- Paper 1 (p value= 0.03)
- Paper 2 (p-value= 0.006)
- Paper 5 (p-value= 0.0185)
- Mannose (p-value= <0.0001)

These results suggest that there is some effect on the measured $\delta^{13}\text{C}$ values of the samples after exposure to the laboratory atmosphere, where moisture is not controlled. The laboratory atmosphere where these samples were left is notoriously inconsistent and prone to large fluctuations in both temperature and humidity.

Mannose (a type of cellulose) was found to be so hygroscopic that it completely liquefied in the presence of moisture over a period of time greater than 16 days (faster if the level of humidity was higher) so these results were not unexpected for this material. The paper results were surprising however, as even those samples that did not change to a statistically significant level, did experience some change over the 16-day period.

As bonded cellulose is stable, the change observed in the paper samples when exposed to the atmosphere is likely caused by the absorption of water into the paper which had an effect on the chemistry of the calcium carbonate filler material. Once absorbed, the water promotes further uptake of atmospheric gases and degradation of the calcium carbonate to form carbonic acid. With further uptake/exposure to water, the carbonic acid forms and releases gaseous carbon dioxide (Al-Hosney and Grassian, 2005). The magnitude of this release would be affected by the natural hygroscopicity of the paper, imparted by the humidity of the location that the paper was made. For papers with a

greater tendency to absorb water, the effect would be larger and more carbon dioxide would be released.

Isotopically, it is predicted that these papers would become more depleted (i.e. heavier) as the reaction would favour the lighter isotope. This is clearly observed in the isotopic abundance values of the paper samples, with little to no effect from atmospheric moisture seen on the cellulose and glucose standard materials.

2.3.3. Sample Handling, Storage and Paper Equilibration

The mean and the standard deviation for each condition and sample are plotted in Figure 2.11. The data used to produce this plot is shown in Table 2.3.

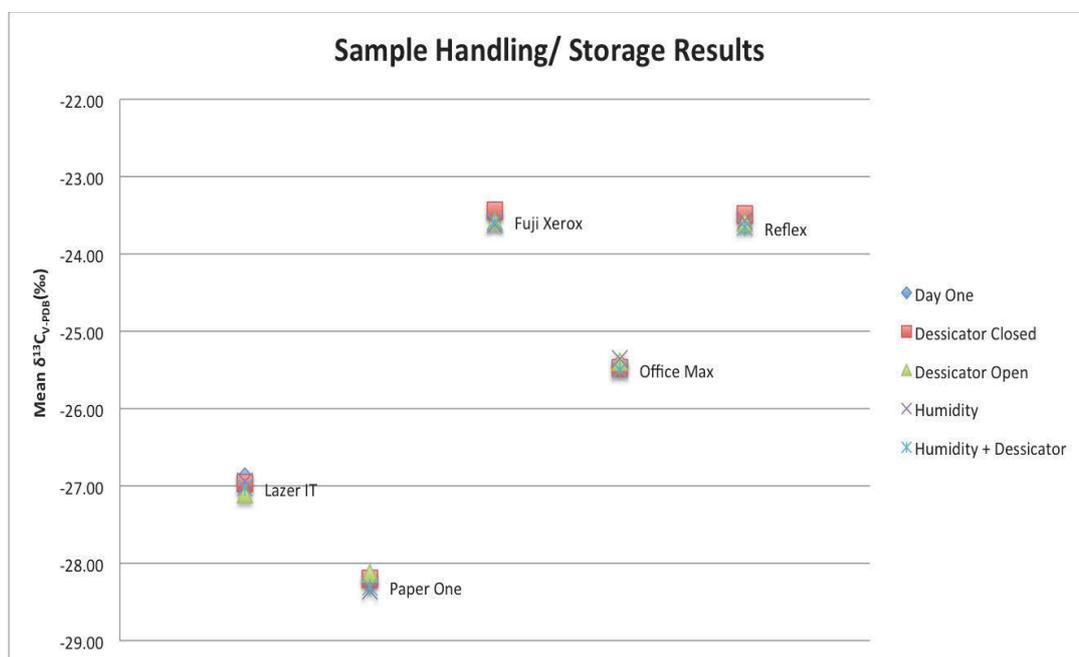


Figure 2.11: Mean $\delta^{13}C_{V-PDB}(\text{‰})$ paper sample values for different treatment and storage procedures

Brand		$\delta^{13}C$ Day One (‰)	$\delta^{13}C$ Desiccator closed (‰)	$\delta^{13}C$ Desiccator open (‰)	$\delta^{13}C$ Humidity only (‰)	$\delta^{13}C$ Humidity + desiccator (‰)
Lazer IT	Mean	-26.88	-26.96	-27.10	-26.93	-27.02
	St Dev	0.14	0.11	0.00	0.08	0.10
Paper One	Mean	-28.21	-28.20	-28.12	-28.35	-28.30
	St Dev	0.01	0.03	0.09	0.00	0.03
Fuji Xerox	Mean	-23.48	-23.44	-23.59	-23.62	-23.58
	St Dev	0.13	0.05	0.05	0.03	0.05
Office Max	Mean	-25.41	-25.47	-25.39	-25.36	-25.50
	St Dev	0.05	0.06	0.02	0.00	0.13
Reflex	Mean	-23.55	-23.49	-23.59	-23.57	-23.66
	St Dev	0.07	0.07	0.12	0.06	0.07

Table 2.3: $\delta^{13}C_{VPDB}(\text{‰})$ results for varying sample handling/storage conditions

From the corrected sample values and the graph of the mean values, there appear to be no trends to be observed in terms of consistent effects on the samples.

The Kruskal-Wallis test followed by Dunn's post hoc test was used to determine if there were any significant differences (at the 95% confidence level) between the sample storage/handling methods. Significant differences were observed within the Paper One sample set ($p = 0.047$), but were not able to be attributed to one sample handling method using Dunn's post hoc test.

Overall, there was no significant effect observed in the results obtained from these tests. As long as the principle of identical treatment is applied to samples within the same analytical run, and that exposure to moisture is reduced where possible, the results can be considered to be stable for comparison.

2.3.4. Design of Sample Run

Number of Standards

Statistically, using a paired t-test, there were no significant differences observed between using two or three international standards to correct the values of the samples. The slope of the calibration line did not change, and its position was only slightly affected with a shift in the intercept of the line. The largest shift in the position of the intercept was by 0.07, however even this difference was not large enough to shift the corrected values of the samples, particularly given that the reported values are limited to being expressed to four significant figures. The results of this experiment are shown in Table 2.4.

Number of Replicates

When the sample data were corrected using three, five or seven replicates of the international standards polyethylene and sucrose placed at the start and the end of the analytical run, differences were detected between the mean values and the standard deviations. A summary of the corrected data is shown in Table 2.4.

The Kruskal-Wallis ANOVA test with Dunn's post hoc test was performed using GraphPad Prism. The standard materials in Table 2.4 were tested for differences between the results corrected using different numbers of replicates of international standards. The experimental work for these tests was undertaken over consecutive days to reduce any potential instrument effects that may have been generated by changing the Flash EA packing materials. The results for NBS-19 and LSVEC were suspected to pose problems as both samples are carbonates and hence are not matrix matched to the method being validated.

The alpha glucose sample corrected using three replicate international standards was found to be significantly different to the seven repeat correction group at the 95% confidence level. The p-value for alpha glucose was 0.039, which is between the 95% and 99% confidence levels. No significant difference was detected for the remainder of the samples corrected.

These results indicate that there is no benefit to running higher numbers (>3 replicates) of international standards at the start and the end of the experimental runs, for the purposes of scale normalisation. It becomes difficult, however, to ensure the quality of the calibration performed when one or more of the international standard samples is excluded during correction using Grubbs test. Meier-Augenstein (2010) recommends that any sample, standard or control that will undergo calculation of likelihood ratio's, as originally outlined in Evett et al. (2000), should be run in even numbered multiples. It makes sense therefore to run the international standards at the start and end of each run in replicates of five, to allow for flexibility, both in the exclusion of outliers and for calculation of likelihood ratio's.

		Published Value (‰)		Replicates = 3				Replicates = 5				Replicates = 7			
				Two Standards		Three Standards		Two Standards		Three Standards		Two Standards		Three Standards	
		Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Standards used for correction	Polyethylene	-32.151	0.05	-32.15	0.05	-32.15	0.05	-32.15	0.02	-32.16	0.02	-32.152	0.042	-32.163	0.040
	Sucrose	-10.449	0.033	-10.45	0.02	-10.45	0.02	-10.45	0.04	-10.45	0.04	-10.449	0.088	-10.456	0.088
	Cellulose	-24.724	0.041			-24.72	0.069			-24.72	0.051			-24.700	0.054
International Standards	NBS-19	1.95	N/A	1.76	0.306	1.76	0.306	1.82	0.128	1.82	0.128	2.010	0.020	2.010	0.029
	LSVEC	-46.6	0.2	-46.09	0.158	-46.09	0.158	-46.33	0.09	-46.34	0.094	-46.310	0.040	-46.330	0.042
	Cellulose	-24.724	0.041	-24.72	0.069			-24.71	0.05			-24.690	0.054		
Standard Materials	Medium Cellulose			-26.85	0.003	-26.86	0.003	-26.86	0.038	-26.87	0.038	-26.870	0.032	-26.880	0.032
	Cellulose Acetate			-32.02	0.008	-32.02	0.008	-31.98	0.015	-31.98	0.015	-32.040	0.037	-32.050	0.037
	Alpha Glucose			-11.42	0.256	-11.42	0.256	-11.21	0.054	-11.21	0.054	-11.090	0.026	-11.100	0.026
Correction Factors	M value			1.001		1.001		1.002		1.002		1.006		1.006	
	B value			-10.794		-10.796		-11.125		-11.128		-10.798		-10.805	

Table 2.4: $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ results for varying international standard number and replicate

Using a Third Set of Standards

A third set of the same international standards (polyethylene and sucrose) were placed in the centre of an analytical run to determine if this increased the accuracy and precision of the calibration of the unknown samples. The same experimental run was used (to reduce instrument effects) and was corrected twice – the first using only the two sets of standards at the start and the end of the run, and the second using all three sets of standards – at the start, middle and end of the sequence.

The results in Table 2.5 show that there is an effect on the accuracy of the correction of the samples but that there is no effect on the precision i.e. even though the sample values are shifting slightly, the standard deviation values are not. The magnitude of the difference with respect to the accuracy is small however and when weighed against the practicalities of including a third set of standards, the difference is not significant enough to warrant increasing a run sequence by a minimum of 16 samples (10 standards and 6 blanks) at the expense of running unknowns.

	Two Sets		Three Sets	
	$\delta^{13}\text{C}$ Mean (‰)	St Dev (‰)	$\delta^{13}\text{C}$ Mean (‰)	St Dev (‰)
Int Cellulose	-24.43	0.06	-24.48	0.06
Cellulose Acetate	-31.87	0.16	-31.96	0.16
Alpha Glucose	-10.91	0.14	-10.92	0.14
Correction Factors				
M value	1.0155		1.0196	
B value	-10.43		-10.433	

Table 2.5: $\delta^{13}\text{C}_{\text{VPDB}}$ (‰) values of samples corrected using two or three sets of international standards.

Inclusion of Blank Capsules

Blank capsules were removed from between all of the standards and samples within a run sequence and the results compared to a prior sequence.

Noticeable effects on the sample measurements were detected when moving from a low $\delta^{13}\text{C}$ value (e.g. international standard polyethylene = -32.15 ‰) to a high $\delta^{13}\text{C}$ value (e.g. international standard sucrose = -10.45 ‰). This commonly resulted in the first and second replicate being lost due to memory effect and being disregarded as an outlier.

To combat this, pairs of blank capsules were placed between the two international standards in both sets of standards (at the start and the end of the analytical run). This resulted in reducing the memory effect observed and the first measurement values could more readily be used instead of automatically being excluded as outliers.

Where unknown sample values are expected to change (e.g. when measuring two different materials in the same sequence), a pair of blank capsules should be inserted between the samples. The use of blank capsules was observed to be of benefit when the unknown samples were expected to be within close proximity to one another (within approximately 10 ‰). This means that for an analytical sequence consisting only of paper samples, no blanks are required between the unknown samples. If an analytical sequence contains both paper and glucose/sucrose samples however, a pair of blanks should be inserted between the different material types.

Use of a Quality Control Check

During this set of experiments it became apparent that it was difficult to know when an experimental run was fit for use. It is recommended that either:

- The international standard IAEA-CH-3 be included within the sequence as an unknown to check the accuracy of the instrument and scale normalisation applied;
or
- A suitable laboratory standard material is calibrated for use as an unknown to check against a laboratory defined consensus value.

Overall, the quality control check should be matrix matched to the target sample and should lie close to the expected value of the other unknown samples i.e. cellulose for cellulose or paper samples, a glucose for sugars.

2.4. Results and Discussion – Method Validation

2.4.1. Precision and Stability

Mean corrected measurements for the polyethylene and sucrose international standards are shown in Figure 2.12 and Figure 2.13. The mean value (solid line) and 2 x standard deviations (dashed lines) are represented on each plot, with 2 x the standard deviation the 95% confidence interval for each standard. Any values considered outliers within the original analytical run were excluded. Table 2.6 is a summary of the individual measurements, mean, standard deviation, 95% confidence interval (95% CI) and number of measurements relating to the data represented for these standards.

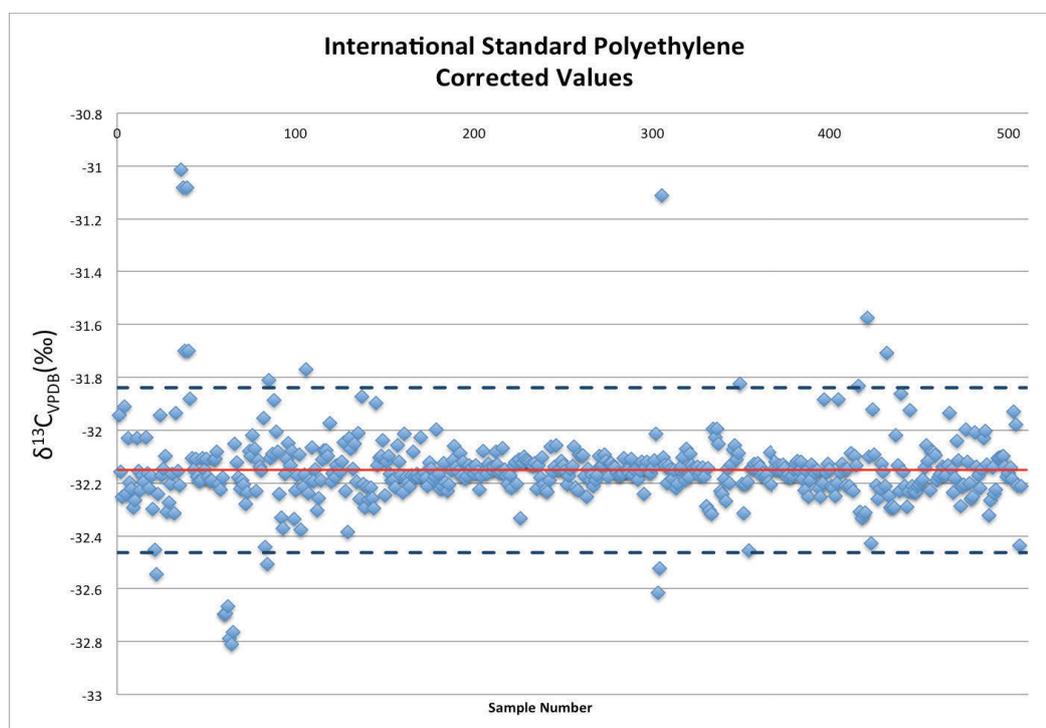


Figure 2.12: Plot of corrected measurements of international standard polyethylene over a 24-month period.

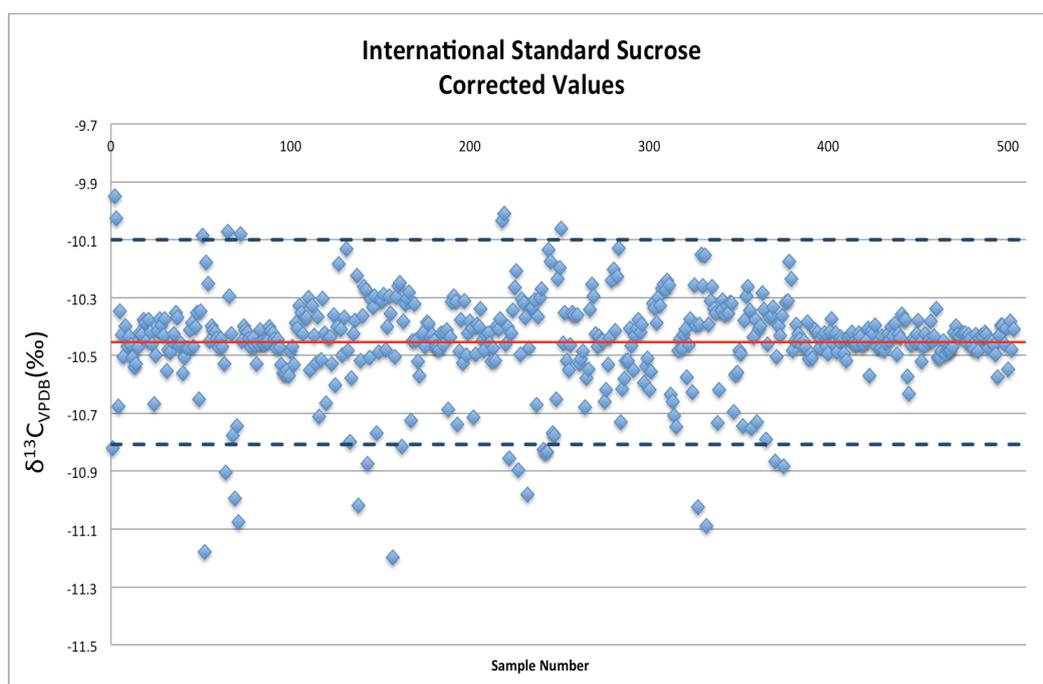


Figure 2.13: Plot of corrected measurements of international standard sucrose over a 24-month period.

	$\delta^{13}\text{C}$ Mean (‰)	1 x St Dev (‰)	95% CI (‰)	Number Samples	$\delta^{13}\text{C}$ Published Value (‰)	Published St Dev (‰)
Polyethylene	-32.15	0.15	0.30	517	-32.15	0.05
Sucrose	-10.45	0.16	0.32	488	-10.45	0.03
Cellulose	-24.69	0.12	0.24	129	-24.72	0.04

Table 2.6: Summary of the analytical results obtained for three international standard materials.

It is important to note with this data that routine maintenance (changing of packing materials, gas tanks and drying agents) and servicing of the IRMS was conducted on numerous occasions during this period. The polyethylene and sucrose measurements here show no issues with precision and the values are stable over time. The corrected mean values are comparable to the published values overall. The standard deviations are close to or below the published expected precision of measurement of 0.15 ‰ for carbon measurements (Brand, 1996).

2.4.2. Repeatability

The international standard cellulose IAEA-CH-3 was used as a quality control standard within analytical runs. The corrected measurements for international standard cellulose as measured over a 12-month period are shown in Figure 2.14. The mean value (solid line) and the 95% confidence interval (dashed lines) are included on the plot. Table 2.6 summarizes the individual measurements, mean, standard deviation, 95% confidence interval (95% CI) and number of measurements relating to the data represented for cellulose.

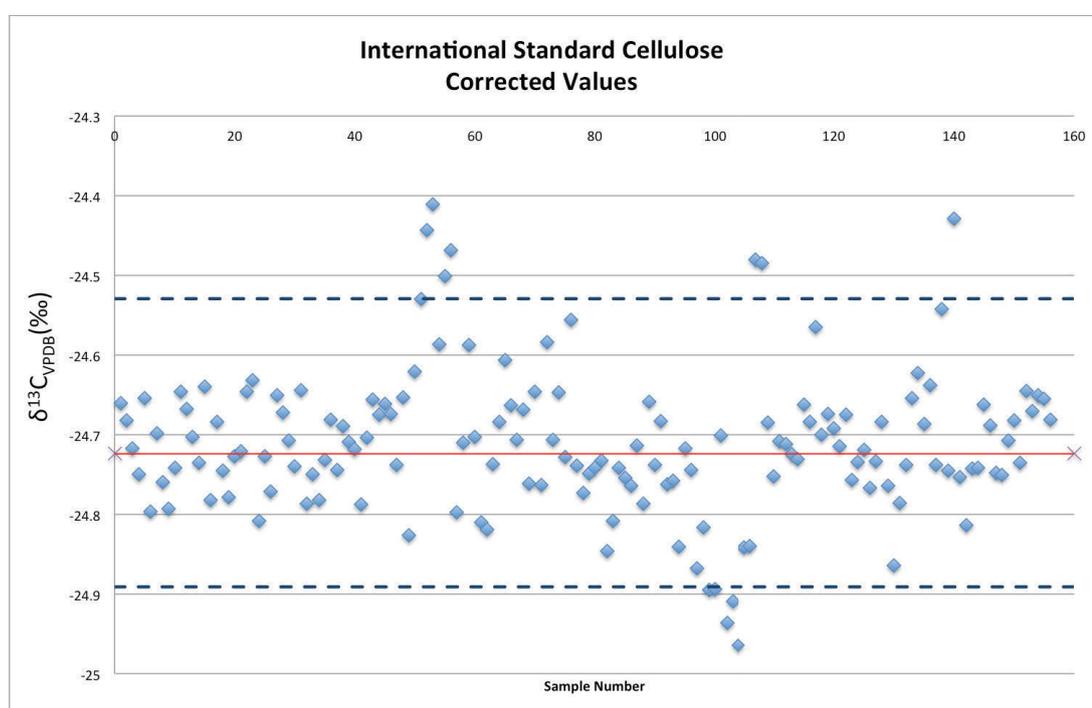


Figure 2.14: Plot of corrected measurements of international standard cellulose over a 24-month period.

2.4.3. Accuracy

Three materials being considered for development as laboratory/working standard materials – cellulose acetate, medium fibre cellulose and alpha glucose – were measured over 18 months and plotted to determine the accuracy of the method in producing comparable results over time. Figure 2.15, Figure 2.16 and Figure 2.17 show the mean measurement for each run, with the error bars denoting the 95% confidence interval (2 x standard deviations) for that sample set. Table 2.7 summarizes

this data. Overall, each material was found to have a standard deviation less than 0.15 ‰ which is less than the error of the instrument for carbon measurements (0.2 ‰).

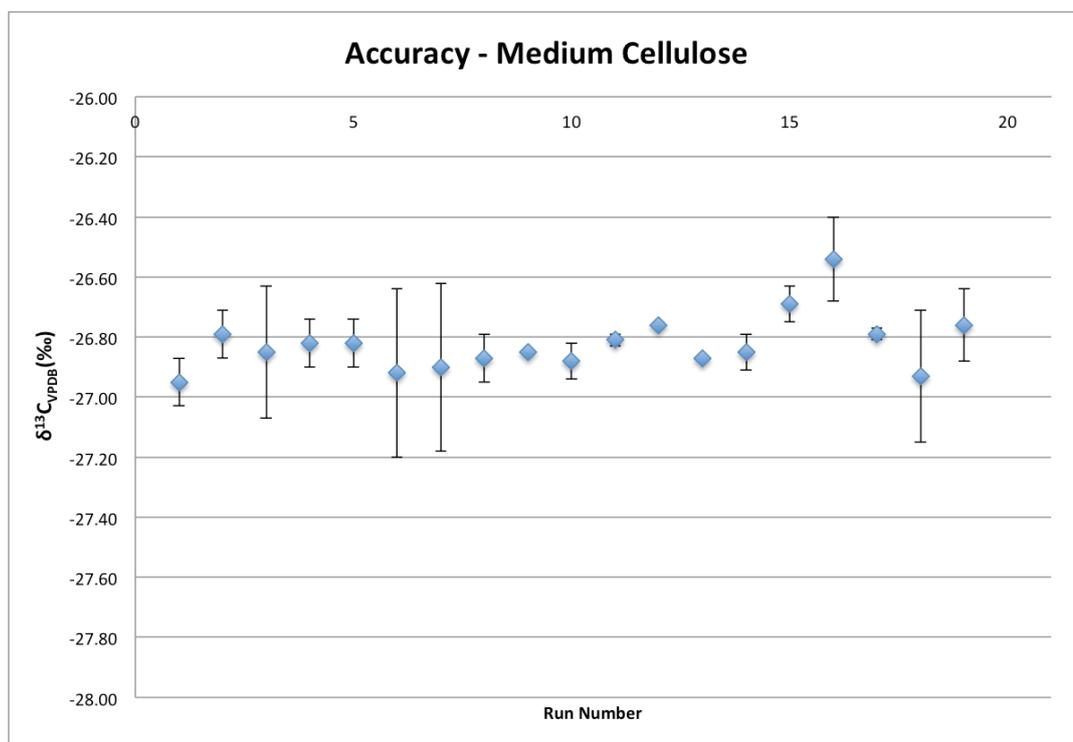


Figure 2.15: Accuracy results for Medium Cellulose

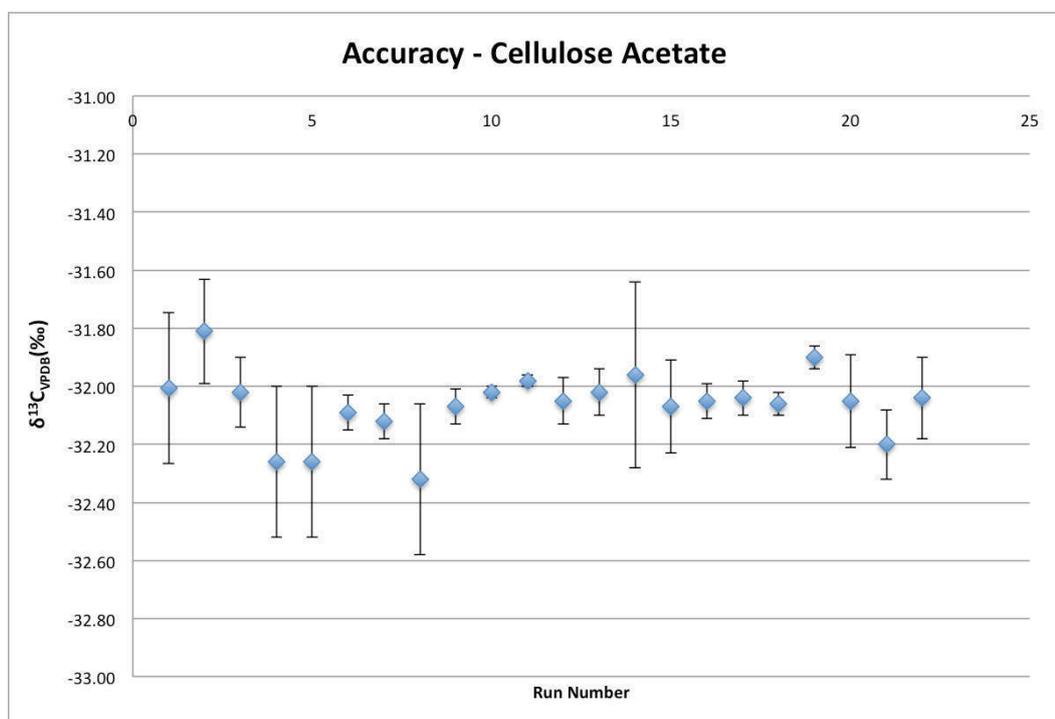


Figure 2.16: Accuracy results for Cellulose Acetate

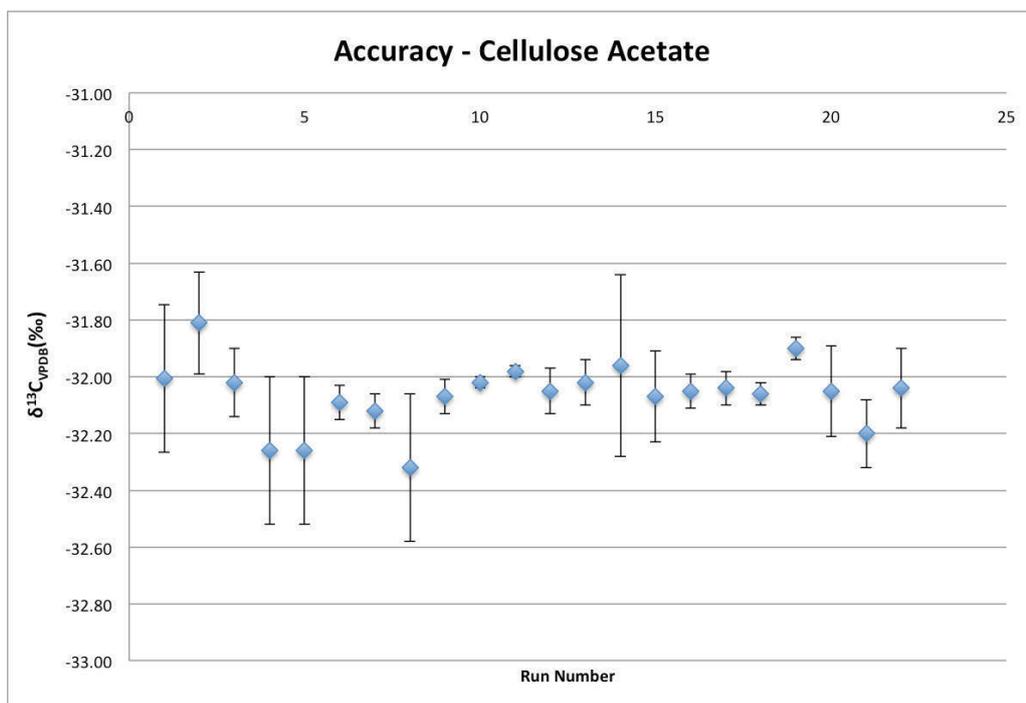


Figure 2.17: Accuracy results for Alpha Glucose

	$\delta^{13}\text{C}$ Mean (‰)	St Dev (‰)	95% CI (‰)	Number of Samples
Medium Cellulose	-26.82	0.09	0.19	63
Alpha Glucose	-11.15	0.14	0.28	94
Cellulose Acetate	-32.06	0.12	0.23	99

Table 2.7: Summarized results from accuracy experiments of cellulose acetate, medium cellulose and alpha glucose.

2.4.4. Robustness

Operators A, B and C are different operators who prepared replicates of the same sequence of samples. Operator D is where operator C has prepared, run and corrected an additional set of the same samples. The mean results have been plotted in Figure 2.18. Figure 2.8 summarizes these results.



Figure 2.18: Mean $\delta^{13}\text{C}_{\text{VPDB}}$ (‰) values obtained for samples prepared and run by three different operators.

		Operator A	Operator B	Operator C	Operator D
International Cellulose	$\delta^{13}\text{C}$ Mean (‰)	-24.59	-24.55	-24.61	-24.47
	St Dev (‰)	0.13	0.28	0.14	0.16
Cellulose Acetate	$\delta^{13}\text{C}$ Mean (‰)	-31.9	-32.2	-32.05	-32.04
	St Dev (‰)	0.02	0.07	0.08	0.07
Alpha Glucose	$\delta^{13}\text{C}$ Mean (‰)	-11.05	-11.05	-11.20	-11.33
	St Dev (‰)	0.02	0.04	0.16	0.36
Medium Cellulose	$\delta^{13}\text{C}$ Mean (‰)	-26.54	-26.93	-26.8	-26.76
	St Dev (‰)	0.07	0.11	0.01	0.07
Paper	$\delta^{13}\text{C}$ Mean (‰)	-27.48	-27.61	-27.56	-27.61
	St Dev (‰)	0.08	0.06	0.01	0.1
ANU Sucrose	$\delta^{13}\text{C}$ Mean (‰)	-11.42	-11.50	-11.48	-11.59
	St Dev (‰)	0.04	0.04	0.13	0.16

Table 2.8: Mean $\delta^{13}\text{C}_{\text{VPDB}}$ (‰) and standard deviation for samples prepared by 3 different operators

The measurement results shows that while all four analytical runs produced comparable results for each measurement, small differences were observed between the standard deviations of the measurement results. The plotted results in Figure 2.18 indicate that the differences observed were not consistent across all sample types however, suggesting that the differences are due to the natural range of the material or the instrument rather than a methodology or sampling error.

The Kruskal-Wallis test for the comparison of the medians of non-parametric populations was performed on these values to identify whether the differences observed between the samples prepared by different operators were significant. Dunn's post hoc test was used to identify which groups were causing the differences. A non-parametric test was utilised as paper was included as an unknown material in these experiments.

Significant differences (to 95% confidence level) were observed between:

- Operators A and B for the Cellulose Acetate samples (actual difference of 0.3 ‰)
- Operators A and B for the Medium Cellulose samples (actual difference 0.39 ‰)

The actual difference in the $\delta^{13}\text{C}$ values for these two samples could be accounted for in the natural variation of the materials, the measurement uncertainty associated with the method or due to some instrumental difference. The last of these options is the more likely given that the instrument is known to be quite variable when technical issues arise. Although statistically these values are different, with regard to the measurement uncertainty of the technique, these differences are not large and are not indicative of a systemic issue with the method.

2.4.5. Measurement Uncertainty

The expanded uncertainty for the measurement of carbon isotope ratios was calculated using the Guide to Expressing Uncertainty in Measurement (GUM) approach (NMI, 2011), technical guidelines produced by NATA (National Association of Testing Authorities, April 2009b) and the equations outlined in Benson et al (Benson et al., 2010).

Using the precision experiments conducted on the potential laboratory standards medium cellulose and cellulose acetate, the expanded uncertainty for the measurement of carbon isotope ratios in cellulose was determined to be 0.26 ‰ (with a coverage factor of 2). The bias (calculated using the precision results for international cellulose) was found to be insignificant.

Likewise, using the precision experiments conducted on alpha glucose, the expanded uncertainty for the measurement of carbon isotope ratios in glucose was determined to be 0.3 ‰ (with a coverage factor of 2).

2.4.6. Inter-laboratory Trial

The 2010 Forensic Isotope Ratio Mass Spectrometry (FIRMS) group trial was undertaken using the experimental method and correction calculations described. Two types of samples were measured and run and the $\delta^{13}\text{C}$ values for 10 replicates of each sample were reported to be included as part of this trial.

For the two materials measured (Glycine and 4-Nitroacetanilide), the z-scores received were within the 2.0 value pass mark, showing that this method produces results that are in consensus with other laboratories.

2.5. Conclusions

Full method validation for $\delta^{13}\text{C}$ to international standards was achieved. The overall measurement uncertainty was calculated to be approximately 0.26 ‰ (with a coverage factor of 2). All performance characteristics including precision, accuracy, repeatability and robustness were within the 0.26 ‰ expanded measurement uncertainty range.

Control of atmosphere/water content was shown to be required for paper samples, even for carbon isotopes, which are generally assumed to be free from water absorption effects.

Chapter 3

3. Carbon Isotopes – Background Population and Homogeneity

The key results of this chapter have been published in Jones et al. (2013a). The published article is included electronically within Appendix 5.

3.1. Introduction

Theoretically, measuring two chemically identical materials and comparing their isotopic values for similarities or differences is a relatively straightforward task. However, a number of questions are raised when it comes to the interpretation phase of this task. These questions are generally related to the variability of the material itself and relating those properties to form a procedure for when it is appropriate to say that a material has originated from a different source. For chemically homogenous materials (e.g. chemically manufactured polymers such as polyethylene) the inhomogeneity of a material from the same source may be lower than the instrumental uncertainty. For naturally derived materials such as paper however, little work has been done with respect to the expected range of variation from material being produced within the same source. While ecological studies, such as those discussed in Chapter 1, are useful as a basis, it cannot be assumed that the same variability would apply to papers due to the significant chemical and physical processing required to produce papers.

Likewise, while there are some challenges with respect to being able to claim that two samples have originated from different sources, there is a greater challenge and need for data to support any examination finding that claims that two samples may have originated from the same source. Firstly, there is a need to determine what the word ‘source’ actually means for a material – does it relate to a singular package, a batch or a certain period of time? Does source relate to that one particular production source (manufacturing facility) or does it refer to the brand, since it could be reasonably assumed that the same starter materials and production processes are likely to be used

by a company with more than one paper making facility. Secondly, is it possible to make judgements about the similarity of a material and are the isotopic abundance values individual enough to be able to make a distinction between a particular source and all other sources on the market?

This chapter aims to build a background population of the carbon isotopes of document papers purchased in the Australian and New Zealand market. To answer a number of the questions posed above, this chapter will also examine the homogeneity of white A4 80gsm papers as a material – from within and between sources. The information gained from these experiments can be used to inform the comparison and opinion formulation process.

3.2. Materials and Methods

The instrumentation, equipment and method used to measure and correct the carbon isotopic abundance values in this chapter are outlined in Chapter 2.

Statistical tests were undertaken using Microsoft Excel® and GraphPad Prism version 5. When assessing whether differences were being observed, the Kruskal-Wallis Non-Parametric Analysis Of Variance (ANOVA) followed by a Dunn's Multiple Comparison Post-Hoc test was used. A non-parametric test was selected as some papers were observed to be non-parametric, even when measured in large number of replicates (n=50) and hence normality could not be assumed for these samples.

To determine the discrimination power (Smalldon and Moffat, 1973), defined as the number of discriminated pairs/total number of pairs of the measured $\delta^{13}\text{C}$ values, Microsoft® Excel was used to perform 7750 pairwise comparisons of the 125 population samples. The benchmark value used for discrimination was determined using the homogeneity experiments below and a "COUNTIF" formula was used to determine if there was a difference between the two measured values.

3.2.1 Samples and Standards

For the purposes of the background population, all paper samples purchased and measured were marketed as white A4 80gsm office papers. Each of the reams consisted of 500 sheets excluding one brand that contained only 250 sheets. For this brand, two reams with the same date stamp were utilised to ensure sampling consistency. 82 reams of paper were collected from common commercial outlets in the Australian Capital Territory and Queensland, Australia over a period of 24 months. These consisted of 38 different brands of papers. In addition, 43 paper samples were collected from printers across Australia and New Zealand, representing an additional 8 unique brands to the ones previously collected. In total, 125 individual paper samples were measured. All of these samples were marketed as virgin pulp, or did not state that they contained any recycled fibre content.

It is important to note that there is only one manufacturer of white A4 80gsm office paper in Australia and that any paper that has been produced nationally originates from the same mill (Australian Paper, Maryvale VIC). This mill however, does produce different grades of paper that vary in fibre and filler composition. All virgin paper samples, their countries of origin and packaging date (where known) are detailed in Table 14.3.

As a preliminary comparison, 7 reams of recycled paper produced in Australia were measured to compare their $\delta^{13}\text{C}$ values against the Australian virgin paper background samples. 5 samples were marketed as containing 10% recycled pulp content while the other 2 samples contained 50% recycled, post consumer waste pulp. These papers are detailed in Table 3.2.

For homogeneity testing, 7 brands of paper that varied in their production location were selected for testing. 8 reams were purchased from each brand, with care taken to ensure that at least 7 reams from each brand were purchased at the same time/location and if possible, containing the same production markings. The brands selected for testing are detailed in Table 3.1.

Brand Name	Country of Origin
Reflex	Australia
Olympic	Australia
Officemax	Sth Africa/China
Docucopy	China
Fuji Xerox	China
Lazer IT	Indonesia
Double A	Thailand

Table 3.1: Brand and origin of paper samples used for homogeneity studies.

To test the homogeneity of the reams, two experiments were conducted with different sampling methods used for each. In the first experiment, one ream from each brand was taken and sampled in the following ways (with ‘one location’ denoting replicate samples taken from the same physical area and ‘random location’ denoting replicate samples taken from across the sheet of paper):

- 3 samples from one page, in one location
- 3 samples from one page, in random locations
- 7 samples from one page, in one location
- 7 samples from one page, in random locations
- 7 samples, one each from 7 different pages
- 25 samples, one each from 25 different pages (representing 5% of the ream)

This protocol was utilised so that in total over 5% of the ream was measured. In addition, sampling in this manner allowed comparison of the variance that may have been contributed by the sampling method itself, so that the most appropriate method for both reams and single sheets could be identified for practical application in casework situations.

In the second round of homogeneity testing, 7 reams from the 7 brands in Table 3.1 were sampled by taking 7 pages, manually selected at random, from throughout each of the reams. From these reams, one sample per page was taken. The decision to collect 7

samples to characterize each individual ream was informed by the results from the first round of homogeneity testing.

As an additional test, homogeneity within the same brand over time was tested using the paper background population samples. Two brands – Reflex and Double A, were collected over a 24-month period and compared to determine the difference in $\delta^{13}\text{C}$ over time. These brands were selected for this comparison, as they were the only two brands that consistently labelled their reams with the packing date. This gave a clearer indication of the time that the paper was produced rather than just using the collection date to infer production date. For comparison, 8 groups from both the Reflex and the Double A samples detailed in Table 14.3 were compared using the Kruskal-Wallis ANOVA with Dunn's post-hoc test.

3.3. Results and Discussion

3.3.1 Background Population Study – Carbon Isotopes

For the 125 virgin paper samples measured, the range of $\delta^{13}\text{C}$ was between -22.5 ‰ and 30 ‰. This range is shown in Figure 3.1, with measured $\delta^{13}\text{C}$ mean and standard deviation values summarised in Table 14.3. When organised by their region of production (obtained from the ream packaging), Figure 3.1 shows that there is a relationship between the $\delta^{13}\text{C}$ and the papers' region of production.

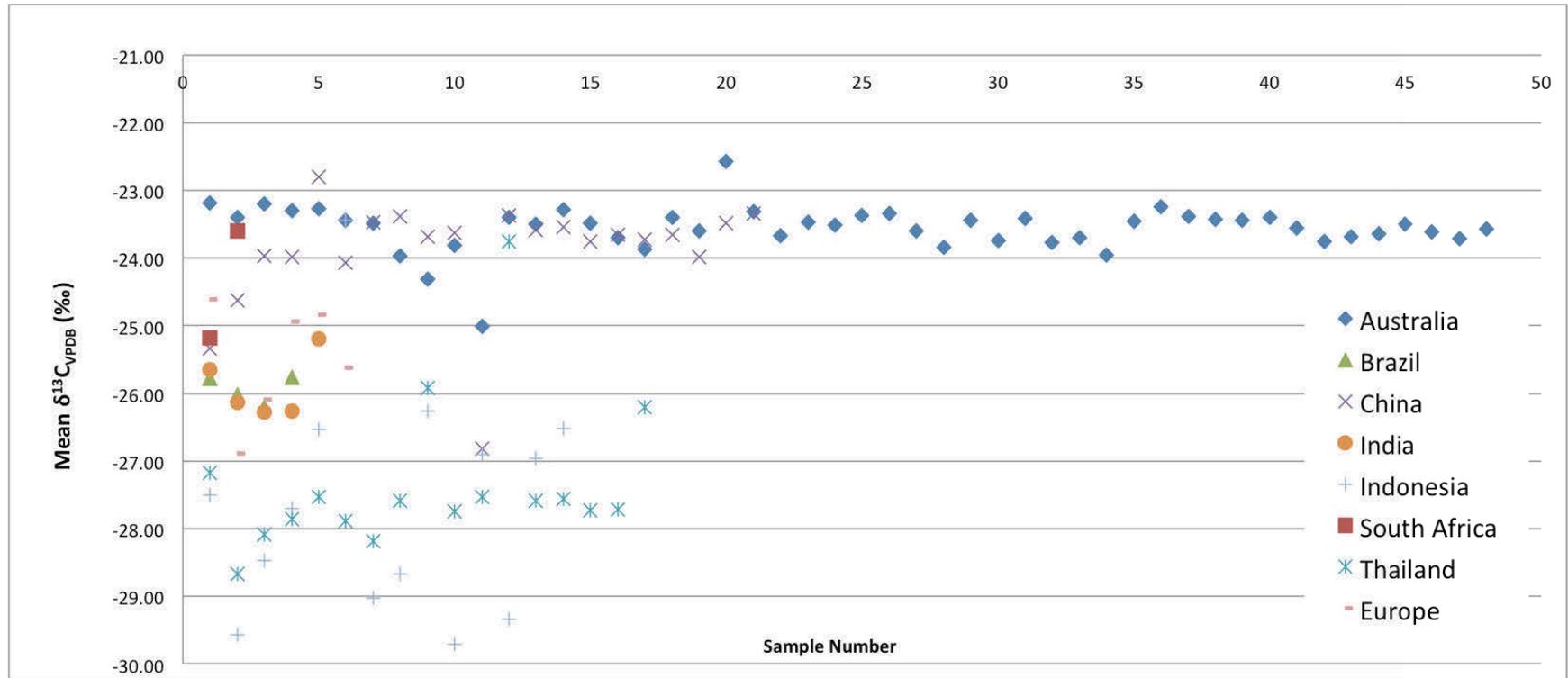


Figure 3.1: Mean $\delta^{13}\text{C}_{\text{VPDB}}$ (‰) values of 125 virgin pulp papers plotted by region of origin

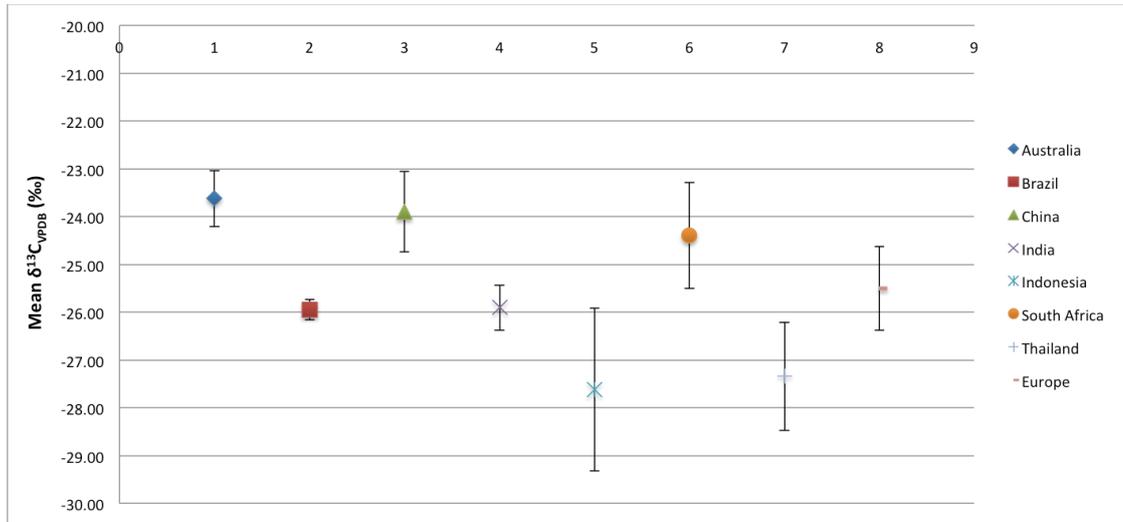


Figure 3.2: Summary of mean $\delta^{13}C_{VPDB}$ (‰) for region of origin of papers, with error bars denoting 1 x standard deviation

When organised by region of production, as shown in Figure 3.1, it becomes clear that the $\delta^{13}C$ value of the paper is likely to be due to the selection of species and growing region for the source cellulose material used in production. In particular, papers that have originated from Australia are consistently found within the -23 ‰ to -24 ‰ range, with only 3 samples falling outside of this. Papers that have been produced in China are also consistently observed within this range and this is due to Australia supplying wood chips of the same species and from the same growing location for China’s paper industry. (<http://www.macquarie.com.au/dafiles/Internet/mgl/au/advisers/campaign /static-file/independent-expert-report.pdf>)

The papers that have been produced in Thailand are similarly consistent in their $\delta^{13}C$ value, with all but 3 values measured in the -27 ‰ to -29 ‰ range. Indonesian papers by comparison are not as consistent and have a larger range between -26 ‰ to -30 ‰, with 3 samples measured outside of this range. The wider range may be due to the use of mixed wood sources, including the use of straws and tropical hardwoods rather than the use of relatively clean paper pulps used in other countries.

The samples measured here are representative of the Australian market and if examined in that context, it could be claimed that the $\delta^{13}C$ measured is indicative of the

region of production of the paper. More widely however, it would be unreasonable to assume that the regions defined here would be unique internationally. The region specificity observed in this study however, could be used for intelligence purposes as long as the assumption that the paper has originated from the Australian market is stated as part of the information provided. A broader background study to determine the true international variability of office paper would have to be undertaken to define background variability and the uniqueness of the growing regions here to remove this caveat.

The $\delta^{13}\text{C}$ of 7 samples of recycled paper produced in Australia were measured to compare against the virgin paper samples. Their mean and standard deviation values are detailed in Table 3.2. Given their low recycled content, papers with 10% recycled content were expected to be closer to the range of virgin pulp Australian papers (-23 ‰ to -24 ‰) than the 50% recycled content papers.

Overall, four of the 10% recycled content samples and one 50% recycled content sample were found to lie within the expected range of values for Australian samples. The two other samples were measured only 0.25 ‰ and 0.14 ‰ outside of the range for the other 10% and 50% recycled content samples respectively. Although further work to characterize recycled papers is required before forensic casework implementation, these preliminary results show comparison of recycled papers is still possible, with the measured values showing that there is no substantial increase in the standard deviation (and hence variability) observed. This means that recycled fibre content may not contribute as much of an effect on the $\delta^{13}\text{C}$ values of the samples than first hypothesized, particularly when the recycled content percentage is low.

Brand Name	Published % Recycled Content	Mean $\delta^{13}\text{C}_{\text{VPDB}}$ (‰)	Standard Deviation (‰)
Australian	10%	-24.25	0.05
Australian	10%	-23.66	0.08
Australian	10%	-23.40	0.07
Australian	10%	-23.82	0.19
Australian	10%	-23.40	0.05
EXPGreen50	50%	-23.76	0.21
Reflex Recycled	50%	-24.14	0.03

Table 3.2: Recycled papers, mean $\delta^{13}\text{C}_{\text{VPDB}}$ and standard deviation for 7 recycled papers collected as samples for comparison to virgin paper background samples

Australian papers within the background population were plotted to observe whether there were any trends in the $\delta^{13}\text{C}$ values over time. These differences may have been due to seasonality (growing seasons), manufacturing changes or source changes. Figure 3.3 shows that there are no repeatable observable patterns in the data and hence changes due to season of production cannot be claimed. There are some observable differences that may be due to source changes however. These samples in this time period (around sample numbers 10-20) were collected in the first half of 2010.

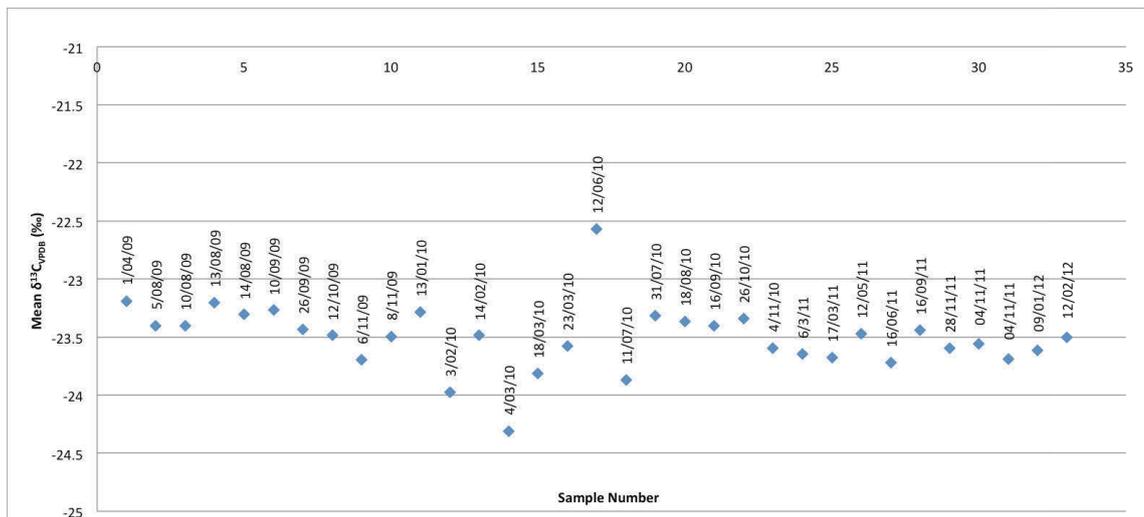


Figure 3.3: Australian papers plotted over time, according to ream packaging date (earliest to latest date).

3.3.2 Paper Homogeneity Within One Ream

The mean and standard deviation results for the paper homogeneity measurements taken within a single ream are detailed in Table 3.3. The sample names have been abbreviated where S = sample, P = page and L = location.

These homogeneity results showed that there are differences in the standard deviation values when between brands of paper. Figure 3.4 shows a plot of the mean values for each of the reams tested b each of the sampling methods. Even though there is variability, each of the reams produced reproducible results that sit within a tight range for each sample type.

		1. DOUBLE A	2. OFFICEMAX	3. REFLEX	4. FUJI XEROX	5. LAZER IT	6. DOCUCOPY	7. OLYMPIC
Mean $\delta^{13}\text{C}_{\text{VPDB}}$ of Paper Samples (‰) $\pm 1\sigma$	3S_1P_1L	-27.46 \pm 0.16	-25.72 \pm 0.06	-23.69 \pm 0.07	-23.77 \pm 0.03	-27.11 \pm 0.13	-22.73 \pm 0.21	-23.1 \pm 0
	7S_1P_1L	-27.03 \pm 0.16	-25.8 \pm 0.09	-23.67 \pm 0.07	-24.0 \pm 0.12	-27.11 \pm 0.1	-23.11 \pm 0.12	-23.27 \pm 0.08
	3S_1P_1R	-26.95 \pm 0.05	-25.72 \pm 0.04	-23.75 \pm 0.1	-23.99 \pm 0.03	-27.06 \pm 0.07	-22.86 \pm 0.01	-23.21 \pm 0.06
	7S_1P_1R	-27.29 \pm 0.08	-25.83 \pm 0.08	-23.88 \pm 0.07	-23.58 \pm 0.07	-27.16 \pm 0.04	-22.95 \pm 0.1	-23.29 \pm 0.08
	7S_7P	-27.47 \pm 0.09	-25.69 \pm 0.09	-23.96 \pm 0.16	-23.54 \pm 0.07	-27.18 \pm 0.12	-22.75 \pm 0.11	-23.31 \pm 0.05
	25S_25P	-27.62 \pm 0.12	-25.57 \pm 0.12	-24.08 \pm 0.1	-23.91 \pm 0.15	-23.31 \pm 0.15	-23.04 \pm 0.2	-23.33 \pm 0.09
Total # samples		52	51	51	50	50	51	49
Grand Mean		-27.43 \pm 0.25	-25.67 \pm 0.15	-23.99 \pm 0.28	23.83 \pm 0.21	-27.22 \pm 0.15	22.98 \pm 0.2	-23.30 \pm 0.09
95% confidence interval		0.5	0.3	0.55	0.41	0.3	0.41	0.18

Table 3.3: Measured $\delta^{13}\text{C}_{\text{VPDB}}$ (‰) abundances from single ream homogeneity testing

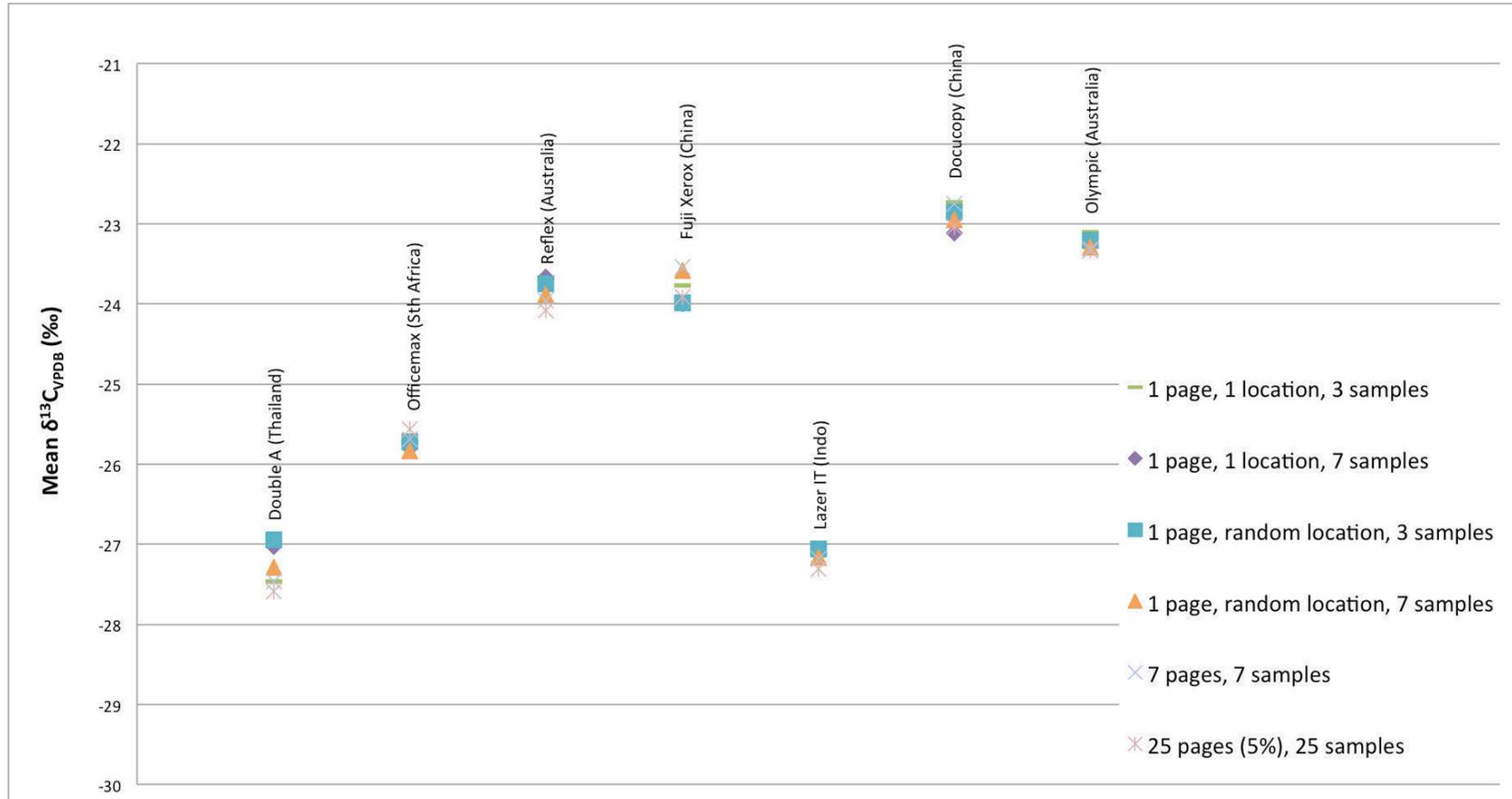


Figure 3.4: Mean $\delta^{13}\text{C}_{\text{VPDB}}$ (‰) values for seven different brands measured during homogeneity and sampling tests

For discrimination, it is recommended that 95% confidence interval results from the single ream homogeneity studies be used first to inform the range of values that a sample could potentially fall within when deciding whether two samples can or cannot be differentiated. Additionally, a stronger level of discrimination can be called using a standardised benchmark/ discrimination value instead of a 95% confidence interval of the paper sample measured, to prevent the false exclusion of samples that have originated from the same source. The homogeneity results from the single and multiple ream studies

Excluding cases where questions are asked about two halves of a single sheet, forensic casework will generally be related to either a partial ream or stack of paper being compared to one or more single sheets, or to page substitutions in multiple page documents. Taking the Fuji Xerox data in Table 3.3 as an example, if we were to compare the single sheet 3S_1P_1L value ($-23.77 \pm 0.03 \text{ ‰}$) to the seven page stack 7S_7P value ($-23.54 \pm 0.07 \text{ ‰}$), using a 95% confidence interval defined by the standard deviation of either of these measurements, we would incorrectly discriminate them as having originated from a different source (difference = 0.23 ‰). This is highlighted visually in Figure 3.5, which also demonstrates that this not be the only comparison pair from this ream that would be discriminated.

To overcome this, it is proposed that an experimentally defined benchmark/discrimination value be used. This would be selected based on the natural variation (and inherent inhomogeneity) of paper, and would lead to a more robust decision point for discrimination and a higher confidence in the results being reported.

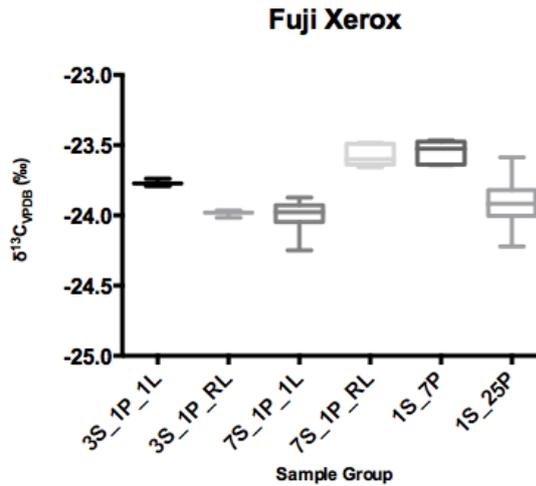


Figure 3.5: Boxplot of Fuji Xerox $\delta^{13}\text{C}_{\text{VPDB}}$ results taken from a single ream. Error bars shown represent 95% range for discrimination.

To set a benchmark/discrimination value from the data Table 3.3, a 95% confidence interval of the largest range measured in one ream is proposed – in this case 0.55 ‰ (or a range of 1 ‰ in total). Figure 3.6 is a plot of the standard deviations measured for each different sampling test. A standard deviation for the grand mean is also included (where the grand mean is the average of all measurements within the ream), as is the calculated measurement uncertainty of the method (0.26 ‰). Six of the seven reams' standard deviations lie below the method measurement uncertainty, indicating that this value is a robust estimate of the maximum variance expected.

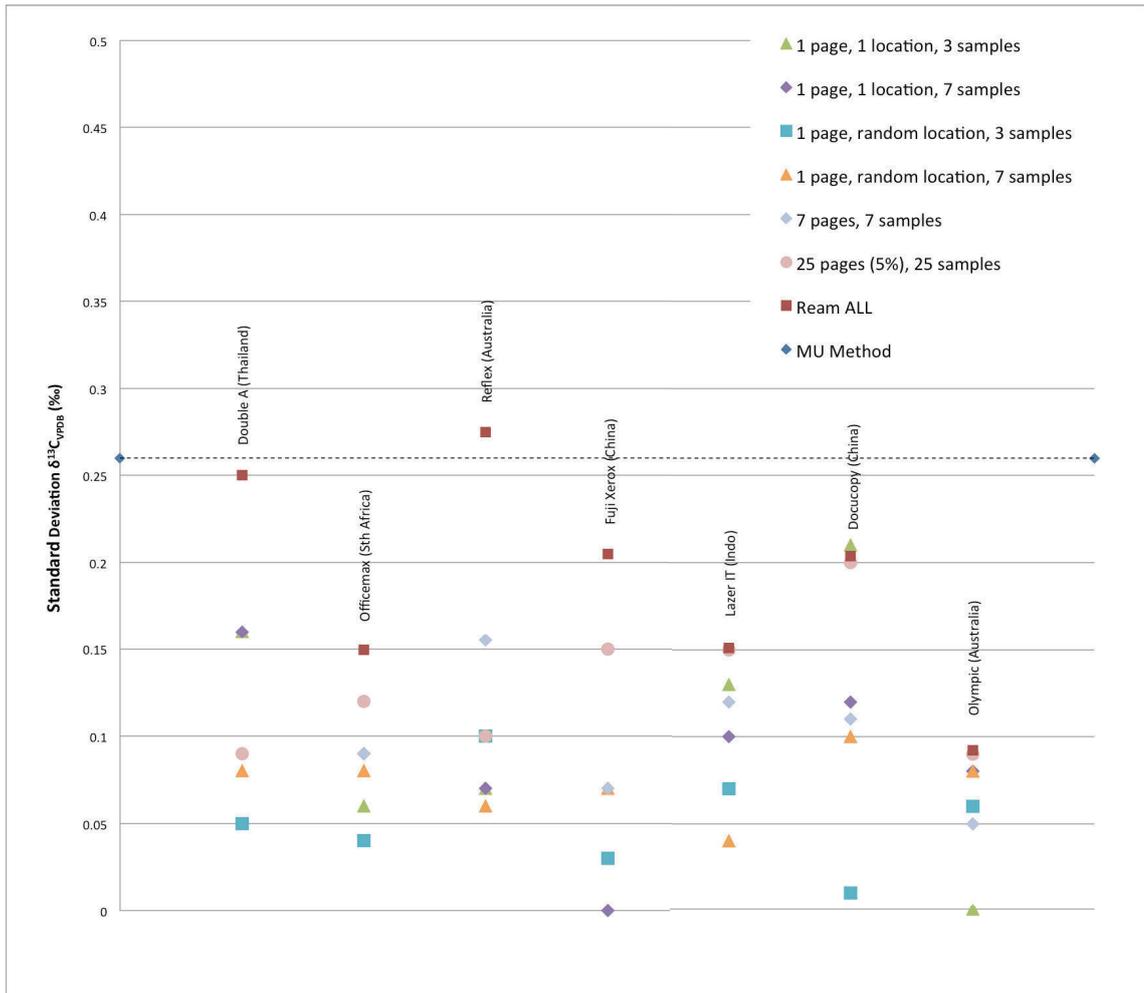


Figure 3.6: Standard deviation values (%) for homogeneity and sampling test

To determine whether one sampling method was representative of the variability within a single ream, a non-parametric Kruskal-Wallis ANOVA test was utilised for comparison. Significant differences were detected between the sampling methods in all reams. The p-values ranged from 0.0449 (Olympic) to <0.0001 (Double A, Fuji Xerox, Officemax and Reflex). Dunn's post-hoc test was used to cross compare each of the sampling methods. If one sampling method was observed to be consistently different to other sampling methods (across the seven brands) then this would indicate that this sampling method was more able to capture the variability of the ream. No sampling group was observed to be consistently different across all of the reams and hence it is suggested that the variance measured is random and likely due to the natural inhomogeneity of the paper.

With respect to ream sampling methods (comparing the standard deviation results in Table 3.3) it is evident that the largest standard deviation was observed in the 25 page results. For all reams however, this value is close to that of the 7 page results and using Dunn's Post Hoc test, only one ream was statistically significant different between these two groups (Fuji Xerox). This result does not justify the need for the additional time and resources required to sample 25 pages rather than 7. Based on this, for forensic casework, it is recommended that at least 7 samples from 7 pages be used to characterize a ream or stack of paper (e.g. a stack collected from a printer).

Comparing the random location sample groups to the single location groups, there is evidence to suggest that sampling from a single location captures more of the variability within a sheet than sampling randomly. This is seen in the higher standard deviation results for these groups (for both the 3 sample and the 7 sample tests), particularly for the Docucopy and Double A reams. Overall however, the standard deviations are relatively small and one method does not clearly produce consistently better results across all brands.

Mixed results were obtained for the number of samples needed to capture the variability of a single sheet of paper. Comparing the 3 sample tests to the 7 sample tests (from both the random location and one location tests), the results indicate that analysing 7 samples leads to smaller standard deviation results. Practically, this sampling method would also be favoured in terms of outlier removal. Again however, the standard deviation values measured here are small, as are the differences between them and hence the number of

samples taken to characterise a single sheet should be a decision taken within the context of the examination question.

3.3.3 Homogeneity Between Reams of the Same Brand

While the background population study previously described examined changes within the same brands over time, to determine the ability of IRMS to discriminate source and to determine what the source level for papers may be, seven reams from seven brands were purchased at the same time and from the same location. A single sample from seven sheets from each ream was measured and the mean and standard deviation results are detailed in Table 3.4.

It should be noted that between the round one and round two homogeneity testing, the Officemax brand changed their production location from South Africa to China.

	Ream	1. DOUBLE A	2. OFFICEMAX	3. REFLEX	4. FUJI XEROX	5. LAZER IT	6. DOCUCOPY	7. OLYMPIC
Mean $\delta^{13}\text{C}_{\text{VPDB}}$ of Paper Samples (‰) $\pm 1\sigma$	A	-27.2 \pm 0.09	-23.49 \pm 0.07	-23.58 \pm 0.1	-23.39 \pm 0.16	-26.86 \pm 0.06	-23.04 \pm 0.27	-23.28 \pm 0.06
	B	-27.22 \pm 0.05	-23.56 \pm 0.08	-23.56 \pm 0.08	-23.54 \pm 0.13	-26.96 \pm 0.1	-23.38 \pm 0.32	-23.33 \pm 0.10
	C	-27.23 \pm 0.08	-23.58 \pm 0.16	-23.6 \pm 0.09	-23.53 \pm 0.13	-27.01 \pm 0.09	-23.08 \pm 0.17	-23.33 \pm 0.08
	D	-27.24 \pm 0.03	-23.49 \pm 0.08	-23.56 \pm 0.05	-23.51 \pm 0.14	-26.94 \pm 0.09	-23.34 \pm 0.28	-23.27 \pm 0.12
	E	-27.25 \pm 0.05	-23.72 \pm 0.16	-23.47 \pm 0.07	-23.46 \pm 0.13	-26.97 \pm 0.06	-23.12 \pm 0.17	-23.31 \pm 0.11
	F	-28.05 \pm 0.10	-23.18 \pm 0.08	-23.69 \pm 0.15	-23.47 \pm 0.10	-27.0 \pm 0.08	-23.25 \pm 0.09	-23.35 \pm 0.10
	G	-28.02 \pm 0.06	-23.21 \pm 0.12	-23.65 \pm 0.07	-23.56 \pm 0.10	-26.99 \pm 0.09	-23.04 \pm 0.15	-23.33 \pm 0.04
Total N		46	46	49	49	46	47	48
Grand Mean		-27.47 \pm 0.38	-23.47 \pm 0.21	-23.59 \pm 0.11	-23.49 \pm 0.13	-26.97 \pm 0.09	-23.16 \pm 0.24	-23.32 \pm 0.09
95% Confidence Interval		0.76	0.42	0.22	0.26	0.18	0.49	0.18

Table 3.4: Mean $\delta^{13}\text{C}_{\text{VPDB}}$ and standard deviation (‰) results for between ream homogeneity

Overall, the brand with the largest variance per ream was Docucopy. The Fuji Xerox and Officemax brands also had higher standard deviations across multiple reams than the other brands.

Each brand was plotted as a box plot, with the whiskers set to represent the 95% confidence interval (2 x standard deviations) for each population. The seven plots are shown below in Figure 3.7 (a-g) with the y-axis plotting the corrected $\delta^{13}\text{C}_{\text{VPDB}}$ (‰) value. For ease of comparison, all of the plots have been standardized with a 2 ‰ range on the y-axis.

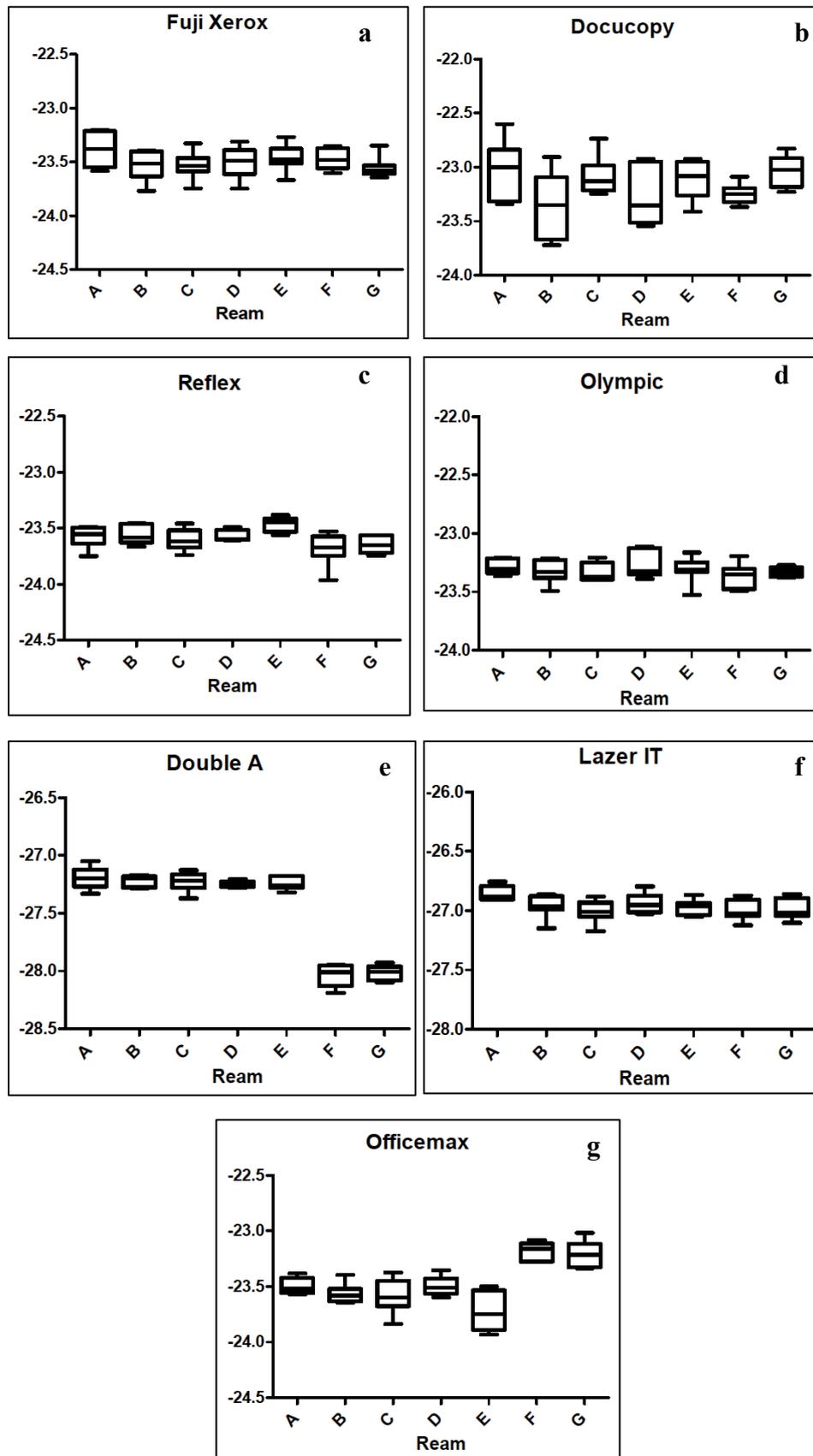


Figure 3.7 (a-g): Boxplots denoting the mean $\delta^{13}\text{C}_{\text{VPDB}}$ (‰) and standard deviation (‰) for seven reams from seven different brands

For the Fuji Xerox, Docucopy, Reflex, Olympic and Lazer IT brands, each of the seven reams measured produced results that were statistically consistent within their own brand group. The range of variation observed for each brand fluctuated however, from tight homogenous ranges of values (e.g. Lazer IT) to larger ranges (e.g. Docucopy) showing that there is inconsistency in terms of how much variance is observed between different brands of reams.

The other two brands, Double A and Officemax showed clear differences within the seven reams measured. For these samples, 5 reams could be discriminated from the two other reams. This is likely due to a difference during sample collection, with the 5 reams originating from one carton of paper (with 5 reams to a carton) and the other two collected as loose reams although at the same time.

Statistically, using the Kruskal Wallis test with Dunn's post-hoc test, significant differences were detected within the Reflex (p-value = 0.007), Double A (p-value = <0.0001) and Officemax (p-value = <0.0001) samples. Setting aside the Reflex ream result, Dunns post-hoc test indicated that the differences in the Double A and Officemax samples were due to the two reams that were different from the others, as observed in the boxplots. This result indicates that even though care was taken during the purchase of the reams, differences can occur due to packaging or longer term storage of the reams in-store. These results also show that differences in the age of the ream will produce variance in the measured values over time. This could be due to small changes in the source materials, or through the use of multiple manufacturing facilities for papers of the same brand.

For the Reflex reams, Dunn's post-hoc test showed significant differences between the E ream and F and G reams. This is due to a slight difference in the sample history of the reams, with the E ream (mean $\delta^{13}\text{C}_{\text{VPDB}} = -23.47 \text{ ‰}$) isotopically lighter than the F (mean $\delta^{13}\text{C}_{\text{VPDB}} = -23.69 \text{ ‰}$) and G (mean $\delta^{13}\text{C}_{\text{VPDB}} = -23.65 \text{ ‰}$) reams. The effect here is much more subtle than in the Double A and Officemax brands however and should be attributed to the natural variation of the Reflex branded samples, given that the difference between the means is only 0.2 ‰, and would not be discriminated given the 0.5 ‰ benchmark value being proposed here.

The grand mean and standard deviation for the 49 samples measured for each brand were also calculated. The Double A brand had the largest standard deviation of 0.38 ‰. Treating the G and H reams as having originated from another source (and removing them from the calculation), the adjusted standard deviation instead becomes 0.07 ‰. This leaves the Docucopy brand as having the largest overall standard deviation at 0.24 ‰. This standard deviation is consistent with the single ream homogeneity study results and provides further evidence that the use of a 1 ‰ total range as a confidence interval for comparison and discrimination of results is robust.

3.3.4 Discrimination of Background Population

Utilizing the 0.5 ‰ benchmark/discrimination value, the 125 background samples were compared to determine how well the carbon isotopic abundance values could discriminate the sample set. As such a large background sample set was being measured, each paper sample was measured in triplicate only. 7750 pairwise comparisons were made, from which 5254 pairs of papers could be discriminated. This equates to a discrimination of 68% of samples. The discrimination table is included electronically within Appendix 5. Although in isolation this value is not high, the discrimination power is expected to increase with the addition of oxygen isotopes and may increase through the use of other paper examination techniques. For a single isotope this test result has yielded a discrimination power that is higher than first expected, especially when considering the close range that Australian and Chinese papers in particular sit within.

3.4. Conclusions

125 paper samples commercially available in Australia and New Zealand were measured for their carbon isotope ratios. The range of values fell between -22 ‰ and -30 ‰. Some relationship could be inferred between the source of the ream and the $\delta^{13}\text{C}$ values measured. For example, the papers made in Australia and China consistently fell within the -23 ‰ to -24 ‰ range.

Seven recycled papers, produced in Australia, were measured and compared against the background paper samples to determine if their $\delta^{13}\text{C}$ abundance varied significantly from papers made with virgin pulps. No significant difference was observed.

Two experiments were undertaken to examine the homogeneity of document papers. The purpose of these experiments was to observe within ream and between ream variance of seven different brands of paper. These results were then utilised to inform the interpretation and comparison process. The maximum standard deviation for a single ream was 0.28 ‰. This translates to a 95% confidence interval of 0.55 ‰ either side the measured value or a total range of approximately 1 ‰. Using carbon isotope ratio values, this is the recommended minimum $\delta^{13}\text{C}$ difference for being able to conclusively differentiate two paper samples.

The between ream homogeneity experiments showed that differences within the same brand can be detected even when purchased from the same store at the same time. This means that variability within the same manufacturing facility or between different manufacturing facilities that produce the same brand of paper, can be detected using isotopic abundance values. Where one manufacturing facility has been used (e.g. in the case of Reflex branded reams) this means that isotopic abundance is not consistent across a production facility indefinitely and varies between production batches or runs.

Overall, the standard deviations within each single ream were low however, demonstrating that there is ream-to-ream variation (within a brand) but that this is generally larger than the single ream variation from which individual sheets are taken. The results also showed that paper produced from the same batch (within the same crate and therefore produced at roughly the same time) had comparable $\delta^{13}\text{C}$ values, within the 1 ‰ confidence interval range, which was deemed to be an appropriate threshold for discrimination after comparison with these results.

Using the 0.5 ‰ benchmark/discrimination value (1 ‰ discrimination range), 68% of the 125 background samples collected could be discriminated when a pairwise comparison was undertaken. This discrimination power is expected to increase with the addition of oxygen and hydrogen isotopes and may be further increased through the use of other paper examination techniques.

In terms of sampling, the recommended minimum number of samples to characterize a ream is 7 samples taken from 7 separate sheets. Measuring 25 samples from 25 sheets within a single a ream (5% of the ream total) did more effectively capture the variance within that ream however the difference between 7 samples and 25 samples was not significant enough to justify investment of the additional time and resources needed to prepare and run 25 replicate samples. In a casework context, 25 sheets may not be available for measurement and would likely hinder within sequence comparisons with the questioned samples, something that is preferential for robust comparison of samples.

Chapter 4

4. Carbon Isotopes – Source, Handling and Forensic Sampling

The main results of this chapter have previously been published in Jones et al (2013b). The published article is included electronically within Appendix 5.

4.1. Introduction

When undertaking any study of the isotopic abundance values of a bulk material, consideration should be given to the source materials and how they are combined to reach the final product being measured. While it is essential to measure and understand the natural variation of a material, in this case using clean papers, questioned paper samples in casework would have undergone some form of writing or printing process prior to being submitted for examination. Understanding the effects of these processes on the $\delta^{13}\text{C}$ abundance values of paper is essential for interpretation and comparison with clean samples, for example in cases where printed documents need to be compared to clean paper located in a printer feed tray.

The work in this chapter will present a range of experiments undertaken so that the source materials, the effects of the production process, the effects of printing and select forensic testing could be observed with respect to their effects on the carbon isotopic abundance values of 80gsm white office papers.

The key source question posed in this study seeks to define what is being measured in a bulk sample of paper and how much of that bulk value is related to the isotopic abundance values of the raw wood chips used in production. This has implications for future forensic intelligence and provenancing work, where models could be developed and used to eliminate regions of interest for suspect papers, if the measurement values of papers could be directly compared to known environmental values (e.g. isoscapes (Bowen, 2010, West et al., 2010)). To do this, samples were taken sequentially from the paper production facility at the Australian Paper Mill (Maryvale, VIC). These

samples ranged from raw wood chips through the pulping, whitening and refinement steps to the final formed and packed paper.

The secondary questions to be addressed in this study involve the alteration of the surface of paper, either by printing or through general use situations, or from the recovery of forensic material (including fingerprints and DNA) and their impact on $\delta^{13}\text{C}$ values measured.

For these types of papers, printing will be by a toner (laser) or inkjet printing process. These two processes vary in their ink type and delivery systems and hence may affect the $\delta^{13}\text{C}$ values in different ways. Toner printing processes utilise polymer particles that are heated to form a layer of material that is pressed and melted on to the surface of the page. Inkjet inks are water-based, fluid inks that are propelled/sprayed by a nozzle on to the surface of the page and subsequently absorbed by the fibres. Both printing methods leave scattered particles of toner or ink across the non-image areas of the page, creating potential contamination issues for IRMS sampling of the base substrate.

A range of handling and usage experiments were undertaken on papers to determine what effect, if any, these situations had on the $\delta^{13}\text{C}$ values of the paper. These experiments ranged from simple handling tests, to immersion in water and exposure to dirt/dust, heat, light charring and UV light.

Although document examination is typically performed prior to examination for fingerprints and collection of trace DNA, there are circumstances where questions regarding the origin or authenticity of documents are not asked prior to other forensic examinations taking place. 1,2-Indandione-Zn treatment for latent fingerprint enhancement is commonly employed in forensic laboratories, followed by tape-lifting to collect trace DNA. 1,2-Indandione-Zn was first investigated in 1997 as an amino acid attractive chemical treatment process that produces very strong luminescence in fingerprints developed on porous surfaces including paper (Wallace-Kunkel et al., 2007, Wallace - Kunkel, 2007, Spindler, 2011). For development of latent prints, after application of the 1,2-Indandione-Zn, heat is applied using a heat press around 150 degrees for up to 10 seconds. As the chemical soaks into the paper sheet, there is the

potential for interference from the chemical during subsequent IRMS measurements, or for fractionation to occur during heating.

Tape lifting is a simple technique, used to collect trace DNA by repeatedly placing and removing adhesive tape on the surface of paper. This tape removes the surface of the sheet, with any environmental debris including epithelial cells. Even though it is not a chemical process, there still exists the potential for alteration of the isotopic abundance due to either inclusion of adhesive residue or incomplete measurement of the paper due to the removal of the surface layer of the sheet onto the tape.

4.2. Materials and Methods

The instrumentation, equipment and method used to measure and correct the carbon isotopic abundance values in this chapter are outlined in Chapter 2.

Statistical tests were undertaken using Microsoft Excel® and GraphPad Prism version 5. When assessing whether differences were being observed, the Kruskal-Wallis Non-Parametric Analysis Of Variance (ANOVA) followed by a Dunn's Multiple Comparison Post-Hoc test was used, consistent with other statistical tests utilised for paper samples.

The fractionation factors associated with a production process were calculated using the following equations. The fractionation factor is an expression of the difference in reaction rates between the light and heavy isotopes through a chemical process or step. In this study, a simple fractionation factor was calculated using the equations in Fry (2006), where fractionation (Δ) is defined as:

$$\Delta = \delta_{\text{source}} - \delta_{\text{product}}$$

And the fractionation factor (α) of the chemical or physical fractionation process is:

$$\Delta = (\alpha - 1) * 1000$$

Mixing is when two or more source materials are combined to form a product with a new isotopic value. Also taken from Fry (2006), the equation used to establish the % contribution $\delta^{13}\text{C}$ of the source materials to the final $\delta^{13}\text{C}$ value of the product is:

$$\delta_{\text{sample}} = (\delta_{\text{source 1}}) * f_1 + (\delta_{\text{source 2}}) * f_2$$

Where f is the mole fraction of each component and is solved for f_1 as:

$$f_1 = \frac{(\delta_{\text{sample}} - \delta_{\text{source 2}})}{(\delta_{\text{source 1}} - \delta_{\text{source 2}})}$$

4.2.1 Source and Paper Production Samples

To examine the effect of the paper production process on the carbon isotopic composition of paper, a site visit at the Australian Paper Mill (Maryvale, Victoria) was conducted in February of 2011. Samples were collected at different stages of the production process, including:

- Whole eucalyptus wood chip (Wood Chip)
- Post-digester unbleached eucalyptus pulp (Unbleach Euc)
- Samples from within the bleaching/whitening process
 - Post oxygen wash (Ex O)
 - Post ozone and peroxide wash (Ex ZD)
 - Post caustic soda, oxygen and peroxide wash (Ex EOP)
 - Post chlorine dioxide wash (Ex D)
- Refined bleached eucalyptus (Refined Bleach Euc)
- Refined bleached pine
- Directly from the paper decal roll, post paper production (Paper Decal)
- Paper packed into reams for shipping (Paper Packed)
- Bulk measurements for packed paper and paper decal
- Paper filler material – Calcium Carbonate (CaCO_3)

One sample from each production stage was collected in sequence from throughout the process over a period of 4 hours. Due to the continuous process, it cannot be assumed

that the source will be identical in all of these samples. Taking the samples on the same day in sequence is as close as possible to following a batch of wood chips through the production process. A number of these samples were collected from wet processes and were dried as soon as practicable in a 60°C oven for 1-2 hours before being re-packaged into glass vials with Teflon lined screw top lids (Sigma Aldrich, Australia) and sealed with Parafilm (Sigma Aldrich, Australia).

As the removal of lignins is the primary function of the pulping and bleaching steps, it was necessary for a cellulose extraction to be performed on these samples so that only one target compound was being measured, rather than a mixture of changing proportions.

The equipment required to perform a cellulose extraction was not available in the AFP laboratory, hence samples from the paper production process were ground using a commercial coffee grinder, packaged into Parafilm sealed glass vials and sent to the University of California, Berkeley for extraction and measurement. The methods used to extract the cellulose are published (Tao et al., 2010, Gaudinski et al., 2005).

Where extracted sample sizes permitted, measurement of the carbon isotope abundance values were performed in duplicate. Results were reported directly to the author in the form of corrected $\delta^{13}\text{C}$ values. Information about the correction process used by this laboratory can be found online (Center for Stable Isotope Biogeochemistry - <http://dawsonlab.synthasite.com/stable-isotope-lab.php>).

In an attempt to replicate the cellulose extraction procedure in the AFP laboratory in a simple manner, 25 paper samples from the background population samples were treated with hydrochloric acid to remove the calcium carbonate filler material prior to measurement. To do this, 1cm samples of paper were cut into small squares and placed into 10mL glass beakers. Concentrated hydrochloric acid (HCl, Sigma Aldrich, Australia) was used to cover the surface of the paper and was left until no further bubbling from the surface of the paper was observed. Whilst being acidified, the beaker was agitated gently by swirling, to ensure that the acid covered both sides of the surface of the paper. After this, the paper was thoroughly washed using deionized

water and dried on Whatman filter paper (Sigma Aldrich, Australia) under vacuum. The papers tested are detailed with the results in Table 4.3.

4.2.2 Usage of Papers – Printing Processes

The second component of this study measured samples taken from office printing processes and forensically treated samples. The equipment and method utilised was the same as the one used to measure bulk paper production samples described in Chapter 3. To sample the papers, a 1.2mm micropunch (Harris Unicore, LabSciTech, Australia) was used to remove circular pieces of paper that weighed between 120 and 150µg of sample.

To assess the effect of printing on the $\delta^{13}\text{C}$ values of papers, a request was made through several networks in Australia and New Zealand for samples to be collected and sent for measurement. The instructions provided asked for a 3-page preformatted document to be printed and sent with a sample of the same un-printed paper so that before and after printing measurements could be made. Information on the make, model and type of printer was requested in addition to information including paper brand, packing date (if available) and purchase/collection location of the paper ream.

46 toner and 13 inkjet test prints were received for measurement. The number of inkjet samples was lower than hoped but this was likely due to business and government departments using toner printers much more frequently for economic reasons. A good range of brands was received and a number of repeats of the same manufacturer and model were also received, increasing the comparability of the test sample base. The papers and printers tested are detailed in Table 14.4 (toner printers) and Table 14.6 (inkjet printers).

Triplicate measurements were taken from each of the unprinted sheets, the printed sheets in the non-image area and as a worst-case scenario, from directly on top of the printed text. The printed text measurement was also made to give a preliminary indicator for the potential for discrimination of the ink/toner using isotopic abundance values using bulk measurements taken directly from a document.

4.2.3 Usage of Papers - Environmental Effects and Forensic Testing

The following testing was undertaken on five different papers, as detailed in Table 4.4. The five reams were labelled A-E and three replicate sheets of paper from the five reams subsequently labelled as pages A, B and C. A 4 cm strip of paper was cut from the bottom of each page as the paper blank. The remaining sheet was sectioned into quarters, for use in the environmental and forensic testing experiments described below. The environmental tests were constructed to span a wide range of usage or damage scenarios. None of the conditions of the tests were conducted quantitatively.

- **Dirt/Dust**

The sheets were soiled by placing them on an outside concrete surface and applying pressure to maximise contact and adhesion of contaminants. Loose particles were dusted from the page prior to sampling.

- **Heat**

The sheets of paper were placed in a 60-degree oven for 4 hours, taken out and allowed to cool prior to sampling.

- **Fire**

The sheets of paper were charred by placing them above an open flame until mild discolouration occurred. Samples were taken from the discoloured sections of the sheets.

- **UV Exposure**

Sheets were placed in a closed container/box with a 254nm UV light source for approximately 2 hours.

- **Handling/Body Contact – Low**

Using a thumb, the sheets of paper were repeatedly touched in the same location by flicking/ turning sheets in the same place 10 times on each sheet. Samples were taken directly from the area of contact.

- **Handling/Body Contact – Medium**

Using multiple fingers, the sheets were touched by creasing/ folding the page by pressing and smoothing with the fingers. Samples were taken directly from the crease/ fold area.

- **Handling/Body Contact – Heavy**

Multiple fingers were held and rubbed on the sheets of paper for 5 minutes. The contact was very visible on the pages as body oils accumulated. Samples were taken directly from the contact area.

For the forensic testing - fingerprint and DNA collection was simulated using one sheet from each brand that was subjected to treatment either with 1,2-Indandione-Zn or DNA tape lifting. Note, no fingerprints or DNA were deposited on the surface of the pages prior to testing in these evaluations. For the tape lifting, a piece of adhesive tape was placed lightly on to the surface of the sheet and removed. This was repeated multiple times.

4.3 Results and Discussion

4.3.1 Source and Paper Production Samples

The results from the cellulose extracted paper process samples are detailed in Table 4.1. This table also includes the bulk measurement values obtained for each of the processing stages. The $\delta^{13}\text{C}$ values for the main additives used in the paper production process are presented in Table 4.2.

Production Stage	Description	Cellulose Extracted Mean $\delta^{13}\text{C}_{\text{VPDB}}$ (‰)	Bulk Mean $\delta^{13}\text{C}_{\text{VPDB}}$ (‰)	Difference Cellulose Extraction vs. Bulk (‰)
Raw Wood Chip		-26.11 ± 0.3	-26.34 ± 0.00	0.23
Unbleached eucalyptus	After pulping	-25.83 ± 0.05	-25.60 ± 0.05	0.2
Ex O	First whitening step	-25.51 ± 0.06	-25.69 ± 0.00	0.18
Ex ZD	Second whitening step	-25.77 ± 0.02	-25.77 ± 0.06	Nil
Ex EOP	Third whitening step	-25.77 ± 0.05	-25.74 ± 0.03	0.03
Ex D	Last whitening step	-25.73 ± 0.13	N/A	
Refined Bleached Euc	Last step before mixing with other pulp and additives	-25.74 ± 0.03	-26.02 ± 0.05	0.28
Formed Paper	After paper production	-25.92	-23.91 ± 0.15	2.01
Packed Paper	After cutting and packing	-25.86 ± 0.05	-23.50 ± 0.13	2.36

Table 4.1: Cellulose extracted and bulk measured $\delta^{13}\text{C}_{\text{VPDB}}$ (‰) ± 1 standard deviation for paper source and production samples

Additive Name	Mean $\delta^{13}\text{C}_{\text{VPDB}}$ (‰)
Dry Bleached Pine (cellulose extracted)	-26.05 ± 0.04
Refined Bleached Pine (cellulose extracted)	-25.82 ± 0.12
Calcium Carbonate Filler (bulk measured)	2.14 ± 0.12

Table 4.2: Cellulose extracted and bulk measured $\delta^{13}\text{C}_{\text{VPDB}}$ (‰) ± 1 standard deviation, for additives used in paper production

Comparing the cellulose extracted values with the bulk measurements in Table 4.1, a decrease in isotopically heavier extraneous material (reflected by an increase in the $\delta^{13}\text{C}$ value) is evident as sample production proceeds. By the second whitening step (Ex ZD) the cellulose extracted and bulk values are comparable (within 2 standard deviations). This means that the lignins and other contaminant materials have been removed or are at trace levels. This also suggests that no fractionation is occurring during these whitening stages.

Additionally, there is a difference between the extracted and bulk samples for refined eucalyptus. This was unexpected as the advice received from the paper mill was that at this stage in production the mixture should only contain pure, whitened cellulose fibres. The difference between the two measurements contradicts this information and suggests that an additive has been placed into the mixture. Alternatively, the refinement process could also have caused a fractionation event. To confirm this, a bulk IRMS measurement was undertaken of the refined bleached pine sample, which is expected to be in the same form as the refined eucalyptus fibres. The bulk refined pine sample was measured at $\delta^{13}\text{C}_{\text{VPDB}} -26.01 \text{ ‰} \pm 0.05$ and when compared to the cellulose extracted refined pine sample in Table 4.1 ($\delta^{13}\text{C}_{\text{VPDB}} -25.82 \text{ ‰} \pm 0.12$), a similar difference of approximately 0.2 ‰ is observed, confirming that there is either an additive present at this stage of the process or that fractionation of the cellulose has occurred during refinement.

The difference between the dry and refined pine samples is due to either a batch difference or a production processing difference. The raw pine fibres were sampled

dry from a storage pallet whereas the refined samples were sampled directly from the wet manufacturing section, ready for input into the paper furnish.

Figure 4.1 is a plot of the $\delta^{13}\text{C}$ values of the extracted cellulose from each step in the production process. Figure 4.2 contains the same data, plotted against bulk measurements for the same samples. This second plot highlights the stages in the process where more than just the cellulose content of the paper is contributing to the $\delta^{13}\text{C}$ values measured. For both plots, error bars were omitted, as they were negligible at this scale.

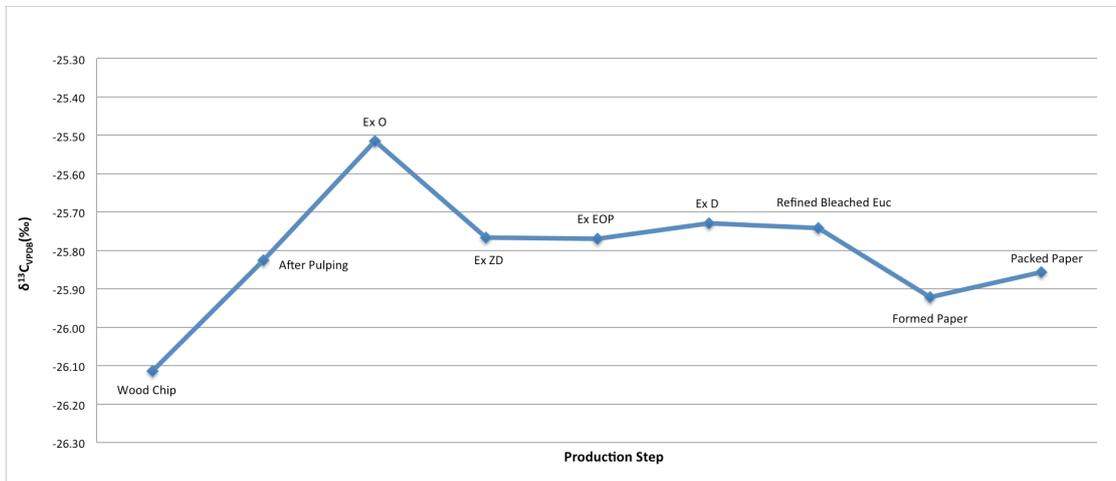


Figure 4.1: Mean $\delta^{13}\text{C}$ values of cellulose extracted from samples collected throughout the paper production process

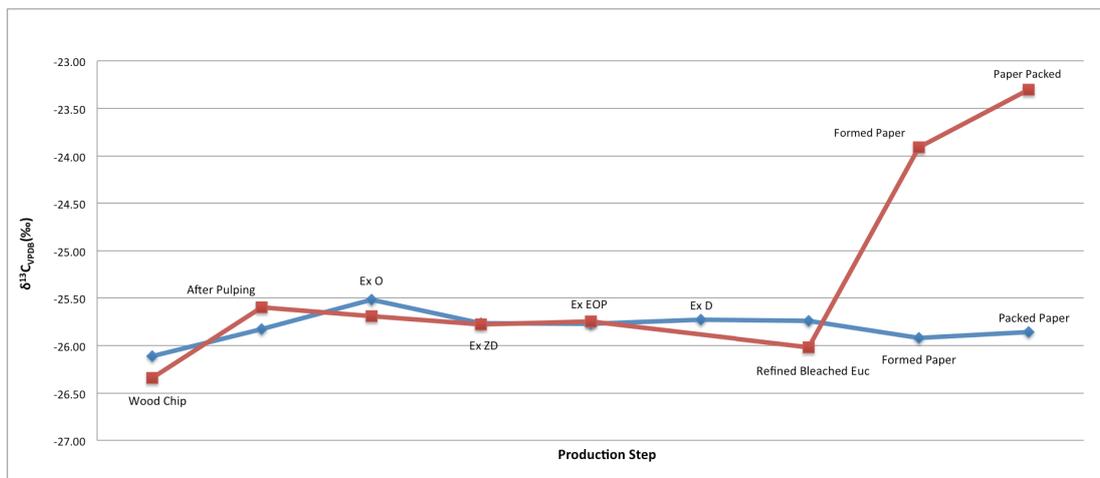


Figure 4.2: Comparison of mean $\delta^{13}\text{C}$ values of extracted cellulose (diamond markers) with bulk values (square markers) collected from throughout the paper production process

The cellulose extracted samples in Figure 4.1 show fractionation of the cellulose occurring as it is processed from raw wood chip through pulping and the first stage of bleaching. The fractionation is small however, at 0.6 ‰ between the two stages. Utilising the equations for calculating the fractionation factor (α) (Fry, 2006), the fractionation factor for this reaction is calculated at 1.006, which reflects a modest fractionation and illustrates that the process was favouring the lighter isotope.

A small difference in the $\delta^{13}\text{C}$ values measured was observed between the raw cellulose (-26.11 ‰) extracted from the wood chip and the formed paper (-25.92 ‰) in Table 4.1. This was a difference of 0.2 ‰ between the whitening and refinement stages of the process and the formed and packed paper samples i.e. between the start and the end of the actual paper production stages. While this may be an additional fractionation process, the magnitude is very small and could equally be explained by a natural variation of the source material or due to addition of the slightly more depleted pine fibres.

A more significant difference can be observed between the bulk and cellulose extracted sample measurements. A difference of over 2 ‰ is observed between both the formed and packed paper samples when the cellulose is extracted and run in isolation. From information obtained about the constitution of the paper, this is due to the inclusion of a significant proportion of calcium carbonate (CaCO_3) filler, added into the paper mixture to smooth, bulk and stabilise the formed sheet.

The CaCO_3 was measured to have a $\delta^{13}\text{C}_{\text{VPDB}}$ $2.14 \text{ ‰} \pm 0.12$ (Table 4.2) - significantly different to the $\delta^{13}\text{C}_{\text{VPDB}}$ -25.7 ‰ value of the eucalyptus fibres. Putting aside the contribution of the addition of the Pine cellulose (which would be minimal due to its similar $\delta^{13}\text{C}$ value, measured at $\delta^{13}\text{C}_{\text{VPDB}}$ -25.82 ‰), and the contribution of other additives (which would be small due to their low percentage in the mixture), using the mixing equations presented in Fry (2006) indicated that the formed paper contains approximately 8% CaCO_3 filler. This is below the expected range of filler composition (10-25%) quoted verbally by the paper mill and published within the literature (Biermann, 1996, Wolfgang et al., 2001, Murphy, 2009).

The measurement results of the bulk versus acidified samples are shown in Table 4.3, with sample numbers retained from the background samples in Chapter 3. Figure 4.3 shows the difference between the two measurements. For Australian papers, the difference in $\delta^{13}\text{C}$ value is close to the difference observed between the cellulose extracted and bulk samples in Table 4.1.

Sample Number	Paper Country of Origin	Bulk Paper Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ St Dev} (\text{‰})$	After Acidification Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ St Dev} (\text{‰})$	Difference (‰)
43	Australia	-23.10 \pm 0.06	-25.89 \pm 0.14	2.79
45	Australia	-23.19 \pm 0.04	-26.67 \pm 0.01	2.08
66	Australia	-24.59 \pm 0.14	-25.92 \pm 0.10	2.87
67	Australia	-23.05 \pm 0.08	-25.87 \pm 0.12	2.57
76	Australia	-23.30 \pm 0.07	-25.62 \pm 0.02	2.21
75	Australia	-23.16 \pm 0.08	-26.01 \pm 0.05	2.84
77	Australia	-23.85 \pm 0.06	-25.90 \pm 0.05	2.04
78	Australia	-23.50 \pm 0.15	-26.39 \pm 0.09	2.88
79	Australia	-23.50 \pm 0.08	-26.18 \pm 0.06	2.67
82	Indonesia	-29.58 \pm 0.04	-29.73 \pm 0.06	0.15
83	Indonesia	-28.37 \pm 0.09	-29.55 \pm 0.09	1.18
88	Indonesia	-27.57 \pm 0.07	-27.90 \pm 0.05	0.33
89	Indonesia	-26.56 \pm 0.04	-26.93 \pm 0.01	0.38
92	Thailand	-27.01 \pm 0.07	-27.19 \pm 0.01	0.18
93	Thailand	-27.88 \pm 0.02	-27.64 \pm 0.08	0.23
94	Thailand	-28.21 \pm 0.02	-27.74 \pm 0.07	0.47
95	Thailand	-28.54 \pm 0.05	-28.34 \pm 0.13	0.20
96	Thailand	-27.55 \pm 0.14	-27.44 \pm 0.03	0.11
102	China	-22.82 \pm 0.17	-25.72 \pm 0.10	2.90
103	China	-24.31 \pm 0.12	-27.45 \pm 0.16	3.14
105	China	-23.78 \pm 0.16	-26.43 \pm 0.07	2.65
109	South Africa	-25.00 \pm 0.09	-25.75 \pm 0.05	0.75
111	India	-25.61 \pm 0.06	-27.08 \pm 0.08	1.48
120	China	-23.40 \pm 0.10	-25.93 \pm 0.06	2.53
121	Brazil	-25.79 \pm 0.01	-25.64 \pm 0.09	0.15

Table 4.3: Papers measured for $\delta^{13}\text{C}_{\text{VPDB}}$ before and after acidification using hydrochloric acid

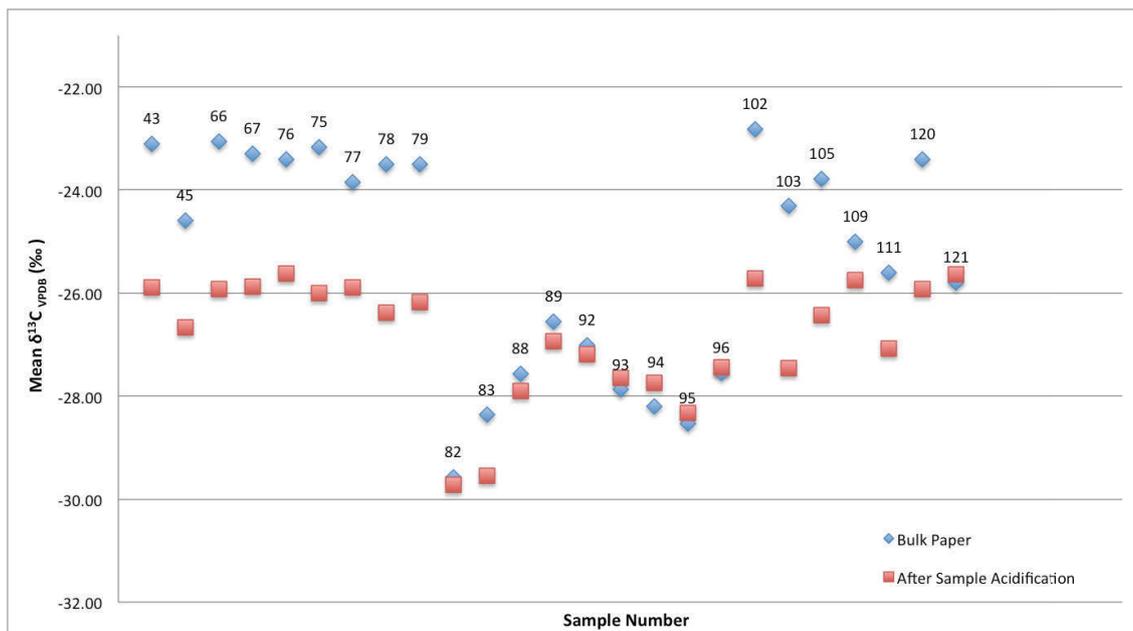


Figure 4.3: Mean $\delta^{13}\text{C}$ values of bulk paper (diamond markers) compared with the same paper samples after undergoing acidification (square markers). Sample number is taken from Table 4.3.

Figure 4.3 shows that the difference between bulk and acidified papers varies depending on the paper sample being measured. From this we can draw two conclusions – that CaCO_3 content varies depending on manufacturing choices made during production (i.e. that the ‘recipe’ for paper is not consistent) and that other filler materials may be being used (e.g. kaolin clay, talc or titanium dioxide) that are either not removed using hydrochloric acid or have little to no effect on the carbon isotopic abundance values of paper.

For comparison of papers this result has wider implications. When two bulk values are measured and are found to not be able to be discriminated (therefore indicating that they may have shared a common source), it may be possible that while one sample has a CaCO_3 filler the other does not but lies in the same $\delta^{13}\text{C}$ range. While these samples may be discriminated using a combination of isotopes, a risk of false association exists when utilising a single carbon isotopic measurement. Therefore, it is recommended that further testing be undertaken via acidification of the samples when the measurement values are found to be within the same range. This will elucidate further information about the sample and may potentially discriminate the samples. Alternatively, chemical identification of the filler materials could be undertaken prior

to measurement with IRMS. A pilot study using X-Ray Diffraction is included for this purpose in Chapter 9.

Overall these results show that paper production contains both fractionation and mixing stages that impact the ability to directly compare (or associate) the bulk carbon measurements to the wood source via published literature (e.g. cellulose isoscapes) without knowing the $\delta^{13}\text{C}$ value and percent composition of the bulk filler material. For some papers, it may be possible to measure the true cellulose value by using hydrochloric acid treatment to remove CaCO_3 prior to measurement and this should be considered as an additional comparison tool when the $\delta^{13}\text{C}$ values of two samples cannot be readily discriminated.

4.3.2 Usage of Papers – Printing Processes

Table 14.5 and Table 14.7 detail the blank paper, printed paper (non-image area) and printed paper (image/text area) $\delta^{13}\text{C}_{\text{VPDB}}$ for the 46 toner and 14 inkjet samples measured. The mean values in these tables are taken from triplicate measurements for each condition.

Figure 4.4 and Figure 4.5 show plots for the data in each table, with the error bars representing the 0.5 ‰ benchmark/discrimination value defined in Chapter 3.

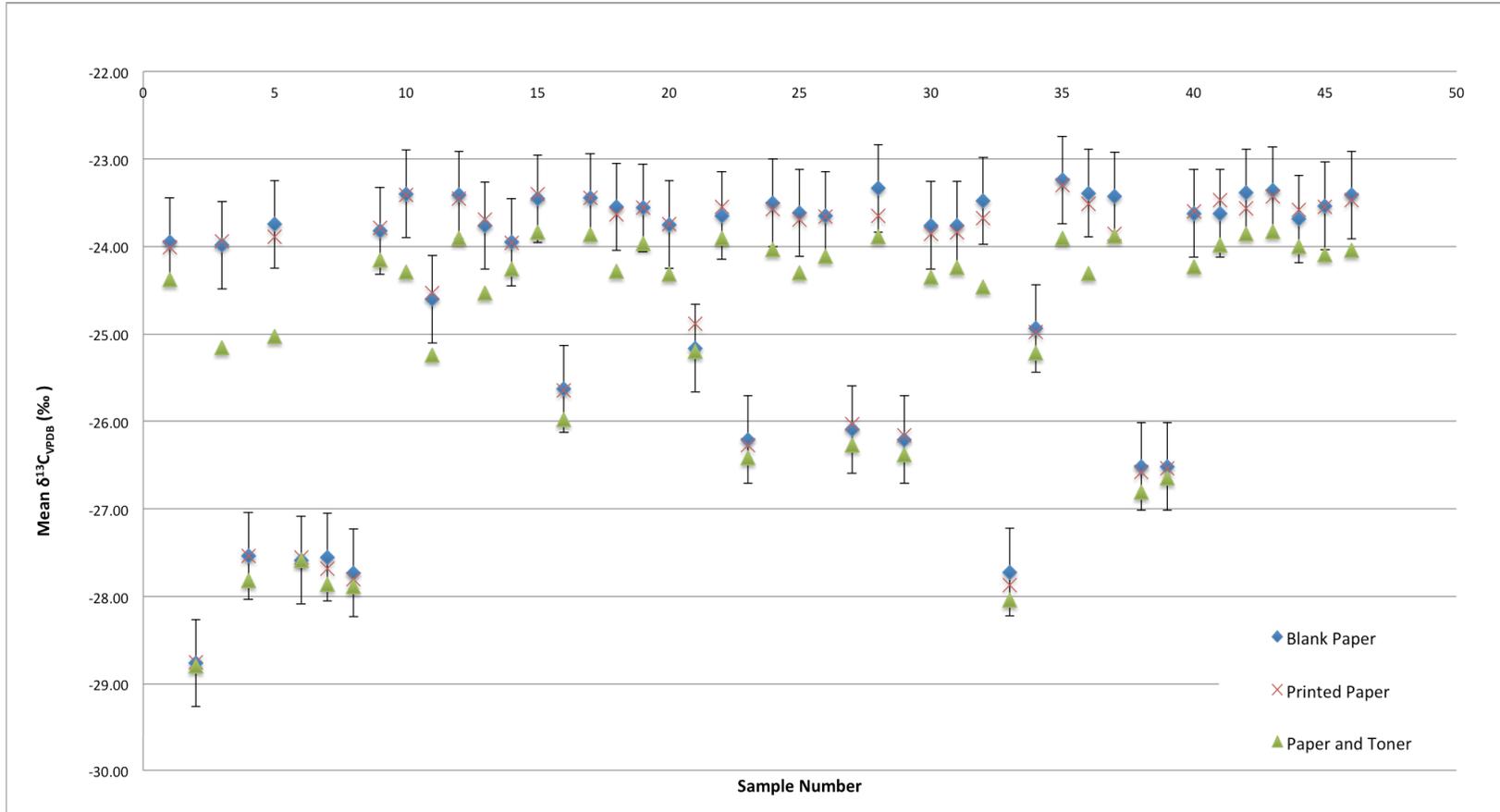


Figure 4.4: Comparison of the mean $\delta^{13}\text{C}_{\text{VPDB}}$ values of papers before and after printing using toner

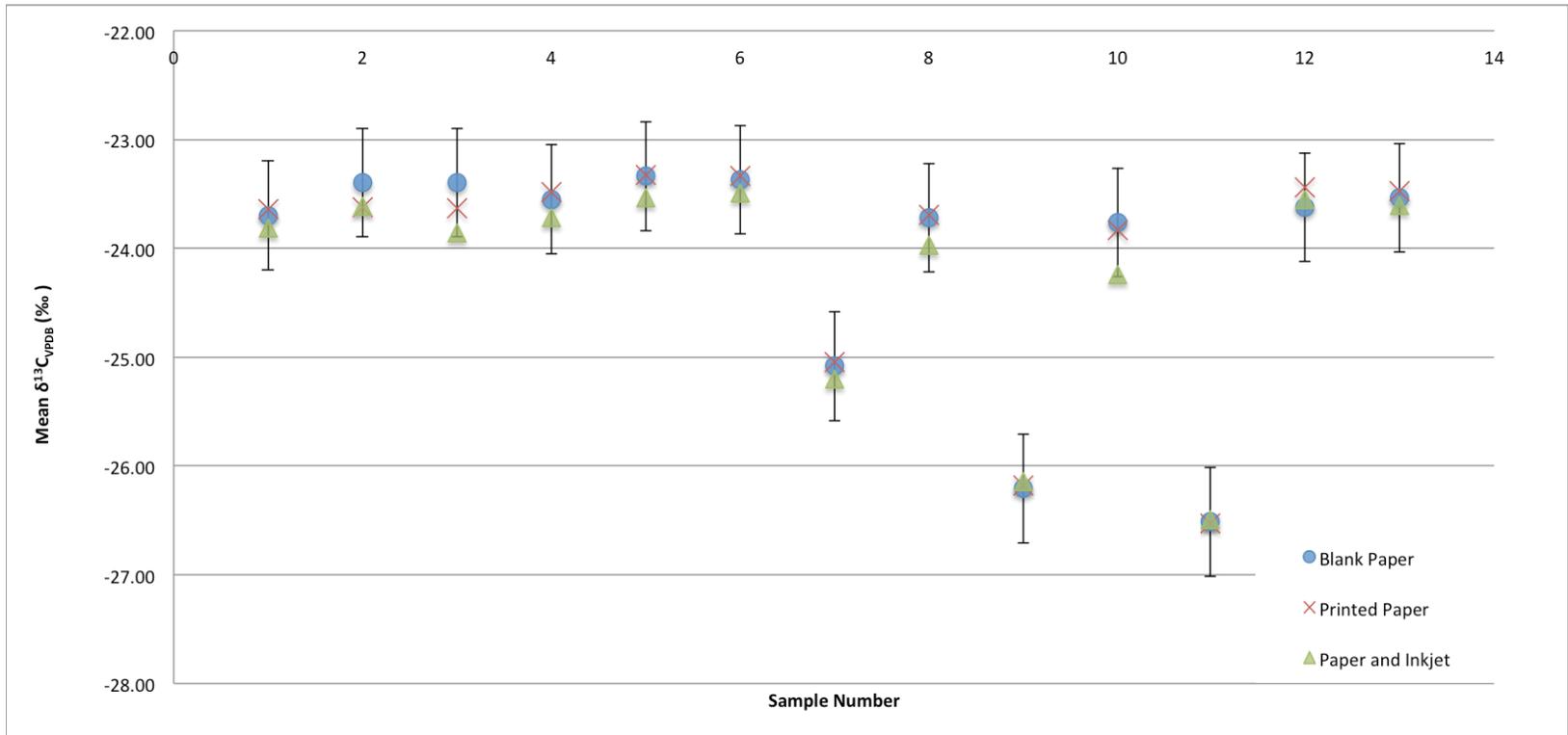


Figure 4.5: Comparison of the $\delta^{13}\text{C}_{\text{VPDB}}$ values of papers before and after printing using inkjet ink

The toner printed samples plotted in Figure 4.4 indicates that the inclusion of toner (either particles or sampling directly from the printing itself) contributes to the $\delta^{13}\text{C}_{\text{VPDB}}$ value measured. In all samples, this measured value was more depleted than the un-printed and printed non-image area samples and in some cases the difference is large enough that the samples would be classified as different during interpretation. The magnitude of the difference was unique to the sample however, showing that even if the principle of identical treatment was used (and both samples were taken from regions that appeared to contain a similar ink density), results from the printed area samples could not be used for comparison of the papers underneath.

Although no consistent difference was observed between the unprinted papers and the samples taken from the non-image areas of the printed pages, for toners that are significantly more depleted than the paper substrates, the presence of extraneous particles from the printing process may be significant. To prevent this, care should be taken during sampling to identify an area with low/no particles using magnification or microscopy, prior to punching the samples. From a practical perspective this would also be essential as different makes, models, age, defects and even the settings used when producing prints from toner machines will impact the amount and location of these particles, which may not be evident without using some form of magnification.

The samples provided from the same manufacturer and model of machine were compared to observe whether the changes to the printed-paper samples were uniform. Some of the repetition of manufacturer and model observed in the sample set was due to one sample collection provider changing the paper utilised and printing multiple test sheets from the same machine. Overall, there was no consistency in the effect that a single machine printing with different papers, or the same make and model of machine, had on the values of the papers or on the toner + paper samples measured. It is hypothesized that this is likely due to the non-uniform nature of sampling for the toner, which would impact on the amount of toner present and therefore the measurement value in relation to the paper that is of a fairly constant mass.

The difference observed between the blank paper samples and the samples that contain toner indicate that if toner can be effectively removed from the paper sheet and

weighed to a controlled sample size that toner as a material has strong potential for comparison and discrimination using IRMS.

The inkjet samples plotted in Figure 4.5 showed depletion of the $\delta^{13}\text{C}$ values when directly sampled from the image areas of the printed pages. This difference however was not as significant as the toner samples and did not result in the sample values moving outside of the 0.5 ‰ benchmark range for discrimination. The non-image area samples were comparable to the unprinted paper samples suggesting that extraneous ink had less impact on the measured values for inkjet inks than toners. With respect to inkjet inks, microscopically this result makes sense as the ink spatter observed with inkjet printing is generally close to the printed characters and only observed outside of that when printing images or scanned documents.

In general, for both printing sample types, the precision of the measurement (standard deviation values) was not affected by the presence of the ink or toner.

4.3.3 Usage of Paper – Environmental Effects and Forensic Testing

The papers used and their untreated $\delta^{13}\text{C}$ values are shown in Table 4.4. The measured values for the environmental samples are shown in Table 4.5.

Paper Brand	Paper Designator	Country of Origin	Untreated Mean $\delta^{13}\text{C}_{\text{VPDB}}$ (‰) for 3 sheets sampled	Untreated Std Dev (‰) for 3 sheets sampled
Double A	A	Thailand	-27.50	0.09
Reflex	B	Australia	-23.62	0.07
Olympic	C	Australia	-23.37	0.12
Lazer IT	D	Indonesia	-26.88	0.12
Fuji Xerox	E	China	-23.27	0.16

Table 4.4: Brands of ream used and pooled mean values for environmental and forensic test evaluation

Treatment	Sample Name	Untreated Paper $\delta^{13}\text{C}_{\text{VPDB}}$ Mean \pm 1 Std Dev (‰)	Treated Paper $\delta^{13}\text{C}_{\text{VPDB}}$ Mean \pm 1 Std Dev (‰)	Difference of Means (‰)
Dirt/Dust	AC	-27.4 \pm 0.11	-25.49 \pm 0.10	1.96
	BA	-23.73 \pm 0.00	-21.9 \pm 0.00	1.82
	CA	-23.49 \pm 0.05	-21.88 \pm 0.14	1.61
	DA	-27.00 \pm 0.11	-25.69 \pm 0.01	1.31
	EA	-23.21 \pm 0.23	-22.25 \pm 0.03	0.96
Heat	AC	-27.4 \pm 0.00	-26.26 \pm 0.06	1.29
	BB	-23.59 \pm 0.05	-26.63 \pm 0.12	0.96
	CA	-23.49 \pm 0.05	-22.51 \pm 0.09	0.98
	DB	-26.87 \pm 0.10	-26.36 \pm 0.01	0.51
	EA	-23.21 \pm 0.23	-22.89 \pm 0.09	0.31
Fire	AB	-27.57 \pm 0.05	-27.12 \pm 0.09	0.45
	BC	-23.58 \pm 0.02	-23.48 \pm 0.03	0.10
	CB	-23.27 \pm 0.06	-23.18 \pm 0.20	0.09
	DB	-26.87 \pm 0.01	-26.94 \pm 0.02	0.07
	EB	-23.30 \pm 0.04	-23.48 \pm 0.05	0.18
UV Exposure	AC	-27.40 \pm 0.00	-26.91 \pm 0.11	0.54
	BB	-23.73 \pm 0.00	-23.29 \pm 0.02	0.30
	CA	-23.49 \pm 0.05	-23.38 \pm 0.10	0.11
	DB	-26.87 \pm 0.01	-27.23 \pm 0.05	0.35
	EA	-23.21 \pm 0.23	-23.84 \pm 0.07	0.63
Handling Low	AB	-25.57 \pm 0.05	-27.16 \pm 0.12	0.41
	BC	-23.58 \pm 0.02	-23.45 \pm 0.12	0.13
	CB	-23.27 \pm 0.06	-23.19 \pm 0.03	0.08
	DC	-26.78 \pm 0.01	-27.00 \pm 0.05	0.22
	EB	-23.30 \pm 0.04	-23.49 \pm 0.07	0.2
Handling Medium	AA	-27.50 \pm 0.11	-27.18 \pm 0.08	0.32
	BB	-23.59 \pm 0.05	-23.43 \pm 0.05	0.17
	CC	-23.36 \pm 0.13	-23.26 \pm 0.16	0.10
	DC	-26.78 \pm 0.01	-26.96 \pm 0.00	0.18
	EC	-23.33 \pm 0.14	-23.48 \pm 0.20	0.16
Handling High	AA	-27.50 \pm 0.11	-27.05 \pm 0.06	0.45
	BB	-23.59 \pm 0.05	-23.54 \pm 0.05	0.05
	CC	-23.36 \pm 0.13	-23.24 \pm 0.08	0.12
	DC	-26.78 \pm 0.01	-26.92 \pm 0.12	0.14
	EC	-23.33 \pm 0.14	-23.58 \pm 0.15	0.29

Table 4.5: $\delta^{13}\text{C}_{\text{VPDB}}$ measurement values for environmental testing experiment

The environmental test that had the largest effect on $\delta^{13}\text{C}$ values was the accumulation of dirt/dust on the surface of the sheets. This test, while not conducted quantitatively in terms of the amount of dirt placed on to the surface of the sheets, produced a universal change in the $\delta^{13}\text{C}$ values of the pages. The heat and UV exposure tests also changed the $\delta^{13}\text{C}$ on each of the sheets, although the change was not uniform. The uneven effect of the heat and UV stimuli may be due to the orientation of the sheets in the oven or the proximity to the UV lamp. The other environmental tests did produce some effect on the papers however it was not consistent and in general not higher than the measurement uncertainty for the method (0.2 ‰).

To determine whether the differences observed were statistically significant, an unpaired t-test was performed on the dirt/dust, heavy handling and heat exposure tests. The statistical results of the five different papers for each of the 3 conditions are shown in Table 4.6. Box plot comparisons are shown in Figure 4.6 (a-c). The t-tests show significant differences for the papers measured for dirt/dust and heat treatments, with a net enrichment of the $\delta^{13}\text{C}$ values being observed. While there is some difference observed for the handling-high test in Table 4.5, the results were not shown to be statistically significant.

	T-ratio	Degrees of freedom	P-value
Dirt/Dust			
AC	25.93	3	0.0001
BA	424.47	3	< 0.0001
CA	18.49	4	< 0.0001
DA	16.60	3	0.001
EA	7.27	4	0.002
Handling - High			
AA	4.96	2	0.038
BB	1.22	4	0.289
CC	1.27	3	0.293
DC	2.90	4	0.044
EC	2.14	4	0.099
Heat			
AC	27.67	3	0.000
BB	12.11	4	0.0003
CA	16.64	4	< 0.0001
DB	7.97	4	0.0013
EA	1.77	3	0.175

Table 4.6: Results of statistical t-tests for three conditions, on five papers.

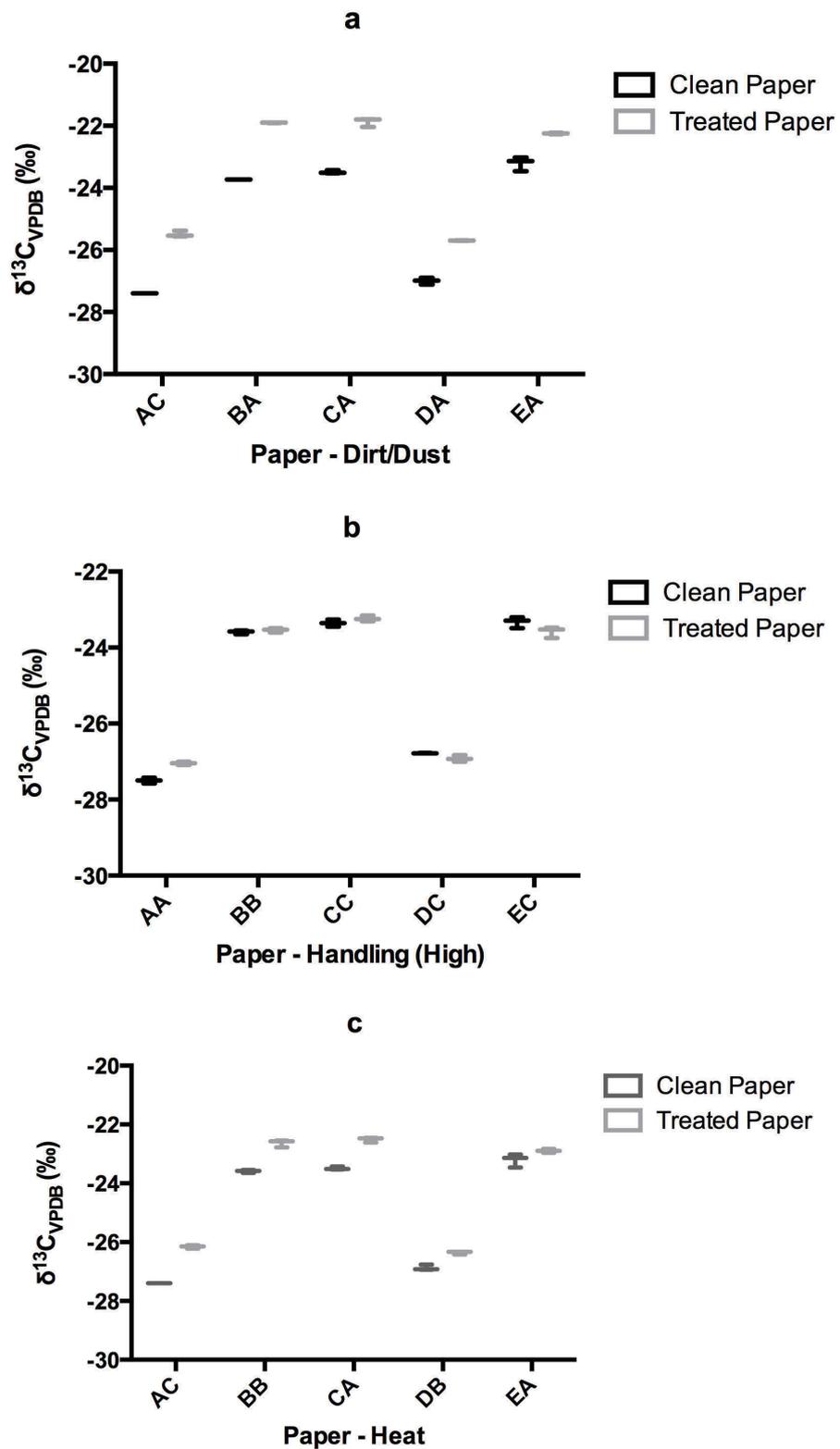


Figure 4.6 a-c: Box plots of five papers measured before and after environmental testing

Table 4.7 details the before and after values for the samples treated for fingerprint detection using 1,2-Indandione-Zn and collection of DNA using a tape-lifting procedure. No significant differences (based on the 0.5 ‰ benchmark differentiation level) were observed between the before and after results for these tests. These results indicate that IRMS of papers is a feasible analytical option that can be conducted following two commonly applied techniques for fingerprint development and collection of DNA.

Treatment	Sample Designator	Untreated Paper Mean $\delta^{13}\text{C}_{\text{VPDB}}$ (‰)	Treated Paper Mean $\delta^{13}\text{C}_{\text{VPDB}}$ (‰)	Treated Paper Std Dev (‰)	Difference of Means (‰)
Indandione	AB	-27.40	-27.24	0.01	0.16
	BC	-23.58	-23.49	0.04	0.09
	CB	-23.27	-23.33	0.17	0.06
	DB	-26.87	-27.10	0.13	0.23
	EB	-23.30	-23.48	0.03	0.18
Tape lifting	AB	-27.40	-27.22	0.07	0.18
	BC	-23.58	-23.55	0.07	0.03
	CB	-23.27	-23.18	0.16	0.09
	DC	-26.78	-27.01	0.04	0.23
	EB	-23.30	-23.47	0.02	0.17

Table 4.7: $\delta^{13}\text{C}_{\text{VPDB}}$ measurement values for forensic testing experiment

4.4. Conclusions

The paper pulping and production process was examined in detail to determine if the chemical and physical manufacturing processing steps had any effect on the $\delta^{13}\text{C}$ values of cellulose. A moderate fractionation was observed within the pulping and first stages of whitening due to the removal of lignin from the raw wood chip. The largest effect however was observed to be a mixing step prior to the paper formation process, with the addition of CaCO_3 filler. This result means that the bulk carbon measurements cannot be compared to cellulose or ecological literature (i.e. to raw

cellulose values) without knowing the $\delta^{13}\text{C}$ value and per cent composition of the bulk filler material. Acidification with hydrochloric acid was shown to remove what is assumed to be the CaCO_3 content, which was subsequently shown to vary considerably between different production locations and batches. As an alternative to the use of isoscapes, a large-scale background study of the area of interest may elucidate trends with respect to the expected values of papers originating from specified regions.

The effect of printing on paper samples was examined and toner printing was found to have the greatest effect on changing the $\delta^{13}\text{C}$ values of the paper. From this study, the following recommendations are made for consideration during the forensic examination of papers:

1. Samples punched from printed documents for paper comparison, should ideally not contain any ink;
2. In particular with respect to toner printed documents, microscopy or magnification should be used to ensure that no extraneous toner particles are included with the paper samples;
3. Based on the paper samples measured that contained toner, if toner can effectively be removed from the paper sheet (using a method that will require further development) and hence the sample size controlled, toner as a material in its own right should be investigated for comparison using IRMS;
4. For inkjet printed samples, the unprinted paper samples are comparable to the printed non-image area samples and as such, microscopy prior to sampling is less important for this type of printing;
5. IRMS of papers is still feasible after 1,2-Indandione-Zn treatment for the visualization of fingerprints and tape-lifting for DNA collection; and
6. Some environmental conditions were observed to change the $\delta^{13}\text{C}$ values of papers. This should be considered in the context provided with casework to avoid misinterpretation of results, particularly when comparing papers with different histories.

Further research is recommended to investigate the types and amount of fillers present in papers to better understand the differences observed in the acidification experiment. A pilot experiment utilising X-Ray Diffraction is included in Chapter 7 to identify the type of filler present in document papers. It would also be advantageous for the paper production study to be repeated, both at the same and different mills to confirm the observations here.

Chapter 5

5. Oxygen Isotopes – Method Development and Validation

5.1. Introduction

To complement the work undertaken for carbon isotopes and to ensure that the methods developed are operationally deployable, method development and validation was undertaken for the measurement of oxygen isotopes. The experiments conducted are similar to those for carbon detailed in Chapter 2.

The same principles, based around international accreditation to ISO17025 and NATA standards (National Association of Testing Authorities, 2015, ISO, 2005) were applied to perform this method development and validation study.

5.2. Materials and Methods

5.2.1 Standards and Samples

To correct all paper samples, international benzoic acid standards IAEA-601 and IAEA-602 (purchased from the International Atomic Energy Agency, Vienna) were run in replicates of 5 at the start and the end of all analytical sequences. The published $\delta^{18}\text{O}_{\text{VSMOW}}$ values of these standards are 23.3 and 71.4 ‰ respectively (Coplen et al., 2006). International standard cellulose IAEA-CH-3 was used as a quality assurance material. While no published/agreed value exists for the oxygen isotopic abundance of this material, inclusion of it in each run will allow a laboratory value to be determined over time.

A Genius ME5 (Sartorius, Goettingen, Germany) analytical balance was used to weigh the international benzoic acid standards IAEA-601, IAEA-602 and IAEA-CH-3. After sample size experimentation, the weight for these standards was set at a sample size of $250\mu\text{g} \pm 20\mu\text{g}$. Paper samples were prepared using a combination of a 1.2mm and

2mm Harris Uni-core micro-punches (Proscitech, Queensland, Australia) to a sample size of 250µg.

For the reproducibility experiment, Polyethylene Glycol (PEG) and coumarin working materials (obtained from other forensic isotopic laboratories) and international standard cellulose samples were weighed to 150, 400 and 200µg respectively. Samples were placed into 3.3 x 4 mm silver capsules for solids (Thermo Fisher Scientific, Sydney, Australia) and stored in a Perspex desiccator with self-indicating silica gel until measurement. Where not specified, such as in the method development experiments, paper samples were selected at random from the background paper population.

VSMOW2 and SLAP2 water samples were obtained as 0.25µL in pre-sealed silver tubes from the United States Geological Service (USGS, <http://isotopes.usgs.gov/lab/referencematerials.html>, USA) with $\delta^{18}\text{O}_{\text{VSMOW}}$ abundance values of 0 and -55.5 ‰ respectively.

5.2.2 Instrumentation and Equipment

Individual samples were dropped from a helium purged Costech Zero Blank autosampler into the graphite furnace of a TC/EA packed with glassy carbon, as shown in Figure 5.1 (Thermo Finnigan, Bremen, Germany). The furnace was set to a temperature of 1400°C and was connected to a 0.5 metre packed molecular sieve heated to 60°C with the helium carrier gas set at a pressure of 120psi. The TC/EA was connected to a ConFlo III and a DELTAplus XP IRMS (Thermo Finnigan). The sample peaks were scaled against pulses of CO reference gas with a purity of 99.999%, set at a peak height of approximately 4000mV, with helium dilution on prior to the sample detection window to ensure that no isobaric interference (peak overlap) from N₂ was encountered. Reference gas pulses were set to 20, 200, 240 and 280 seconds (for 20 seconds duration), with a total measurement time of 340 seconds.

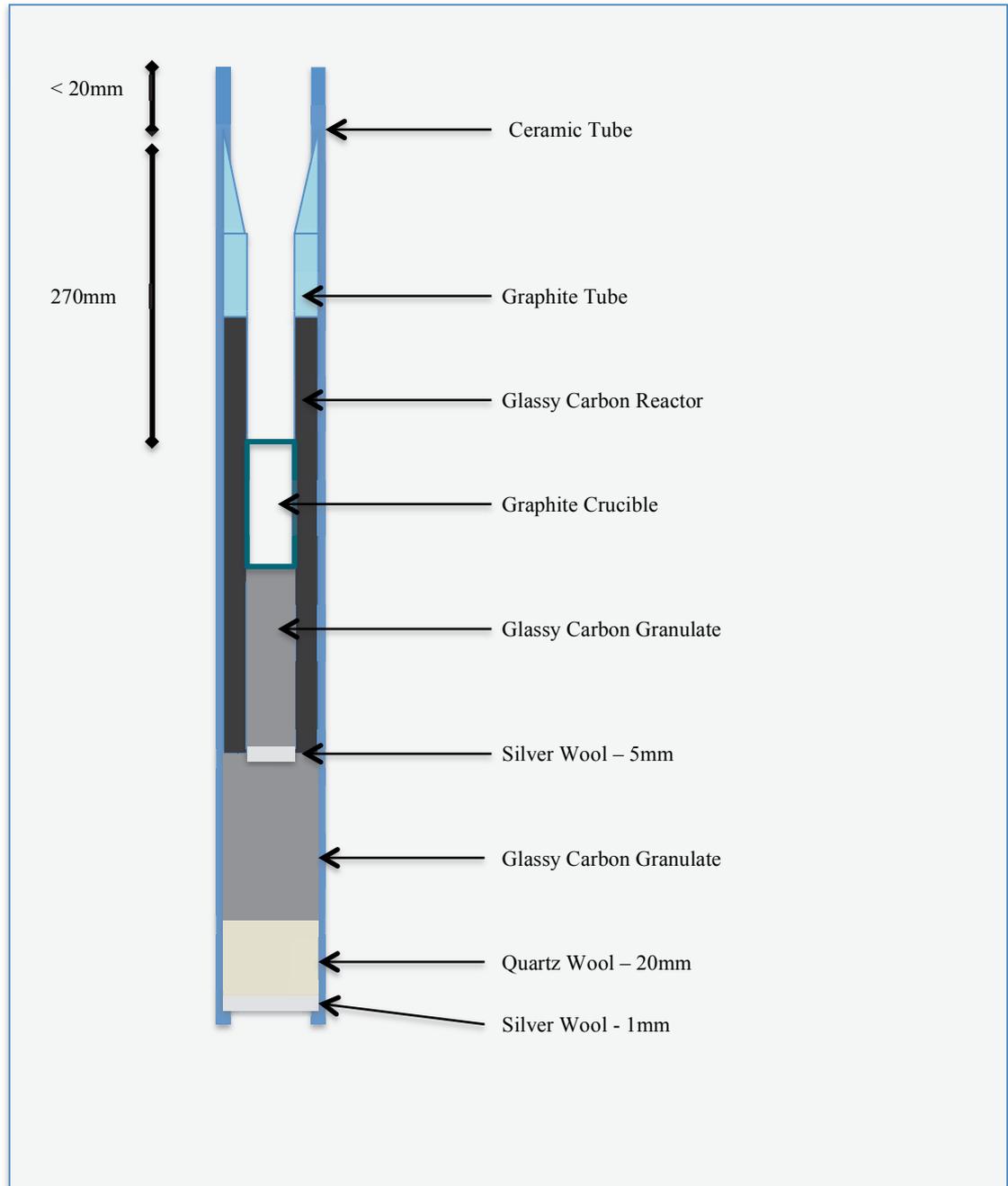


Figure 5.1: Configuration and materials used in the furnace of the TC/EA

5.2.3 Correction of values and statistical tests

The correction of unknown results and the statistical tests undertaken are consistent with those previously described in Chapter 2.

5.2.4 Method Development

The target performance characteristics of the oxygen method are similar to those presented for the carbon method in Chapter 2. To avoid repetition, the sample design experiment results for carbon were adopted and utilised for the measurement of oxygen. This included the number of replicates, placement and use of standard and QA materials within analytical sequences. Throughout these experiments and the measurement of paper values in the following chapters, the adoption of these parameters was observed to produce results that were fit for purpose for oxygen measurements.

The focus of the method development experiments was to ensure that the instrument response was linear for increasing sample weight and hence to determine the most appropriate sample size based on the linear range of the materials. Concurrently, the sample response was observed for isobaric interference to ensure that if it occurred, that the method was optimised to reduce the effects on the sample peak.

The samples utilised for the linearity/sample size experiment are included in Table 14.8. As this experiment aimed to measure instrument response only, a single replicate was measured at each weight point. Additionally, no data correction was made and the raw instrument δ values were used to investigate linear range.

International standard waters VSMOW2 and SLAP2, used to anchor the calibration scale for hydrogen and oxygen isotopic measurements, were run using the method under development here. This was done for two reasons – firstly to check that the results produced using this method, including isotopic scale normalisation processes described in section 1.6, produced results that were traceable to the primary scale anchors. The second reason was to begin a quality control protocol that will be used to monitor instrument performance over time.

5.2.5 Method Validation

The method validation activities described in Chapter 2 were combined into a single reproducibility experiment that was conducted over a period of 12 weeks. The precision and stability of the instrument was assessed during this period, with replicates of the same materials - international standard cellulose IAEA-CH-3, Polyethylene Glycol (PEG) and Coumarin – run in replicates of 8 in weeks 1, 2, 3, 4, 6, 9 and 11. The accuracy of the method was determined by comparing these results to the laboratory or working values either determined in-house (for international cellulose) or by another forensic laboratory (for PEG and Coumarin). To test robustness, the preparation of the samples was shared between three operators over the 12-week period.

The aim of the sample drying experiments was to determine the most appropriate sample handling procedure for removing all absorbed water from paper samples prior to measurement. IAEA-601, IAEA-602, IAEA-CH-3 and three paper samples were dried in an oven for 24 hours at 60°C prior to measurement, with the capsules either crimped or uncrimped. Replicate samples for comparison were prepared but held in a desiccator at room temperature prior to measurement. All samples were measured in triplicate and in one analytical sequence. As some of the target materials are the international standards used for correction, uncorrected raw delta values were used for comparison.

Precision and stability were determined by monitoring the international benzoic acids IAEA-601 and IAEA-602 over an 18-month period. International cellulose was also measured over this period both for monitoring purposes and to calibrate a laboratory value.

The method developed was utilised in two FIRMS inter-laboratory trials, for the measurement of nylon, magnesium stearate and vanillin. Raw values were scale normalised using IAEA-601 and IAEA-602 and 10 replicates were measured and used to compare the AFP laboratory method with the values obtained by a number of different laboratories internationally.

5.3. Results and Discussion

5.3.1 Oxygen Method Development

Linearity

A linearity experiment was conducted to determine whether a consistent instrument response and $\delta^{18}\text{O}$ value was observed for four materials – paper, IAEA-601, IAEA-602 and IAEA-CH-3. The instrument responses vs. sample size values (termed linear response) are shown in Figures 5.2 to 5.5. The linear range values are shown in Figures 5.6 to 5.9. The numerical values are included in Table 14.8.

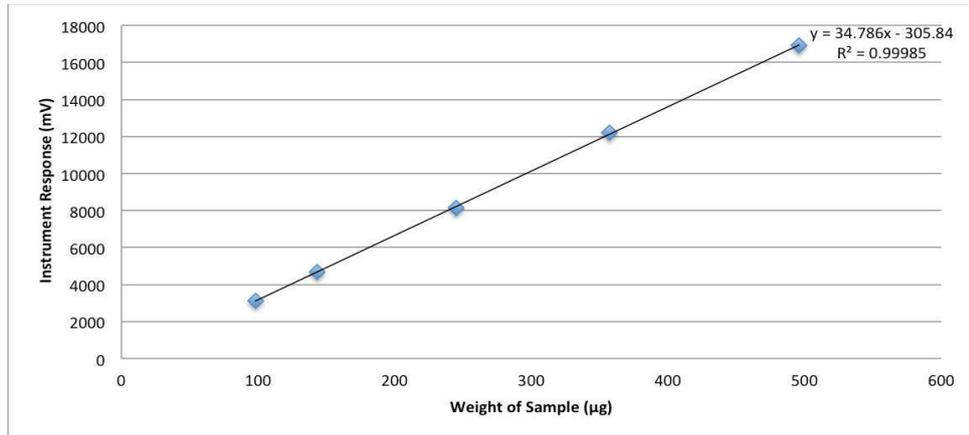


Figure 5.2: Linear response for $\delta^{18}\text{O}$ of paper

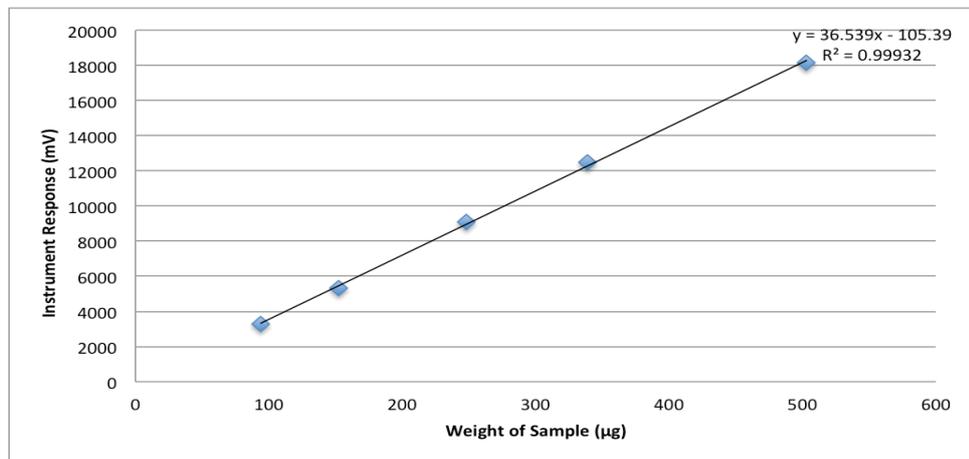


Figure 5.3: Linear response for $\delta^{18}\text{O}$ of IAEA-CH-3

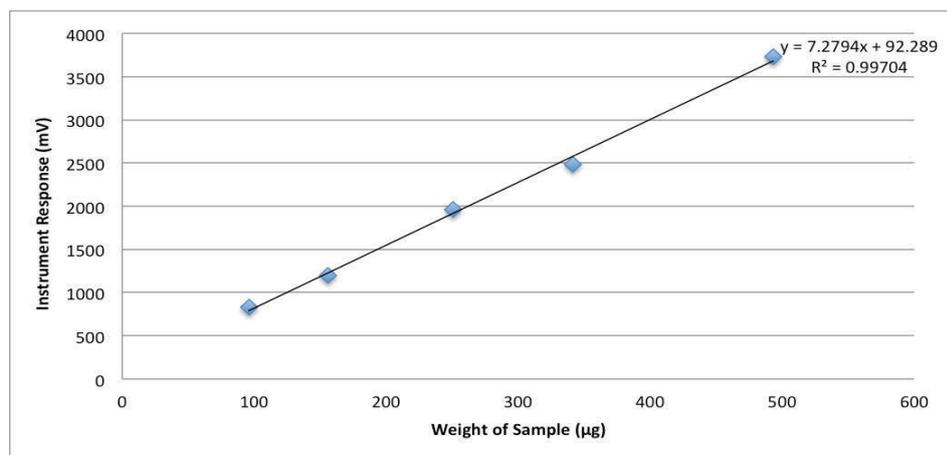


Figure 5.4: Linear response for $\delta^{18}\text{O}$ of IAEA-601

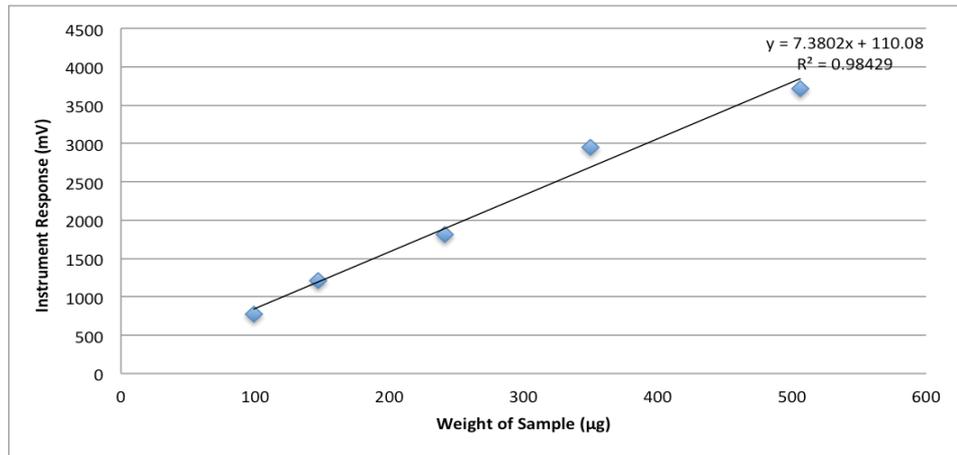


Figure 5.5: Linear response for $\delta^{18}\text{O}$ of IAEA-602

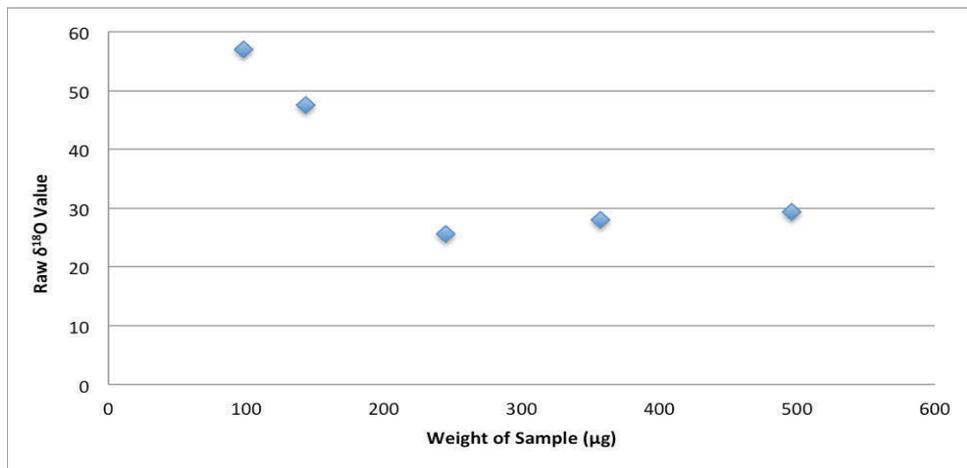


Figure 5.6: Linear range for $\delta^{18}\text{O}$ of paper

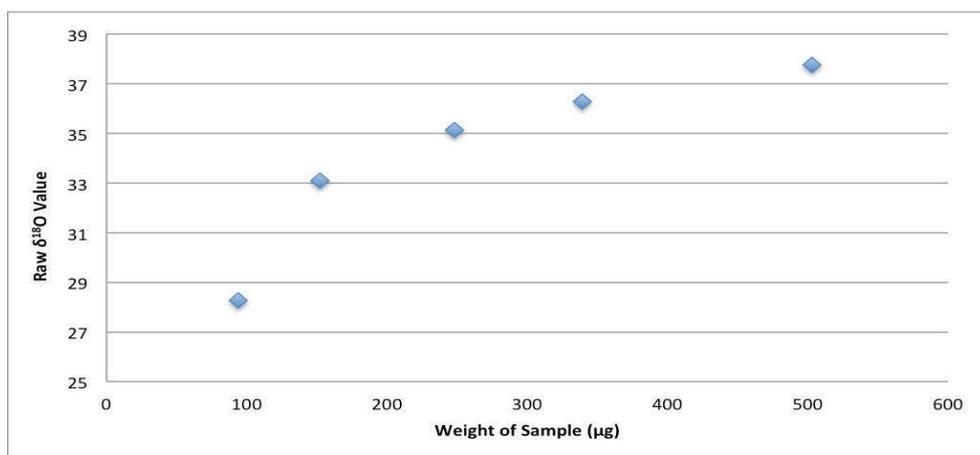


Figure 5.7: Linear range for $\delta^{18}\text{O}$ of IAEA-CH-3

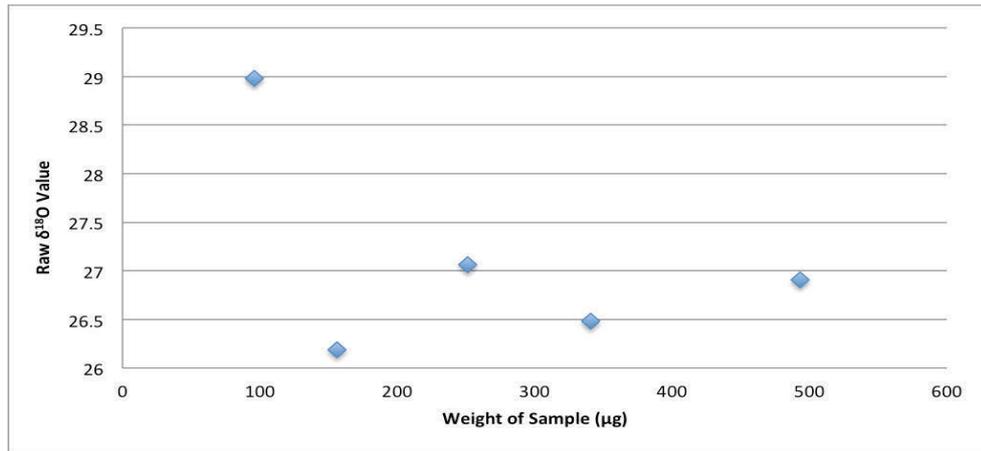


Figure 5.8: Linear range for $\delta^{18}\text{O}$ of IAEA-601

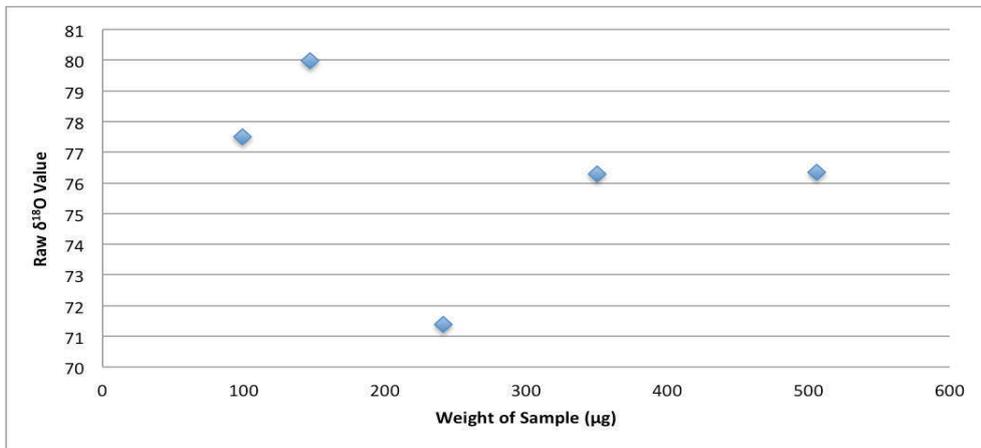


Figure 5.9: Linear range for $\delta^{18}\text{O}$ of IAEA-602

The linear response plots show that the instrument response is linear for increasing sample size. This indicates equal sample conversion with increasing weight, with all linear trend lines except IAEA-602 higher than 0.99. Utilising IAEA-602 over a large number of runs has shown that this standard is quite variable and may be more prone to either absorption of water or sublimation, as shown in its slightly lower linearity value.

The delta values for IAEA-602, as shown in the linear range plot, are also much more variable than observed for the other materials. It was hypothesised that this variability may be due to the crystalline structure of the material, which when being sampled was seen to be more uniform but larger in size than IAEA-601, resulting in more surface area for absorption of water. Attempts to homogenise the material to a finer powder with a mortar and pestle proved unsuccessful, with a large amount of static being produced. This resulted in the powder being impossible to dispense into capsules.

Overall, the linear range plots showed that a sample size greater than 250 μ g is required to achieve reproducibility and stability for the delta values measured. Based on this sample size, the method was shown to produce consistent combustion, with all peaks baseline resolved. An example of a chromatogram for the oxygen measurement of IAEA-601 and for a document paper is shown in Figure 5.10 and Figure 5.11 respectively.

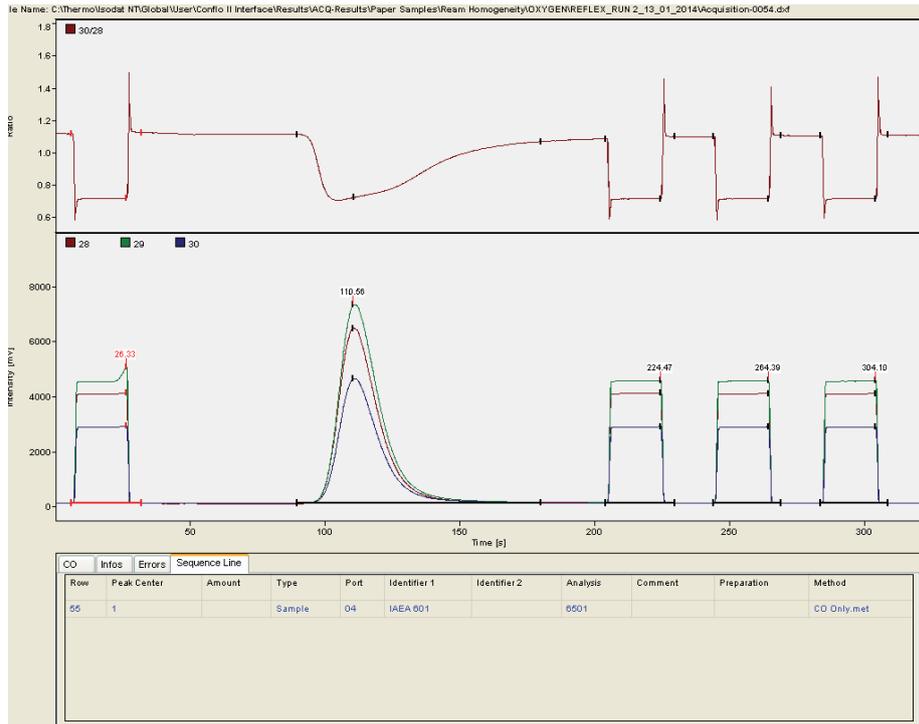


Figure 5.10: Typical chromatogram for measurement of oxygen in IAEA-601

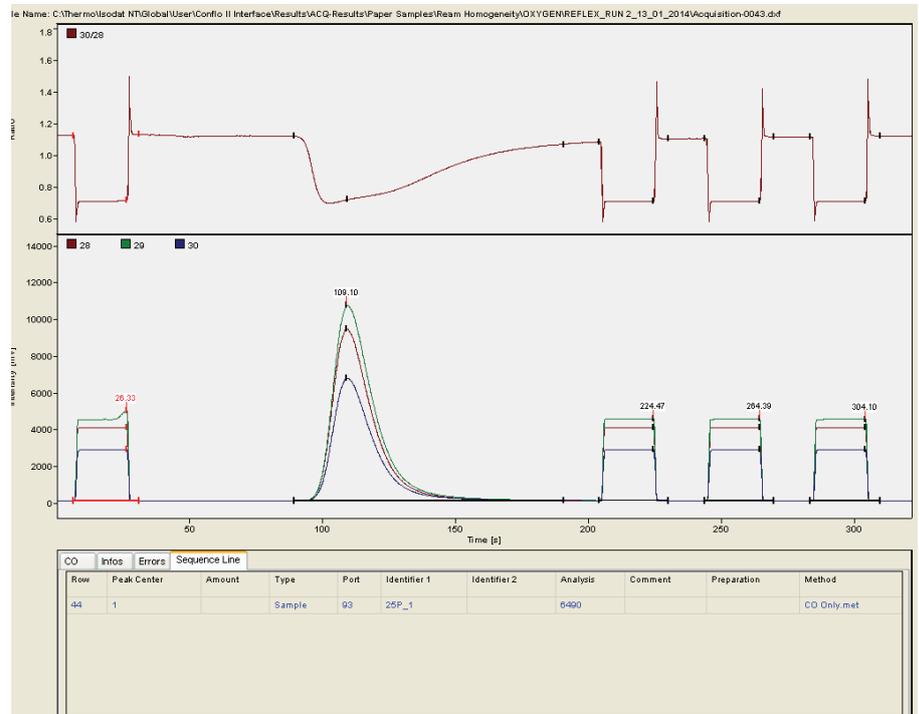


Figure 5.11: Typical chromatogram for measurement of oxygen in document paper samples

VSMOW/SLAP Calibration

Pre-weighed samples of the standard waters VSMOW2 and SLAP2 were obtained from the USGS and run in-house to determine the slope of the calibration line for the VSMOW/SLAP scale. The results for the calibration are shown in Figure 5.12.

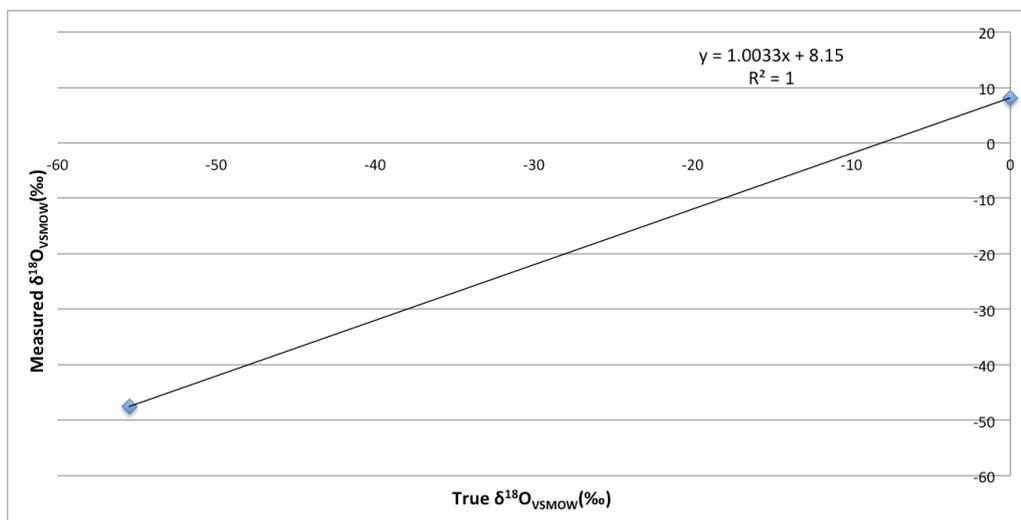


Figure 5.12: Calibration of VSMOW and SLAP2 International Standards

The calibration of VSMOW2 and SLAP2 shows good correlation between the measured and true values (stretch factor of 1.0033). Isotopic scale normalisation procedures were observed to be fit for purpose. The results of these measurements will be used as the start of an internal quality assurance check that will be conducted on the instrument at least every 6 months to ensure consistent instrument performance, which is traceable to international primary reference standards over time.

5.3.2 Method Validation

Reproducibility

The purpose of calibration of unknown samples using standard materials that have been scaled to a master scale (for oxygen this is the VSMOW/SLAP scale) is to ensure international traceability between laboratories. While precision of measurement is most easily observed through the standard deviation of replicate measurements for homogenous materials, comparability of values between runs is essential for tasks such as constructing background databases, international comparison/publication of results

and conducting experiments on large datasets that cannot be measured in a single analytical sequence.

The measurement results for the three unknown materials coumarin, PEG and cellulose over a period of 11 weeks are shown in Table 5.1.

		WEEK 1	WEEK 2	WEEK 3	WEEK 4	WEEK 6	WEEK 9	WEEK 11	Overall Results	Overall, excluding Wk 2
Coumarin	Mean $\delta^{18}\text{O}_{\text{VSMOW}}(\text{‰})$	15.77	14.87	15.56	15.44	15.45	15.78	15.31	15.45	15.55
	St Dev (‰)	0.24	0.20	0.11	0.10	0.14	0.27	0.88	0.31	0.19
PEG	Mean $\delta^{18}\text{O}_{\text{VSMOW}}(\text{‰})$	-17.72	-22.03	-17.93	-18.76	-18.67	-16.98	-16.88	-18.42	-17.82
	St Dev (‰)	0.04	0.18	0.09	0.09	0.34	0.34	0.25	1.75	0.80
Cellulose	Mean $\delta^{18}\text{O}_{\text{VSMOW}}(\text{‰})$	31.26	33.15	31.48	32.08	31.90	31.92	33.19	32.14	31.97
	St Dev (‰)	0.12	0.17	0.05	0.11	0.03	0.17	0.93	0.76	0.67
Stretch factor		0.99	1.09	0.99	1.01	1.01	0.99	0.96		
Shift factor		-8.25	-11.28	-8.46	-8.78	-8.92	-8.62	-6.23		

Table 5.1: Mean $\delta^{18}\text{O}_{\text{VSMOW}}(\text{‰}) \pm 1$ Std Dev (‰) results for replicates of coumarin, PEG and international cellulose run over 12 weeks.

As homogeneity was assumed for these materials, a parametric ANOVA with the Holm-Sidak's multiple comparisons test was utilised to examine the results of the weeks using GraphPad Prism.

Significant differences were observed from the results of the ANOVA test for all three of the unknowns. What was apparent amongst these results however was that significant differences were observed between week 2 and the measurements obtained from other weeks. Significant differences were also observed for the comparisons that included the week 9 results.

Looking more closely at the week 2 results, issues were encountered with the correction of the data due to inconsistent measurement results for the benzoic acid international standards. This was most evident in the large standard deviation values calculated for both standards -1.44% for IAEA-601 and 1.65% for IAEA-602. Due to the large range of variation observed, no single values were identified as outliers using Grubbs test.

Manually excluding a number of the values, reproducibility was most accurately met by manually removing all but two measurement values for each benzoic acid standard. This is not ideal however as reproducibility was clearly not being met in this analytical sequence. The linear calibration equation was calculated to be $y = 1.0174x - 7.0249$. Using this correction, the values obtained for each of the unknowns were still significantly different to the other weeks' values and offset around 2% .

Observing the chromatography of the benzoic acid standards for this week, the average peak height (for mass 28) was 524 and 414mV for IAEA 601 and IAEA 602 respectively. This was mirrored in both sets of standards, at the start and at the end of the analytical sequence, suggesting that this was not an instrument issue. The peak heights for the unknown materials were around 4000mV, as expected, confirming that this was either a sample or sample handling issue originating from the international standard materials. This issue is more closely examined in the drying experiments discussed below and the week 2 results have been excluded from inclusion in the reproducibility results here, as they would have been in a casework situation.

The week 4 results were also observed to have low peak heights for the international benzoic acid standards, which were again reflected in the corrected results for the unknown samples. The peak heights are not as low as the week 2 peaks however, and if the results had been obtained in a casework situation the samples would have been re-measured.

The original examination protocol for the reproducibility experiment included drying of the samples in open silver capsules at 60°C for at least 1 hour prior to crimping and loading onto the autosampler. Ensuring that residual moisture is removed prior to measurement is highlighted in a number of publications (Carter, 2011, Carter and Fry, 2013, Qi and Coplen, 2011), by drying over a desiccant material with or without the addition of heat and vacuum.

In week 2, the prepared analytical sequence was placed in a 60°C oven for approx. 4 hours, with the sample capsules open. After the lower peak heights were observed on the chromatograms for the international standard benzoic acids, a closer look at the properties of the materials was undertaken which found that at 60°C the benzoic acids are likely to partially sublime and hence sample loss would occur (NIST, 2011). The amount of sample loss would increase with the amount of time the sample is held at higher than room temperature, explaining the increased sample loss for the week 2 samples that were left in the oven for longer than the one-hour time period specified in the initial procedure.

In response to these observations, a drying experiment was undertaken to more closely observe the sample loss over time. In addition, a decision was made to remove the drying stage completely from the reproducibility experiments from week 5 onwards.

Overall, reproducibility was not achieved when directly comparing the values of the materials week to week. This is seen in the ANOVA results and in the multiple comparison tests, with all groups included or with outlier/problematic groups removed. This is partially due to the good repeatability and low standard deviation values in individual weeks, but also due to correction issues which have made the results largely incomparable between weeks. The collated standard deviations for the materials, with the week 2 results removed, are higher than the published tolerance for oxygen

measurements of 0.5 ‰ (Carter, 2011) calculated at 0.54 ‰ for coumarin, 0.84 ‰ for PEG and 0.52 ‰ for international cellulose. This indicates that reproducibility was not achieved in this series of measurements.

While reproducibility was not demonstrated, there are a number of key points which once improved would likely solve and meet the requirements of reproducibility. These include:

- More stringent monitoring and exclusion of analytical sequences when the results obtained deviate from expected behaviours;
- The use of a QC chart for international standard materials; and
- The use of matrix matched QC materials (and QC charts).

It is evident that the selection of materials used in this reproducibility experiment was not ideal. The use of PEG as an unknown material, which is clearly outside of the bracketed values of the international benzoic acid standards is not recommended. Likewise, coumarin sits outside the bracketed values, though not to the extent of PEG. This is highlighted as an issue for calibration in a number of publications (Werner and Brand, 2001, Meier-Augenstein, 2010) as a core principle of standard material selection. Realistically, the inclusion of these materials as unknowns was a decision that was made in an attempt to validate materials that were the most similar to materials of forensic interest with values that had been defined by other laboratories.

It should also be noted here that while reproducibility was not achieved, the precision of the method was demonstrated within single weeks with low standard deviation values for all materials observed. This means that samples that are run within the same analytical sequence are readily comparable and can be used to characterise and compare samples for casework purposes. Future work to address these reproducibility issues is required, which will need to focus primarily on international and quality control standards so that cross comparison of results between runs will be able to be achieved. Overall, these experiments showed that reproducibility cannot be *assumed* and that continuous close monitoring and exclusion of sequences that show issues with the chromatography or scale normalisation is required.

The implications of these findings on this project are minimal as issues with reproducibility and between-sequence comparisons were noted and accounted for in the designation of the benchmark values for discrimination presented in Chapter 6. Reproducibility was also increased due to the consistency of the sequences run and the use of a single material type, which was properly bracketed within the international standards utilised. The measured value for the QC material used within the paper sequences (IAEA-CH-3) was also comparable across runs.

Robustness was included as a factor in these experiments by varying the operators used to prepare and run the samples. No observable difference could be attributed to the operator.

Sample Drying

To determine a method for removing excess water from the samples, within the current capabilities and set-up of the AFP laboratory, an oven drying experiment was undertaken with IAEA-601, IAEA-602, IAEA-CH-3 and three papers. The mean raw $\delta^{18}\text{O}_{\text{VSMOW}}$ values are shown in Table 5.2.

		$\delta^{18}\text{O}_{\text{VSMOW}}$ Uncrimped, Oven Dried	$\delta^{18}\text{O}_{\text{VSMOW}}$ Crimped, Oven Dried	$\delta^{18}\text{O}_{\text{VSMOW}}$ Uncrimped, Undried
IAEA-601	Mean	Nil sample peaks	32.47	32.07
	St Dev		0.10	0.08
IAEA-602	Mean	Nil sample peaks	79.23	81.39
	St Dev		0.77	0.39
International Cellulose	Mean	41.10	41.01	40.37
	St Dev	0.10	0.11	0.30
Paper 1	Mean	35.61	35.44	35.40
	St Dev	0.11	0.08	0.12
Paper 2	Mean	38.24	38.05	37.92
	St Dev	0.02	0.01	0.00
Paper 3	Mean	33.61	33.37	34.26
	St Dev	0.18	0.05	0.00

Table 5.2: Summarized data for sample drying experiments

The drying experiments showed that there was complete sample loss for IAEA-601 and IAEA-602 when the capsules were left uncrimped in the oven. For the crimped samples that were dried in the oven, the peak heights were reduced by almost 50%. This shows that drying the sample at 60°C is not a suitable preparation for benzoic acids, or that the drying duration used in this experiment was too long.

There were no visible trends in the standard deviations for the international cellulose or paper samples between dried and undried. Overall, holding the samples in a desiccator was observed to be an appropriate sample handling method for these types of samples with relatively low standard deviation values observed.

Precision and Stability

Throughout the experiments detailed here for the measurement of the oxygen isotopes of paper, the results for IAEA-601, IAEA-602 and IAEA-CH-3 were collated as quality control charts to understand the performance and accuracy of the instrument over time. These are included as Figure 5.13, Figure 5.14 and Figure 5.15 with the data summarised in Table 5.3. For these figures, the red line denotes the grand mean value for this sample set, while the dashed lines are 2 x the standard deviations.

Overall the instrument was found to produce comparable results, with some trends over time. Deviations away from the mean expected value are observed particularly in the IAEA-CH-3 plot where values can be seen to rise in groups. This is due to correction and in particular, correction issues due to the benzoic acid international standards. The comparability of results appears to be increasing over time however as the results are becoming more consistent and trending closer to the grand mean of all measurements.

Instrument stability is observed to be increasing over time as the method was developed, as shown in the IAEA-601 plot. The reproducibility of IAEA-602 is an issue however, with a number of values consistently sitting outside of the 95% confidence intervals plotted.

Due to the stability issues and the linearity results, caution needs to be used when using IAEA-602 for correction of casework samples. While it would be prudent to remove this standard from use all together, there were no alternative organic international standard materials available as a substitute. To address these issues, a number of water standards (VSMOW, SLAP2 and GISP) will be used to calibrate a suitable working material to replace IAEA-602 in the future, outside of this project.

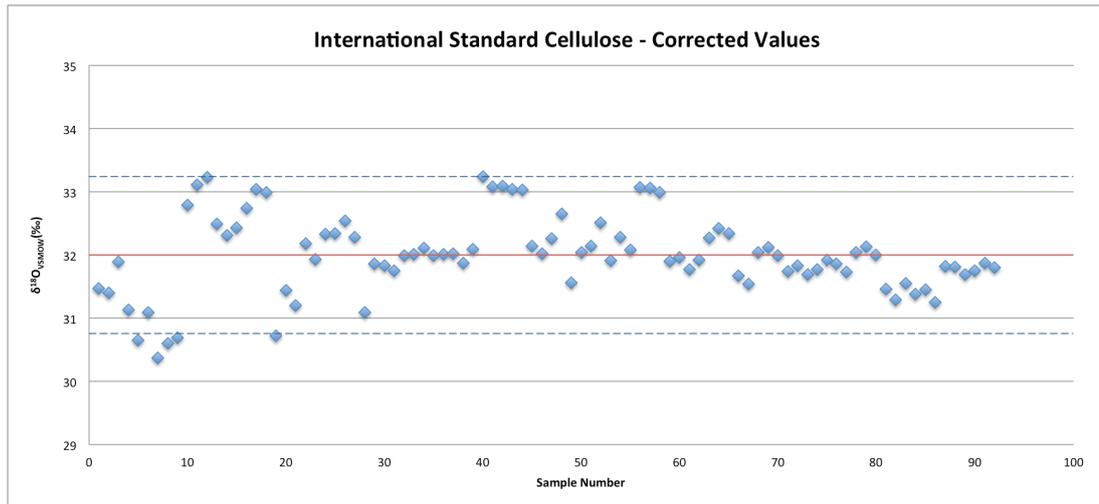


Figure 5.13: IAEA-CH-3 measured as a quality control material over the course of 18 months of experiments.

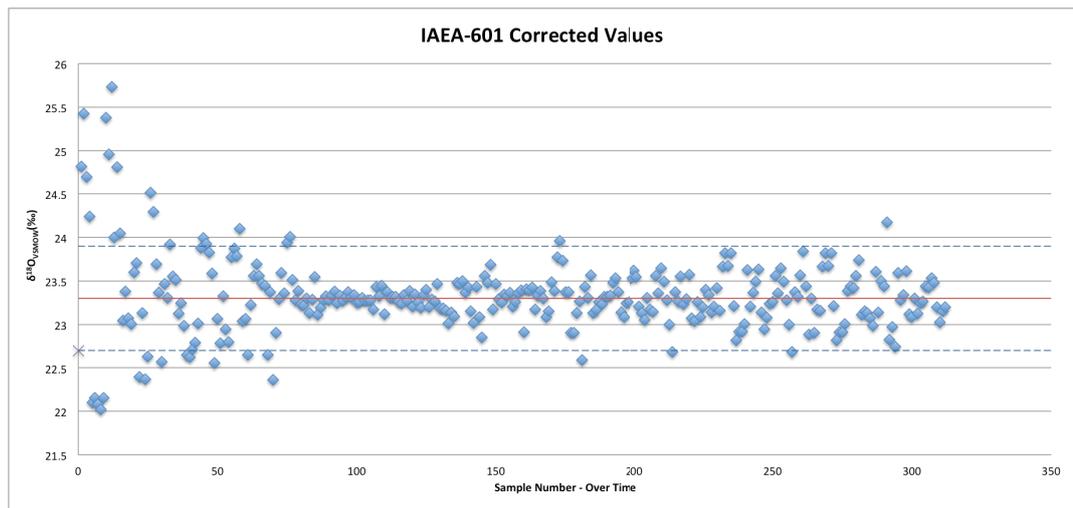


Figure 5.14: Corrected values for IAEA-601, used as a standard material over the course of 18 months of experiments.

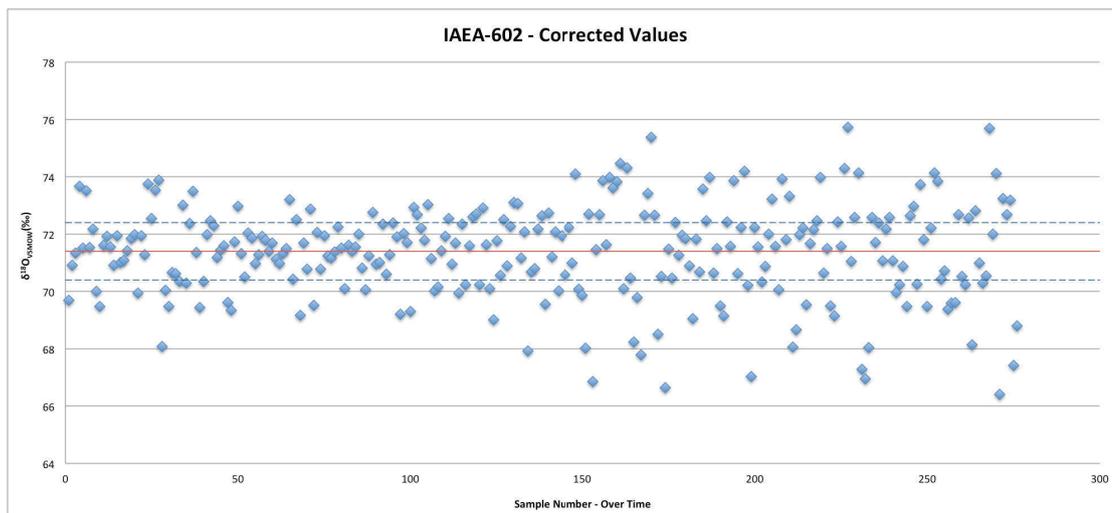


Figure 5.15: Corrected values for IAEA-602, used as a standard material over the course of 18 months of experiments.

	$\delta^{18}\text{O}$ Mean (‰)	St Dev (‰)	95% CI (‰)	Number of Samples	$\delta^{18}\text{O}$ Published Value (‰)	Published St Dev (‰)
IAEA-601 Benzoic Acid	23.33	0.44	0.89	312	23.3	0.3
IAEA-602 Benzoic Acid	71.38	1.63	3.26	276	71.4	0.5
International Cellulose	31.99	0.62	1.24	92	N/A	N/A

Table 5.3: Summary of the analytical results obtained for three international standard materials, measured over time for their $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values

Inter-laboratory trials

The method validated here was utilised in two inter-laboratory trials over the course of 12 months to analyse three materials. An explanation of the calculation of z-scores is included in Chapter 2.

The z-scores obtained for the three materials nylon, magnesium stearate and vanillin were 0.32, -0.09 and -0.62 respectively. These scores show that the results obtained are comparable to other laboratories with all z-scores calculated at a score less than 1.

5.4. Conclusions

The method presented here was shown to be precise, accurate and to produce results that are comparable within an analytical sequence. Issues with comparability between sequences were encountered with reproducibility not being met between cellulose and other organic samples run on different days. Future work to be undertaken to improve these results will include the calibration of matrix matched laboratory standard materials with greater homogeneity than the current international standard benzoic acids, in particular IAEA-602.

Sample handling was shown to be a key issue for the measurement of oxygen isotopes, in particular for the benzoic acid international standards that were shown to sublime if left above room temperatures (i.e. if dried in an oven) for any length of time. There was no effect on the $\delta^{18}\text{O}_{\text{VSMOW}}$ values of paper or international cellulose from the drying processes tested indicating sample stability for these sample types. Storing samples to dry in a desiccator was shown to be an appropriate alternative treatment for samples.

Chapter 6

6. Oxygen Isotopes – Background Population and Homogeneity

The key results of this chapter have been published in (Jones et al., 2016). The published article is included electronically within Appendix 5.

6.1. Introduction

As prior studies utilizing isotope ratio mass spectrometry of paper have shown, a greater level of discrimination can be gained with the use of multiple stable isotopes (van Es et al., 2009). For an organic sample such as paper, the high cellulose content of the material lends itself towards the use of oxygen and hydrogen isotopes in addition to the carbon isotopes previously studied. These two isotopes, and hydrogen in particular, come with inherent problems related to absorption and exchange of water from the atmosphere to the surface of the cellulose molecules. Given this, oxygen was selected as a more consistent isotope to add to carbon in an attempt to increase the discrimination of papers using IRMS.

Once a precise and accurate analytical method was defined, the primary aim of this chapter was to conduct a background study of the same paper sample set measured for carbon isotopes. To aid in interpretation, the homogeneity studies utilised for carbon isotopes were also undertaken here to determine the relative consistencies or inconsistencies within and between different brands or manufacturers of paper.

6.2. Materials and Methods

The analytical method used for these measurements is outlined in Chapter 5. All samples were run against international standard benzoic acids IAEA-601 and IAEA-602, with international cellulose IAEA-CH-3 acting as a quality control material.

The background paper samples run are detailed in Table 14.3. Likewise, the paper samples measured for the homogeneity experiments are the same as those in Table 3.1, excluding the substitutions of HP Everyday and Paper One branded samples (originating from Brazil and Indonesia respectively) which have replaced the Double A and Olympic samples from the first round of testing.

The standards and samples measured in these experiments did not include any drying prior to measurement, except being held in a desiccator at room temperature before immediately being transferred to a sealed helium purged auto-sampler.

6.3. Results and Discussion

6.3.1. Background Population Study – Oxygen Isotopes

123 of the 125 background samples measured for carbon were measured for their oxygen isotopic values, and were measured between 17 ‰ and 32 ‰. The background papers measured are plotted in Figure 6.1 as a scatter plot, and organized via their region of origin in Figure 6.2 and Figure 6.3. The summarized data is included in Table 14.3.

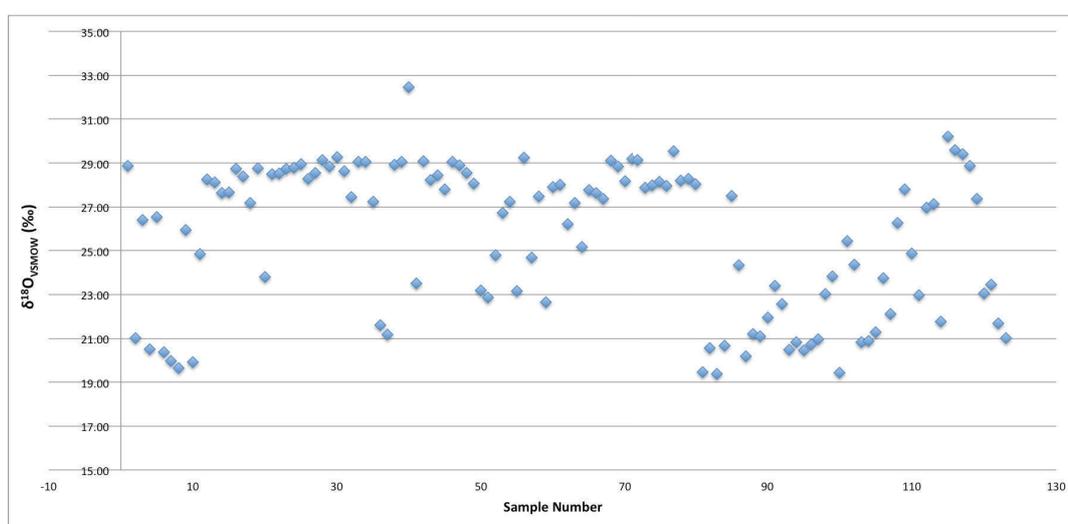


Figure 6.1: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values of 123 virgin pulp papers collected from Australia and New Zealand

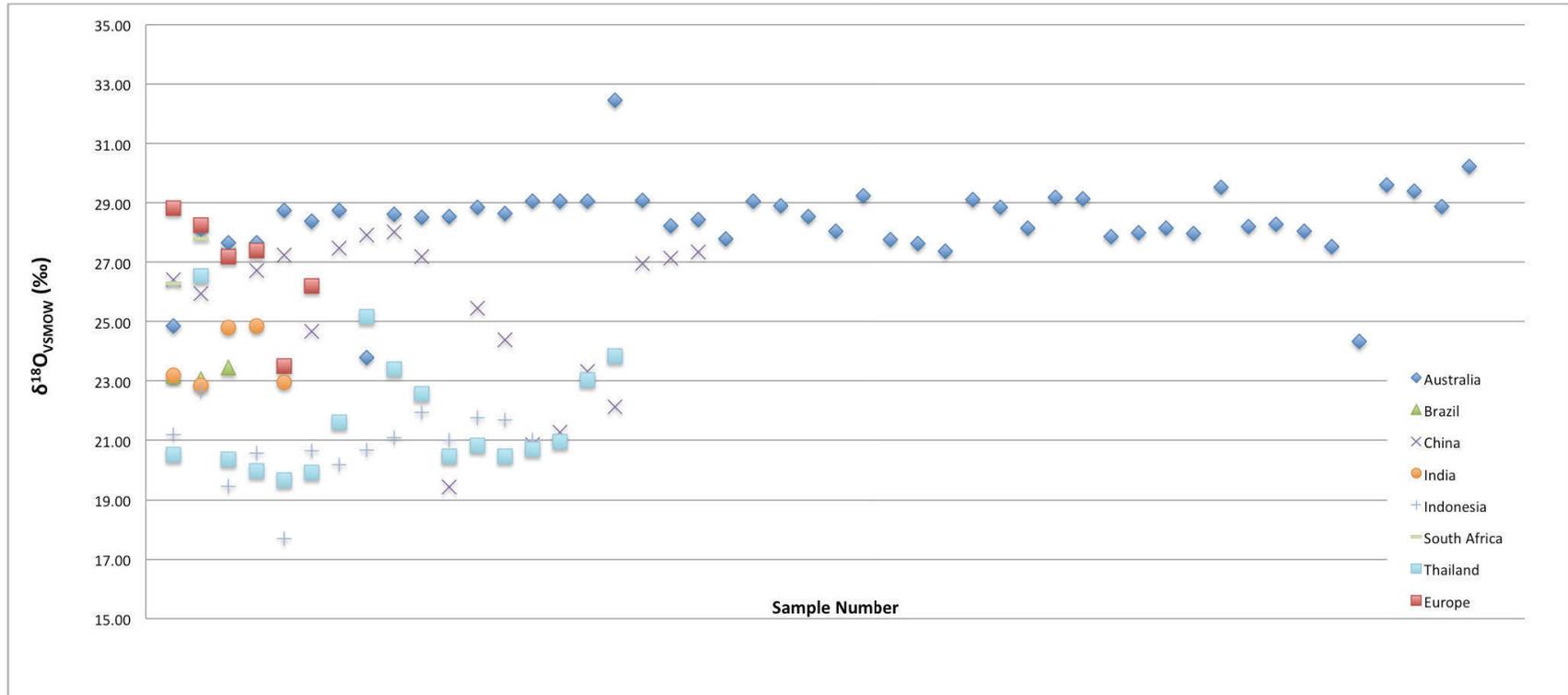


Figure 6.2: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values of 123 virgin pulp papers plotted by region of origin

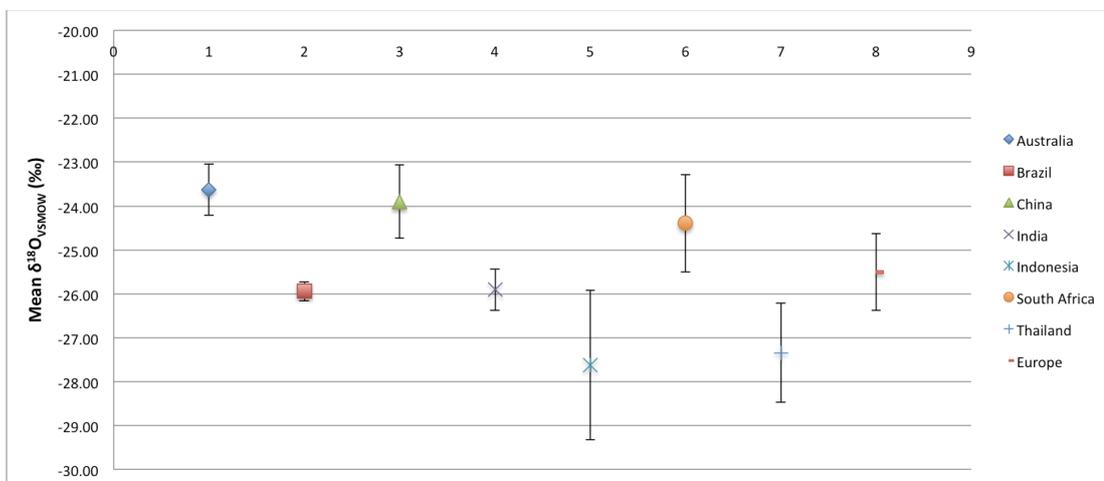


Figure 6.3: Summary of mean $\delta^{13}C_{VPDB}$ (‰) for region of origin of papers, with error bars denoting 1 x standard deviation

A relationship is observed between region of origin and $\delta^{18}O$ value. The strength of this relationship appears to be weaker for oxygen than for carbon however.

Due to the wider range observed, there is likely a larger potential for discrimination of samples using oxygen isotopic abundance values than carbon. With respect to the ability to assign a region of origin, this appears to be more complex than for carbon, due to the increased overlap of values. This is clearly shown in Figure 6.3, with for example, the Australian, Chinese and South African papers overlapping.

Comparing Australian virgin pulp papers with the Australian recycled papers in Table 6.1, the seven samples measured sit within the expected range for this region. This indicates that recycled papers are comparable to the Australian region for provenancing purposes.

Brand Name	% Recycled Content	Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰)	Standard Deviation (‰)
Australian	10%	27.31	0.11
Australian	10%	27.13	0.00
Australian	10%	27.35	0.17
Australian	10%	28.51	0.17
Australian	10%	27.46	0.09
EXPGreen50	50%	28.17	0.13
Reflex Recycled	50%	27.65	0.04

Table 6.1: Recycled papers, mean $\delta^{18}\text{O}_{\text{VSMOW}}$ and standard deviation for 7 recycled papers collected as samples for comparison to virgin paper background samples

Australian papers within the background population were plotted over time utilising the packing date contained on the ream packaging, as shown in Figure 6.4. The reams were produced between April 2009 and February 2013. No significant trends can be observed that could be interpreted as cyclical or representative of seasonality however, there are some differences in the values. This difference lies in the same time period as was previously observed for the carbon results. This suggests that there may have been a manufacturing or source material value shift during this period. Sample numbers 10 and 15 in Figure 6.4 refer.

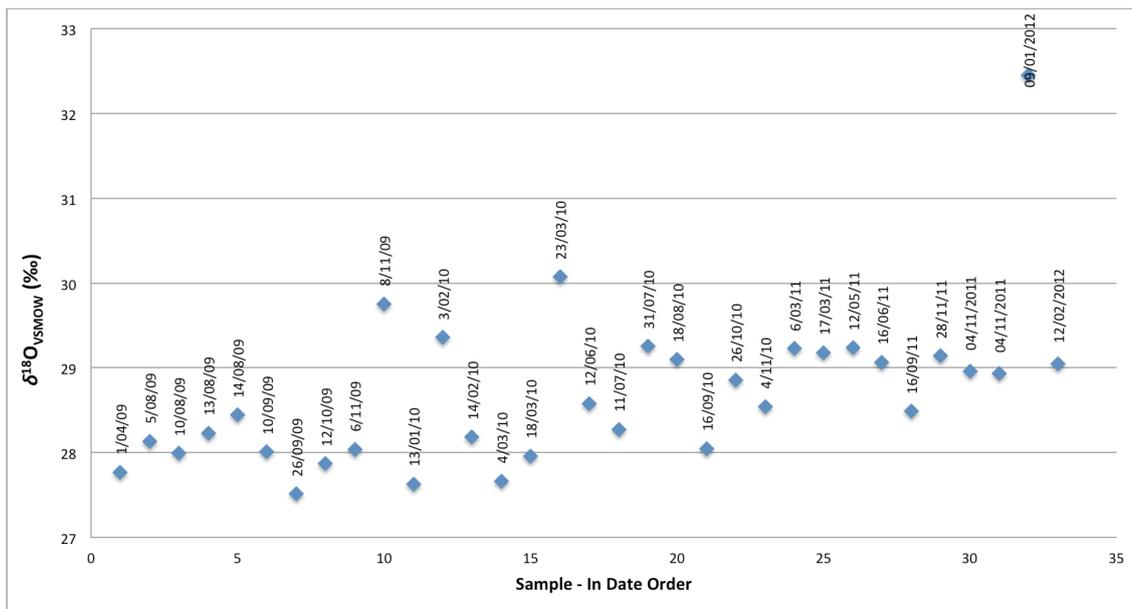


Figure 6.4: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values of Australian paper samples collected over time, plotted in order of their packing date

To further assess whether the shifts in values are due to seasonality, an average per season was calculated and plotted in Figure 6.5. While there does appear to be some movement in the values between groups, it is not cyclical or repeatable between seasons. There does appear to be an overall trend towards enriched values however, which is likely to be due to a source material change since the enrichment is consistent across seasons.

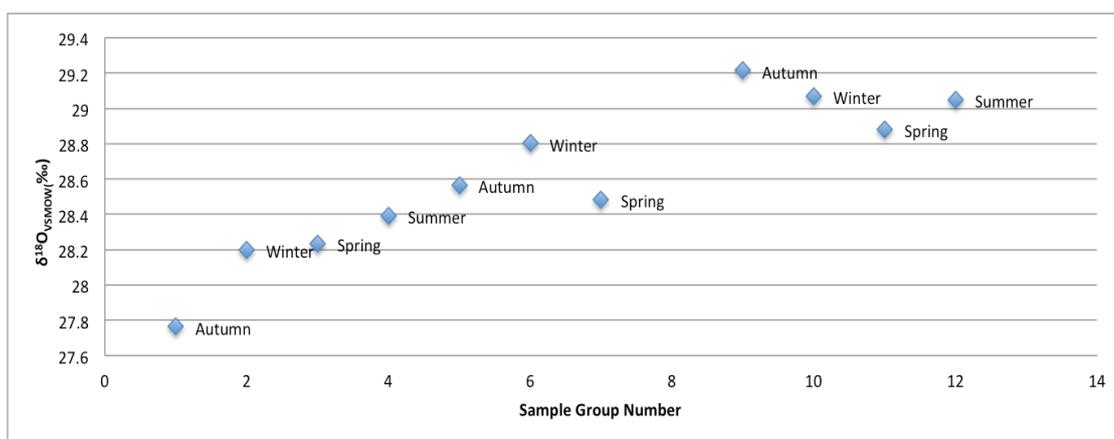


Figure 6.5: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values of Australian paper samples averaged per season, from 2009 until 2012.

6.3.2. Paper Homogeneity Within One Ream

The mean and standard deviation results for the paper homogeneity measurements taken within a single ream are detailed in Table 6.2. The sample names have been abbreviated where S= sample, P = page and L = location.

		1. FUJI XEROX	2. OFFICEMAX	3. REFLEX	4. LAZER IT	5. DOCUCOPY	6. HP EVERYDAY	7. PAPER ONE
Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ of Paper Samples (‰) $\pm 1\sigma$	3S_1P_1L	26.56 \pm 0.10	26.56 \pm 0.11	29.33 \pm 0.21	21.16 \pm 0.15	22.45 \pm 0.11	24.12 \pm 0.10	21.82 \pm 0.06
	7S_1P_1L	26.65 \pm 0.10	26.45 \pm 0.04	29.58 \pm 0.08	21.24 \pm 0.15	22.23 \pm 0.31	24.24 \pm 0.15	21.79 \pm 0.10
	3S_1P_1R	26.64 \pm 0.10	26.59 \pm 0.11	29.44 \pm 0.08	21.26 \pm 0.11	22.11 \pm 0.20	23.95 \pm 0	21.77 \pm 0.13
	7S_1P_1R	26.74 \pm 0.06	26.56 \pm 0.11	29.65 \pm 0.07	21.32 \pm 0.13	21.96 \pm 0.12	24.08 \pm 0.15	21.73 \pm 0.35
	7S_7P	26.77 \pm 0.10	26.53 \pm 0.06	29.67 \pm 0.12	21.22 \pm 0.10	22.10 \pm 0.30	24.19 \pm 0.32	21.80 \pm 0.07
	25S_25P	26.42 \pm 0.16	26.81 \pm 0.17	29.52 \pm 0.11	21.59 \pm 0.27	22.86 \pm 0.31	23.16 \pm 0.21	21.86 \pm 0.23
Total # samples		49	49	51	49	49	47	50
Grand Mean		26.57 \pm 0.19	26.68 \pm 0.19	29.55 \pm 0.14	21.42 \pm 0.27	22.51 \pm 0.47	23.67 \pm 0.54	21.82 \pm 0.21
95% confidence interval		0.38	0.38	0.28	0.54	0.93	1.02	0.42

Table 6.2: Measured $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) abundances from single ream homogeneity testing.

With reference to a 95% confidence interval for discrimination, the results demonstrate the potential for discrimination of some brands within a single ream. This means that, in agreement with the carbon results, a 95% confidence interval is inappropriate for use. Avoiding the use of this kind of discrimination will prevent the false exclusion of results, aid in establishing an appropriate ‘source level’ for this material and take into account the inhomogeneity of the material. The potential for discrimination within a single ream is shown in the box-plot for the Docucopy ream below in Figure 6.6e where discrimination between a number of groups (that had varied in sampling) would have occurred if a 95% confidence interval was used.

Figure 6.6 (a-g) also demonstrates that for the reams with groups that are inhomogeneous, there is a greater magnitude of difference between the groups than seen in the carbon results. The standard deviations for each of the groups, plotted in Figure 6.7, show that for the measurement of oxygen isotopes the standard deviation is likely to sit between 0.1 and 0.2 ‰ but not higher than 0.35 ‰. Defining the benchmark value as 2 x the highest likely standard deviation for a sample group would derive a value of 0.7 ‰ on either side of the sample measurement (a total range of 1.4 ‰). Taking into account the grand mean of the sample for each brand however, this would be extended to a value of 1 ‰ (total range of 2 ‰).

Looking at the results for the Kruskal-Wallis ANOVA test and the boxplots in Figure 6.6 (a-g), statistically significant differences were observed between the means of the 25-page sample group when compared to other sampling groups for five of the seven brands (all brands excluding Reflex and Paper One). There are two likely reasons why these consistent differences were observed. When conducting this experiment, due to the number of samples being measured, two experimental sequences were utilised. The first contained the smaller sampling groups while the second contained only the 25-page sample group. This may have contributed to a small offset being introduced between the analytical sequences that are likely due to small differences in isotopic scale normalisation between sequences.

Alternatively, the 25-page sample group may be better at recording the true mean of the ream being measured. While the use of international standards for correction is designed to protect against these kinds of effects, the behaviour of IAEA-602 in

particular makes this a credible explanation. Given that the offset is seen in a number of brands (as shown in the box plots), an instrument effect seems more likely than a sampling effect. Where differences between the paper samples run in different sequences, the effect was mirrored in the international cellulose values. As international benzoic acid standards were responsible for correction of unknown values, they could be corrected to the assigned value without issue.

Given these observations, while the use of a higher benchmark value does introduce a limitation to the discrimination power achieved, a second extended benchmark value is proposed to allow for these findings. This proposal is supported by the reproducibility experiment in chapter 5, that showed that the method may not produce results that are as directly comparable between runs as first assumed.

In terms of sampling, there was no significant difference between the mean values obtained when sampling 3 or 7 samples and when taking samples from a single or multiple (random) locations on a page. Observationally, there is some suggestion from the plot of the standard deviations (Figure 6.7) that sampling from a single location may capture more of the variation seen in a single page, however it is not apparent that this is the case.

While there is no statistical evidence to support it, best practice would suggest that when able, 7 samples should be taken to characterize a single page. This should be considered in terms of sequence length however, with preference given to running samples for comparison within one analytical sequence to maximise comparability.

Logically it would seem to be inappropriate to take less than 7 measurements from 7 pages as an indication of the true value of a ream, given that a ream is comprised of 500 sheets. In terms of characterizing a ream, there is no statistically significant benefit derived from taking 25 samples from 25 pages to characterize a ream over 7 samples from 7 pages. Given this, in casework situations the recommended number of sheets to sample for a ream (or close to a ream) of paper is as many samples as practicable but no less than seven sheets.

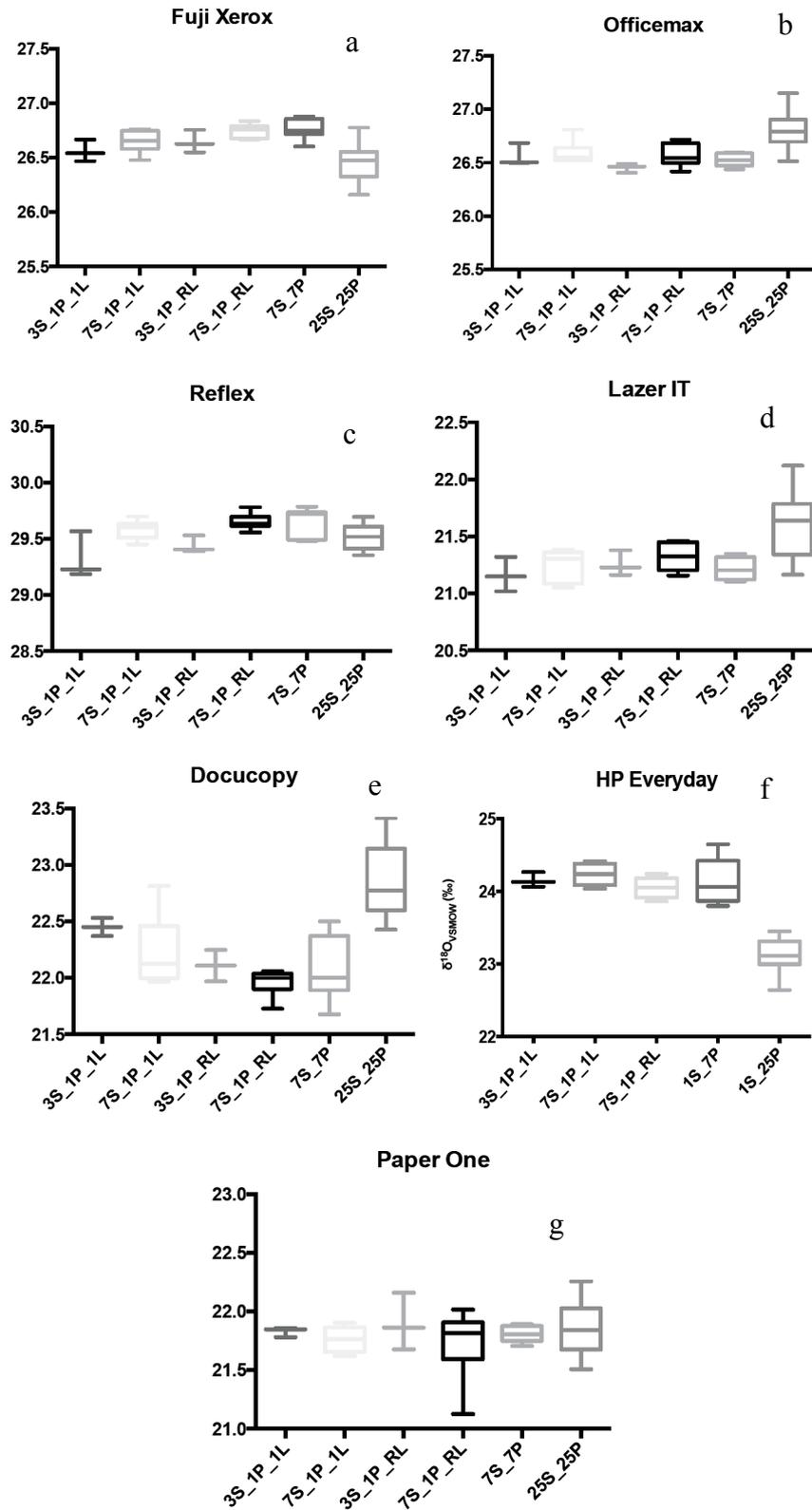


Figure 6.6 (a-g): Boxplots of $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values for single ream homogeneity

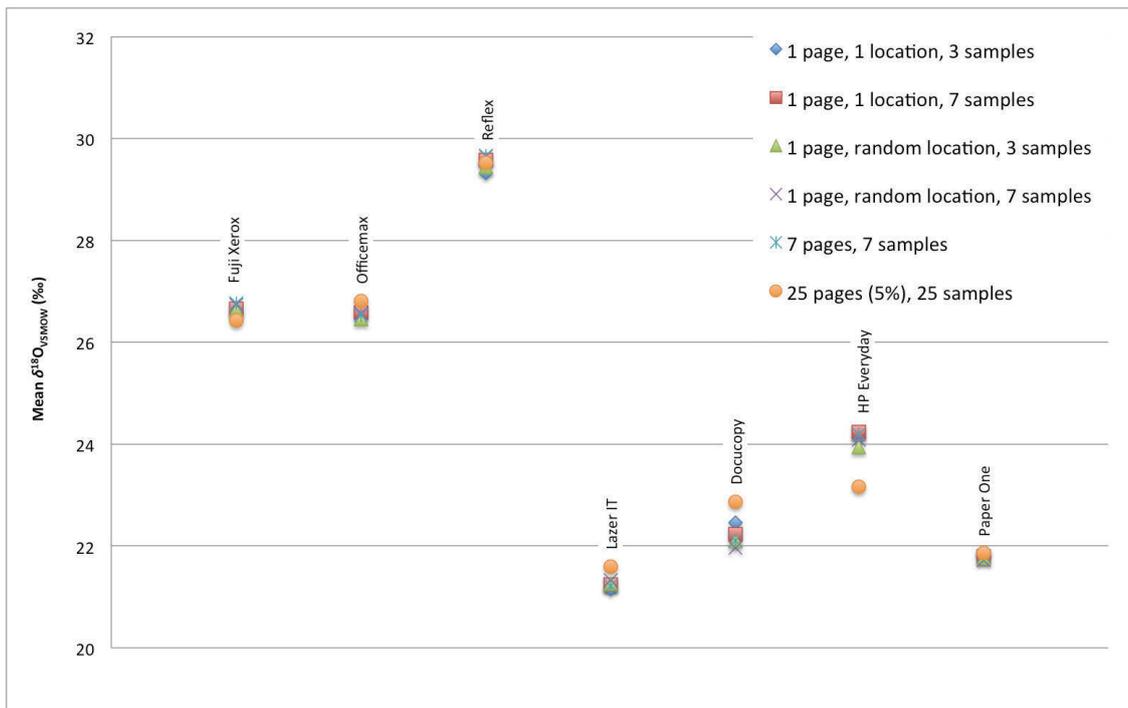


Figure 6.7: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) values for seven different brands measured during homogeneity and sampling tests.

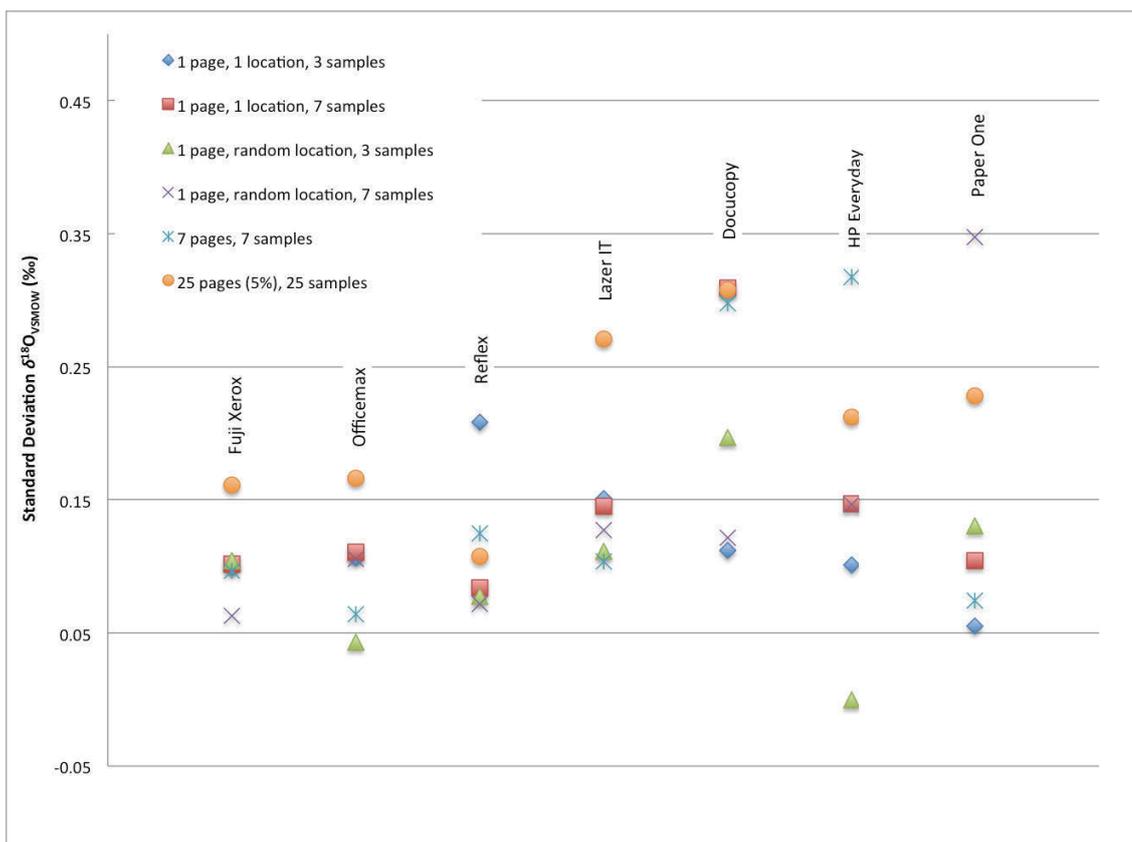


Figure 6.8: Standard deviation values (‰) for homogeneity and sampling test

6.3.3. Homogeneity Between Reams of the Same Brand

To determine the ability of isotopic abundance to discriminate between reams from the same source, and to examine the expected level of homogeneity within a single brand of paper, 7 reams from the same 7 brands were measured for their $\delta^{18}\text{O}$ values. The mean and standard deviation for the seven reams is detailed in Table 6.3.

The standard deviation per ream was the highest for the Docucopy branded reams. This was followed by the Lazer IT and Officemax brands. The Docucopy and Officemax reams also had the highest standard deviations when measured for $\delta^{13}\text{C}$, as reported in Chapter 3. The boxplots for these reams are shown in Figure 6.9 with the error bars set to represent the 95% confidence interval.

	Ream	1. FUJI XEROX	2. OFFICEMAX	3. REFLEX	4. LAZER IT	5. DOCUCOPY	6. OLYMPIC	7. DOUBLE A
Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ of Paper Samples (‰) $\pm 1\sigma$	A	26.59 \pm 0.4	26.53 \pm 0.46	29.48 \pm 0.20	20.48 \pm 0.36	21.64 \pm 0.44	29.02 \pm 0.21	23.95 \pm 0.10
	B	26.40 \pm 0.16	26.24 \pm 0.23	29.54 \pm 0.22	20.46 \pm 0.11	21.23 \pm 0.30	28.52 \pm 0.18	23.91 \pm 0.15
	C	26.36 \pm 0.06	26.28 \pm 0.40	29.38 \pm 0.23	20.65 \pm 0.30	21.16 \pm 0.53	28.59 \pm 0.19	23.83 \pm 0.11
	D	26.32 \pm 0.08	26.19 \pm 0.24	29.33 \pm 0.11	20.43 \pm 0.34	20.92 \pm 0.49	28.57 \pm 0.18	23.76 \pm 0.24
	E	26.26 \pm 0.03	26.16 \pm 0.29	29.27 \pm 0.14	20.31 \pm 0.33	20.99 \pm 0.23	28.40 \pm 0.17	23.63 \pm 0.27
	F	26.31 \pm 0.06	25.09 \pm 0.24	29.18 \pm 0.15	20.52 \pm 0.30	21.19 \pm 0.40	28.45 \pm 0.10	24.25 \pm 0.12
	G	26.32 \pm 0.06	25.02 \pm 0.20	29.07 \pm 0.07	20.18 \pm 0.34	21.04 \pm 0.58	28.45 \pm 0.14	21.61 \pm 0.17
Total N		41	48	49	48	43	42	36
Grand Mean		26.37 \pm 0.12	25.92 \pm 0.64	29.32 \pm 0.22	20.43 \pm 0.32	21.18 \pm 0.47	28.57 \pm 0.25	23.53 \pm 0.9
95% Confidence Interval		0.24	1.27	0.44	0.64	0.93	0.50	1.8

Table 6.3: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) and standard deviation (‰) results for between ream homogeneity.

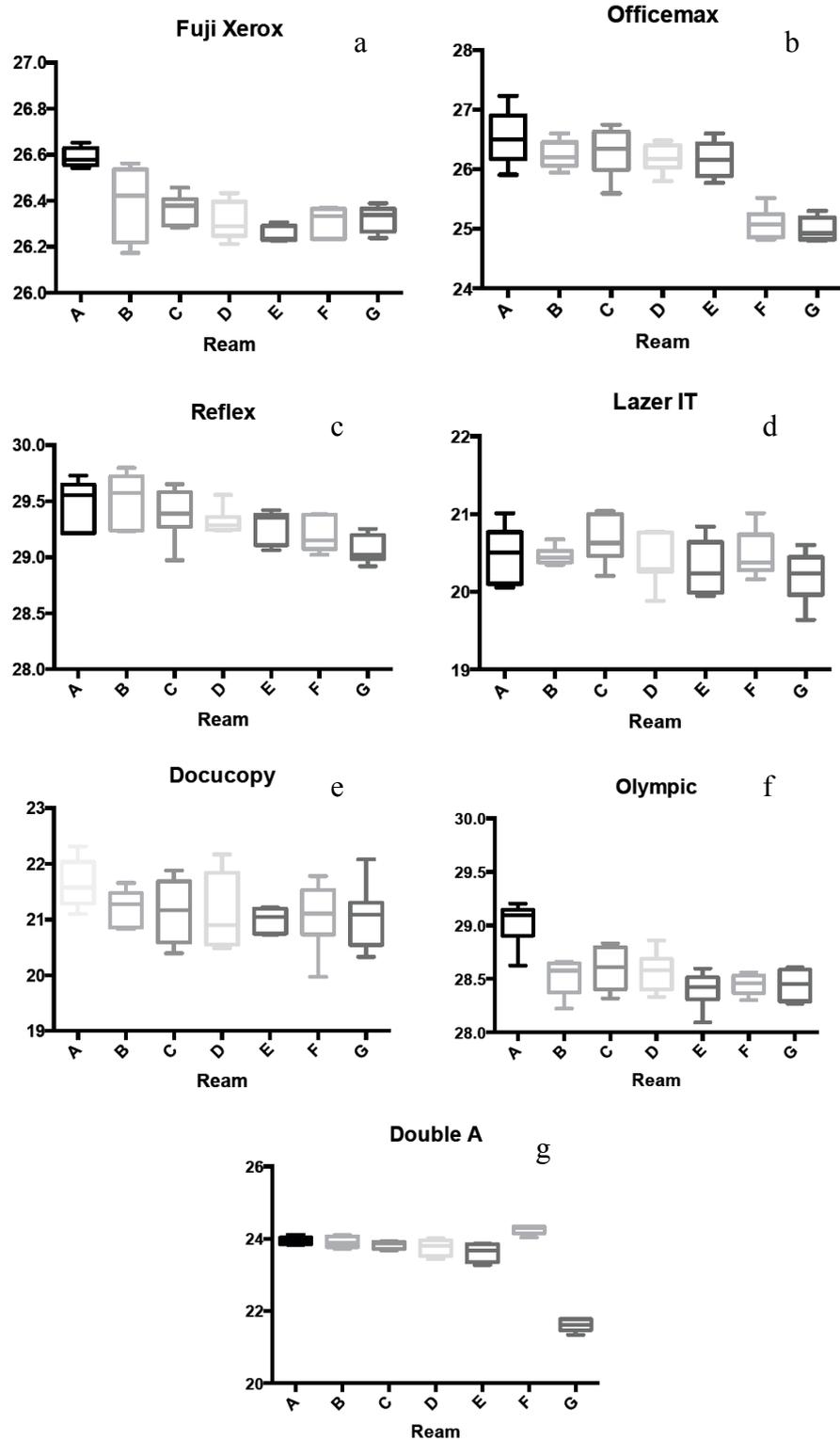


Figure 6.9 a-g: Boxplots denoting the mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) and standard deviation (‰) for seven reams from seven different brands

The results for oxygen are very similar to the carbon results. The Reflex, Lazer IT, Docucopy and Olympic reams would not be discriminated from one another utilizing either a 95% confidence interval or a benchmark value. As these reams were all packed on the same day/time, this implies that measurement and hence discrimination of paper reams is at the batch level rather than the singular ream level.

Officemax and Double A showed clear differences in the F and G reams, which are consistent with the carbon results. While both the Officemax F and G reams would be discriminated using a 95% confidence interval and either of the 0.7 and 1 ‰ benchmark values, only the G ream in the Double A brand would be easily discriminated.

The F ream for this brand would be discriminated by a 95% confidence interval of the measured samples but would not be discriminated with the more conservative use of a benchmark value. The difference between the F ream and the other sample groups is approximately 0.3 to 0.4 ‰, which was seen within the natural variation range for a number of ream samples in the single ream homogeneity experiment described in the previous section. Therefore, if this were a casework situation, we would not be able to discount the proposition that those two groups of samples had come from the same batch of paper.

Statistically, Dunn's post-hoc test showed differences in a number of brands including the Fuji Xerox (A vs. E and F), Reflex (B vs. F and G, C vs. G) and Olympic (A vs. E, F and G) brands. Though statistically significant, the magnitude of the difference is small, and given the small standard deviation values it is inappropriate to discriminate given the inhomogeneity of paper. The Lazer IT and Docucopy brands did not have any statistically significant different reams. This is likely due to the increased heterogeneity recorded for these reams, making differences more difficult to register given the larger standard deviation values.

6.3.4. Discrimination of Background Population

Two benchmark values are being proposed for use for discrimination of paper– a value of 0.7 ‰ for samples that have been run within the same analytical sequence and a

value of 1 ‰ for samples that are being compared between analytical sequences to account for potential instrument shift.

Using these two benchmark values, the 123 background samples were compared to determine the effectiveness of discrimination using oxygen isotopic abundance values. 7504 pairwise comparisons were made, from which 6135 and 5658 pairs of papers could be discriminated from each other utilizing the 0.7 and 1 ‰ benchmark values respectively. The discrimination power is 82% for 0.7 ‰ and 75% for a 1 ‰ benchmark value. Both of these discrimination powers are larger than the carbon discrimination power of 68%. A combined discrimination using carbon and oxygen isotopic abundance is assessed in Chapter 10.

6.4. Conclusions

123 paper samples were measured for their oxygen isotopic abundance values. Correlations were observed between the source of the ream and the $\delta^{18}\text{O}$ values measured, however there was more overlap observed than in the carbon results.

The variability and hence potential for discrimination of the oxygen isotope ratios of paper was found to be larger than the carbon results and this is not unexpected given the wider range that the oxygen values were observed to sit within. The maximum standard deviation for over 50 samples taken from a single ream was observed to be 0.54 ‰. This figure is likely exaggerated however due to differences introduced by using two different analytical sequences to measure the samples rather than one. For the same ream, the standard deviation from measuring 25 pages was observed to be 0.21 ‰.

Based on these observations, two benchmark values are proposed for use – 0.7 ‰ for papers that have been measured within the same analytical sequence and 1.0 ‰ for between sequence comparisons.

These benchmark values were observed to be robust within the inter-ream homogeneity experiments. Utilised to cross compare the 123 background samples, a discrimination power of 82% for 0.7 ‰ and 75% for a 1 ‰ was calculated.

Chapter 7

7. Oxygen Isotopes – Source, Handling and Forensic Sampling

7.1. Introduction

Focused on repeating the work conducted for carbon isotopes, this chapter aims to examine the external influences that may impact on the use of oxygen isotopes for the discrimination of document papers. While a benchmark/discrimination value has been experimentally defined for the comparison of the $\delta^{18}\text{O}$ values of papers, there are a number of other factors that require consideration during interpretation. Examining these factors experimentally, allows the consideration required for interpretation to be identified, quantified and recorded for future casework reporting.

The experiments conducted in this section mirror those undertaken Chapter 4, including an examination of the effect of the production and printing processes on oxygen isotopic abundance values. In addition, a short series of experiments was undertaken to determine the effect of the environment and forensic testing on paper $\delta^{18}\text{O}$ values.

7.2. Materials and Methods

The instrumentation, equipment and method used to measure the oxygen isotopic abundance values in this chapter are detailed in Chapter 5. Additional detail regarding sample collection and preparation is included in the material and methods section of Chapter 4.

Statistical tests were undertaken using Microsoft Excel® and GraphPad Prism version 5. When assessing the statistical significance of differences observed between groups of values, the Kruskal-Wallis Non-Parametric Analysis Of Variance (ANOVA) followed by a Dunn's Multiple Comparison Post-Hoc test was used. Otherwise, the

experimentally defined benchmark value determined in Chapter 5 (0.7 ‰) was used to assess whether values could be discriminated.

7.2.1. Source and Paper Production Samples

As for the carbon experiments, the following samples were collected and measured for their $\delta^{18}\text{O}$ abundance values from different stages of the production process:

- Whole eucalyptus wood chip (Wood Chip)
- Post-digester unbleached eucalyptus pulp (Unbleach Euc)
- Samples from within the bleaching/whitening process
 - Post oxygen wash (Ex O)
 - Post ozone and peroxide wash (Ex ZD)
 - Post caustic soda, oxygen and peroxide wash (Ex EOP)
 - Post chlorine dioxide wash (Ex D)
- Refined bleached eucalyptus (Refined Bleach Euc)
- Refined bleached pine
- Directly from the paper decal roll, post paper production (Paper Decal)
- Paper packed into reams for shipping (Paper Packed)
- Bulk measurements for packed paper and paper decal
- Paper filler material – Calcium Carbonate (CaCO_3)

Cellulose extracted and bulk samples were measured by the University of California at Berkeley with additional bulk samples measured at the AFP laboratory. Measurements made in both laboratories were referenced against international standards and as such, are comparable to demonstrate the effect of the production process on sample values.

The acidification experiment was undertaken within the AFP laboratory and was repeated separately to the carbon experiment. While there was sample remaining after the carbon experiments were undertaken, over 18 months had elapsed before the oxygen samples were analysed. Given the time difference, it is feasible that trace residues of acid could have further degraded the paper cellulose over that time, skewing the abundance values measured. A small change to the experimental method was made from the carbon experiments with the samples washed then dried on Whatman filter paper (Sigma Aldrich, Australia) without the use of a vacuum. This

sped up the acidification process and was not observed to change the reaction or ease with which the samples were washed to remove the residual acid. Fewer samples were measured than for carbon, with the number of replicate samples reduced. The samples selected and measured are shown in Table 7.3.

7.2.2. Usage of Paper – Printing Processes

250 μ g of sample was taken from a selection of the office printed samples utilised in Chapter 4 with a 2mm micropunch (Harris Unicore, LabSciTech, Australia). Repetition of brand and model of printed was removed in these experiments, based on the results obtained in the carbon experiments. This reduced the number of samples from 46 to 22 for the toner printed samples and from 13 to 8 for the inkjet printed samples. The samples measured for their oxygen isotopic content are shown in Table 14.9 (toner printers) and Table 14.11 (inkjet printers). For ease of comparison, the sample numbers were retained from the Chapter 4 experiments.

Triplicate samples were taken from the non-image areas of the printed pages (i.e. from the blank sections) and from directly on top of heavily inked areas that had been printed using a bold font. These samples were compared to the plain unprinted paper results measured in the background population study to determine whether there was any effect on the oxygen isotopic values of the papers from the printing process. The printed samples measured will also be used to give a preliminary indication of whether IRMS may be a suitable technique for discrimination of the toner or inks.

7.2.3. Usage of Papers – Environmental Effects and Forensic Testing

The range of tests measured for this section have been refined and targeted based on the results obtained in Chapter 4. For the tests that were removed from this section, further work may be required should these situations be presented in casework. The samples measured were retained from the carbon experiments and stored in plastic sheet protectors in the laboratory until measurement.

The following tests were undertaken:

- **Dirt/Dust**

The sheets were soiled by placing them on an outside concrete surface and applying pressure to maximise contact and adhesion of contaminants. Loose particles were dusted from the page prior to sampling

- **Fire**

The sheets of paper were charred by placing them above an open flame until mild discolouration occurred

- **Handling/Body Contact – Heavy**

Multiple fingers were held and rubbed on the sheets of paper for 5 minutes. The contact was very visible on the pages as body oils accumulated. Samples were taken directly from the contact area.

For the forensic testing, samples were taken from sheets that had been subjected to fingerprint treatment using 1,2-Indandione-Zn and DNA collection via tapelifting. A more detailed description of these techniques is included in Chapter 4.

7.3. Results and Discussion

7.3.1. Source and Paper Production Samples

The results from the cellulose extracted paper process samples are detailed in Table 7.1. This table also details the bulk measurements collected for each of the processing stages. The main additives used in the process and their measured $\delta^{18}\text{O}$ values are detailed in Table 7.2.

Production Step	Step in Process	Extracted Cellulose Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰)	Bulk Sample Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰)
Raw Wood Chip		29.96 ± 0.46	25.28 ± 0.51
Unbleached eucalyptus	After pulping	31.04 ± 0.35	28.60 ± 0.48
Ex O	First whitening step	31.33 ± 0.12	29.98 ± 0.10
Ex ZD	Second whitening step	31.02 ± 0.30	28.93 ± 0.20
Ex EOP	Third whitening step	30.99 ± 0.24	29.81 ± 0.86
Ex D	Last whitening step	31.43 ± 0.13	30.26 ± 0.08
Refined Bleached Euc	Last step before mixing with other pulp and additives	32.54 ± 0.21	30.34 ± 0.11
Paper Decal	After paper production	30.91 ± 0.08	28.77 ± 0.10
Packed Paper	After cutting and packing	30.97 (single measurement only)	28.59 ± 0.07

Table 7.1: Cellulose extracted and bulk measured $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) ± 1 standard deviation for paper source and production samples

Additive Name	Extracted Cellulose Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰)	Bulk Sample Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰)
Dry Bleached Pine	30.99 ± 0.20	30.51 ± 0.00
Refined Bleached Pine	31.11 (single measure only)	30.38 ± 0.04
Calcium Carbonate Filler	N/A	22.92 ± 0.10

Table 7.2: Cellulose extracted and bulk measured $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) ± 1 standard deviation, for additives used in paper production

A number of bulk measurement results obtained from the University of California Berkeley were used to assess the difference in values to the bulk measurement results obtained in the AFP laboratory. This duplication allows for an assessment of comparability and accuracy between the two laboratories. Any differences observed would indicate instrumental or correction differences between the two laboratories, which are expected, and would be used to inform the comparison between the extracted and bulk values. The bulk measurements made at the University of

California for the formed paper and packed paper were 29.41 ‰ and 29.31 ‰ respectively. The same measurements undertaken at the AFP laboratory were 0.64 ‰ and 0.73 ‰ different. This offset in values needs to be considered during interpretation of Figure 7.2 below.

Examining the bulk measurements in Table 7.1, a fractionation in the material or a significant decrease in extraneous material is inferred with an enrichment of 3 ‰ observed between the values for whole wood chip and the unbleached eucalyptus (i.e. between raw wood chip and pulp). While an enrichment of values is observed in the cellulose extracted samples, the magnitude is much smaller. This indicates that there has been a significant removal of non-cellulosic contaminants during pulping but that the pulping step had only had a moderate fractionation effect (1 ‰) on the cellulose itself. From here, through the whitening and production process to the packed paper, the $\delta^{18}\text{O}$ abundance values remain relatively stable for both the cellulose extracted and the bulk measurements. There is a 0.5 ‰ shift in the values at the known mixing stage, when pine fibres and other additives, such as the isotopically lighter calcium carbonate filler material, are added to the eucalyptus base fibres. This is consistent with the carbon results at the same stage in the production process.

Figure 7.1 is a plot of the $\delta^{18}\text{O}$ values of the extracted cellulose from each stage in the production process. Figure 7.2 contains the same data, plotted against bulk measurements for the same samples. For both plots, error bars were omitted, as they were negligible at this scale.

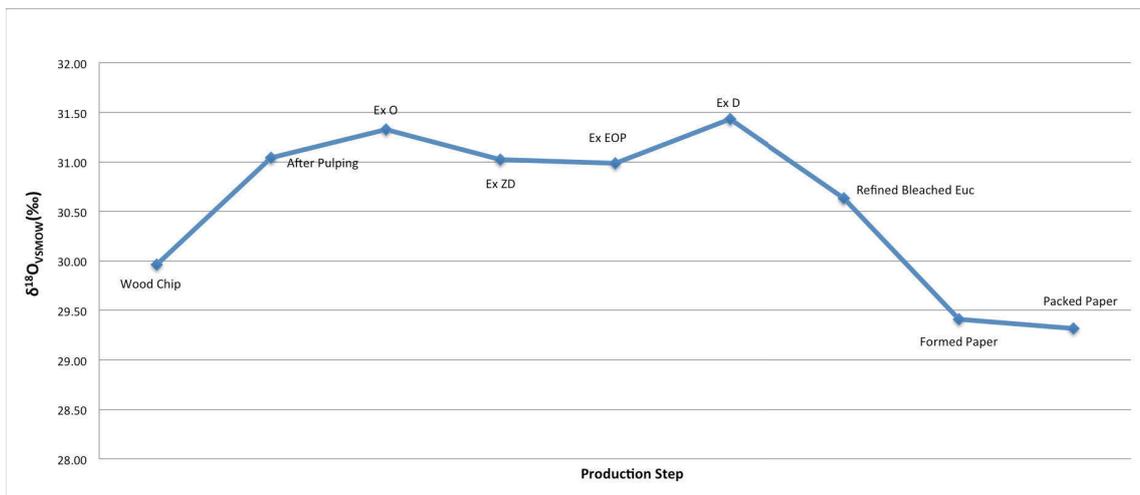


Figure 7.1: Mean $\delta^{18}\text{O}$ values of cellulose extracted from samples collected throughout the paper production process

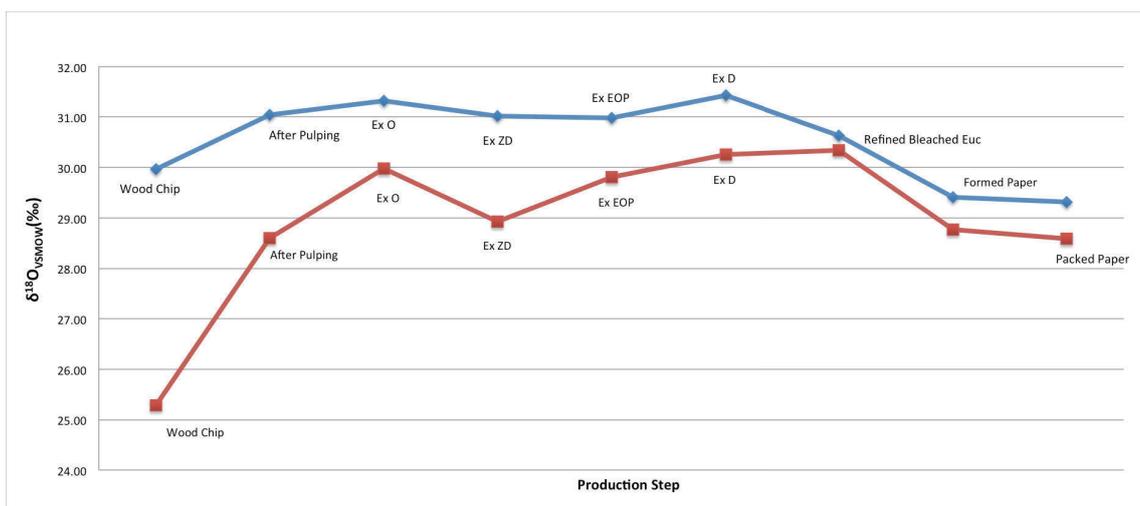


Figure 7.2: Comparison of mean $\delta^{18}\text{O}$ values of extracted cellulose (diamond markers) with bulk values (square markers) collected from throughout the paper production process

The plotted cellulose extracted samples in Figure 7.1 show more clearly the first fractionation event due to the pulping process, with a difference between the raw wood chip and the unbleached eucalyptus. This is more pronounced in the bulk samples, indicating that there is significant loss of non-cellulosic material that is isotopically heavier than the cellulose in the wood chips. There is also a noticeable effect in the bulk sample in the later stages of the process with the bleached eucalyptus

samples and the paper decal being offset by over 1 ‰ compared to the refinement stages. This can be attributed to the contribution of calcium carbonate, which as a bulk material is significantly lighter at 22.92 ‰.

A shift in values is observed after the addition of refined bleached eucalyptus for the cellulose extracted samples, with the value becoming 1 ‰ enriched relative to the Ex D sample. This shift is not retained during the next stage however with the paper decal value dropping below the other refinement stage results to 30.91 ‰. It is unclear from this data whether this shift is due to fractionation within the production process or due to an erroneous data point. The bulk samples do not reflect this shift, with the refined bleached eucalyptus sample staying consistent with the results observed at other stages of refinement. The difference between the refined bleached eucalyptus sample and the paper decal sample is observed in both the cellulose extracted and the bulk paper sample data.

The changes throughout the production process are similar for both oxygen and carbon isotopes - a difference is observed in the pulping stage, followed by relatively consistent results throughout the refinement and whitening stages. A subsequent difference is observed between the mixing and actual paper production stage. The difference between the refined eucalyptus and the formed paper is larger than in the carbon results however and this is likely due to the calcium carbonate filler having a larger effect given the higher ratio of oxygen to carbon in the calcium carbonate compound (3:1).

Overall, these results indicate that it is not possible to directly compare the results of bulk $\delta^{18}\text{O}$ measurement with the source cellulose and hence a cellulose isoscape. The overall difference between raw cellulose and formed paper was observed to be approximately 3.5 ‰, which is expected to vary between brands depending on the type, amount and isotopic value of the filler material used.

To determine whether hydrochloric acid acidification could be used as a simple method to measure the source cellulose value, acidification was undertaken on a number of papers using concentrated hydrochloric acid. The measurement results of the bulk versus acidified samples are shown in Table 7.3. Figure 7.3 shows the

difference between the two measurements. Australian papers showed a difference that was approximately equivalent to the production process results in Table 7.1.

Sample Number	Paper Country of Origin	Bulk Paper Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ St Dev} (\text{‰})$	After Acidification Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ St Dev} (\text{‰})$	Difference (‰)
43	Australia	28.23 \pm 0.18	30.71 \pm 0.16	2.47
66	Australia	27.76 \pm 0.17	30.70 \pm 0.06	2.93
75	Australia	27.99 \pm 0.19	30.73 \pm 0.07	2.74
77	Australia	27.96 \pm 0.09	30.83 \pm 0.07	2.87
79	Australia	29.36 \pm 0.31	30.50 \pm 0.17	1.14
82	Australia	28.19 \pm 0.21	30.60 \pm 0.14	2.41
83	Indonesia	19.46 \pm 0.24	25.27 \pm 0.08	5.81
88	Indonesia	20.57 \pm 0.09	25.22 \pm 0.04	4.65
89	Indonesia	20.18 \pm 0.00	25.27 \pm 0.04	5.10
96	Indonesia	20.68 \pm 0.88	25.42 \pm 0.15	4.74
103	Thailand	23.41 \pm 0.12	26.69 \pm 0.12	3.28
109	Thailand	20.46 \pm 0.08	26.69 \pm 0.11	6.23
111	China	24.18 \pm 0.35	27.39 \pm 0.07	3.21
121	China	20.87 \pm 0.21	25.22 \pm 0.18	4.35

Table 7.3: Paper $\delta^{18}\text{O}_{\text{VSMOW}}$ values before and after acidification

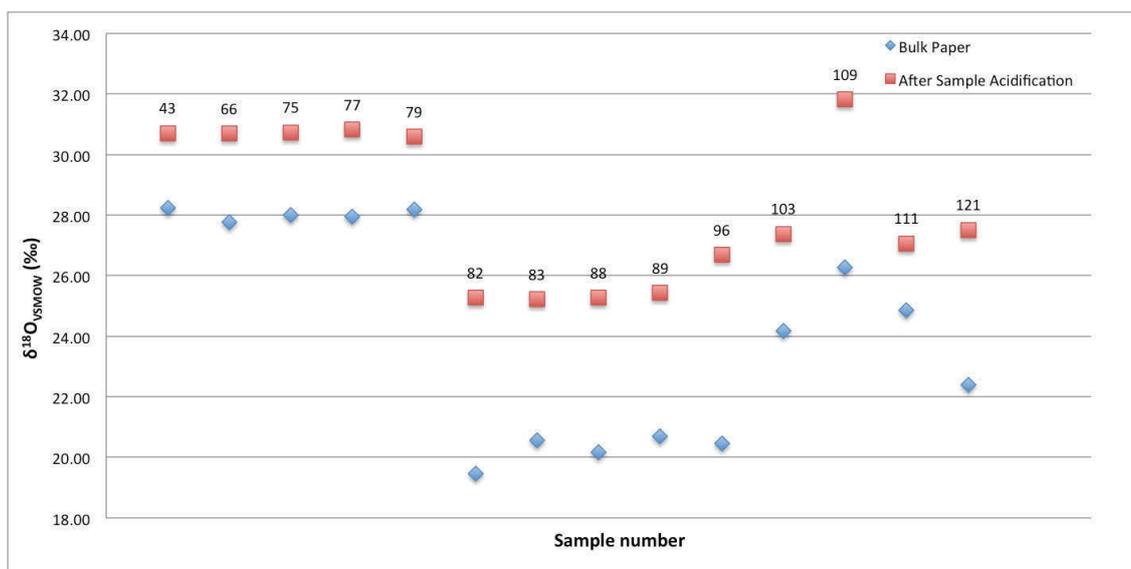


Figure 7.3: Mean $\delta^{18}\text{O}$ values of bulk paper (diamond markers) compared with the same paper samples after undergoing acidification (square markers).

Figure 7.3 shows that all samples had some effect from the acidification process on the $\delta^{18}\text{O}$ values measured. From these results, it appears that the magnitude of the difference observed after acidification would depend on the type, amount and isotopic value of filler material used. The only papers with a ‘known’ filler type are the Australian made papers (calcium carbonate), and these react as expected, becoming ~ 3 ‰ enriched when compared to the bulk paper values. For other filler types, further investigation using different analytical techniques is needed to determine the type and quantity of filler material present. Consideration also needs to be given within the examination protocols and interpretation structures proposed, given that it is evident that IRMS is being used to examine materials that can no longer be considered to be chemically identical.

While these and the carbon results show that there is potential for the use of acidification for the further discrimination of samples, additional experiments are required to show that the analytical method is consistently removing the same material in the same quantities (i.e. that the acidification process is repeatable) before being used for casework. At this stage, the acidification method seems only applicable for papers that utilise calcium carbonate type filler. Use of an additional analytical method prior to IRMS analysis could also be argued, to identify and discriminate samples based on the filler content. One suitable technique is X-Ray Diffraction (XRD), which can be used to identify crystalline substances. A pilot study of this technique is presented in Chapter 9.

Overall, these results re-iterate that the paper production process contains both fractionation and mixing stages. Additionally, and more strongly than was seen in the carbon results, a direct comparison between the bulk paper value and a database or predicted raw cellulose value (e.g. through cellulose isoscapes) cannot be made for oxygen isotopes. A material specific isoscape containing measured values would need to be created for comparison to unknown values in future, if provenancing work was to be undertaken.

7.3.2. Usage of Papers – Printing Processes

Table 14.10 and Table 14.14 detail the blank paper, printed paper (non-image area) and printed paper (image/text) $\delta^{18}\text{O}_{\text{VSMOW}}$ values for the 22 toner and 8 inkjet samples measured. The mean values in these tables are taken from triplicate measurements for each condition. **Error! Reference source not found.** and Figure 7.5 show plots for the data from each table, with the error bars representing the 0.7 ‰ benchmark/discrimination value utilised for discrimination of samples, as defined in Chapter 6.

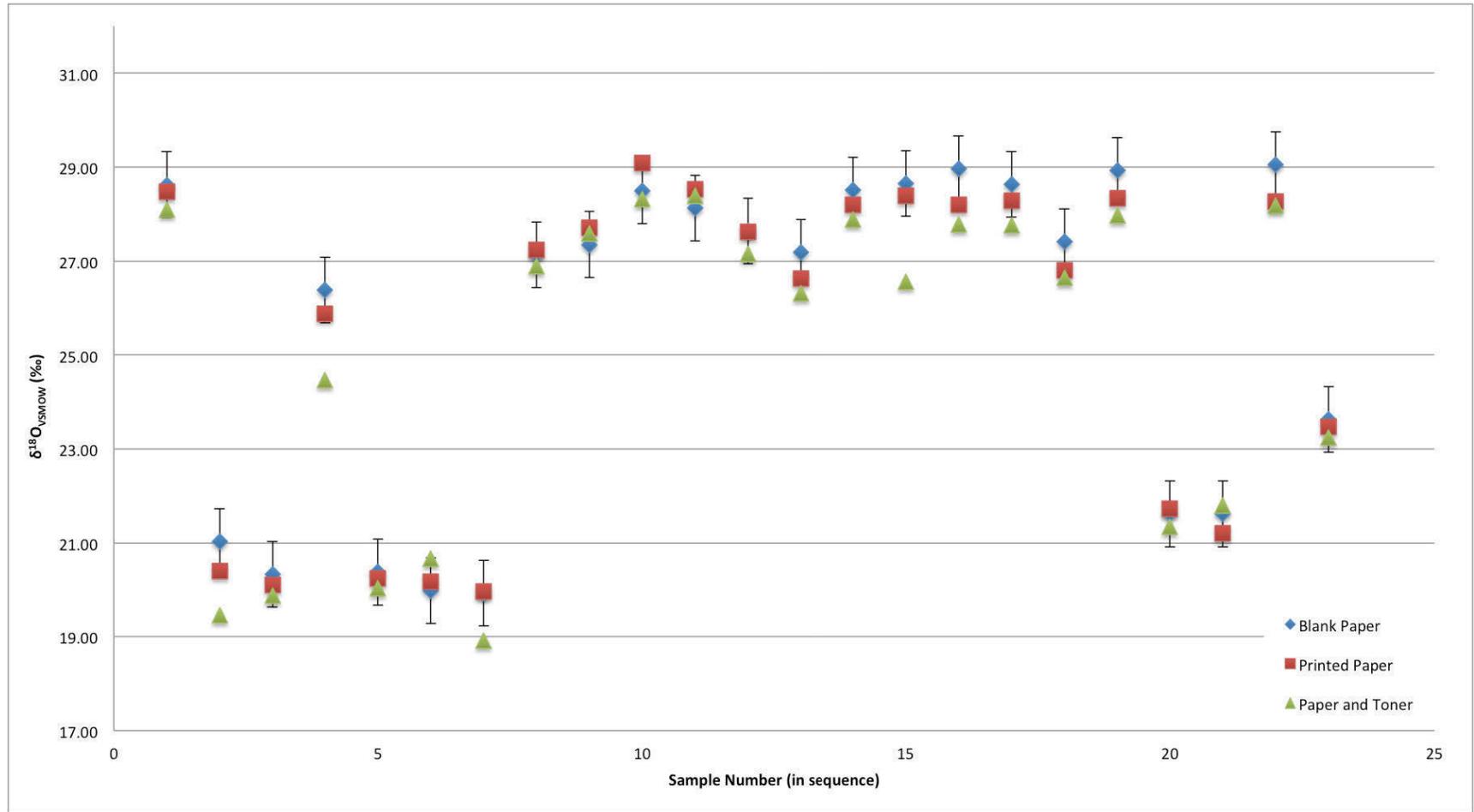


Figure 7.4: Comparison of the mean $\delta^{18}\text{O}_{\text{VSMOW}}$ values of papers before and after printing using toner

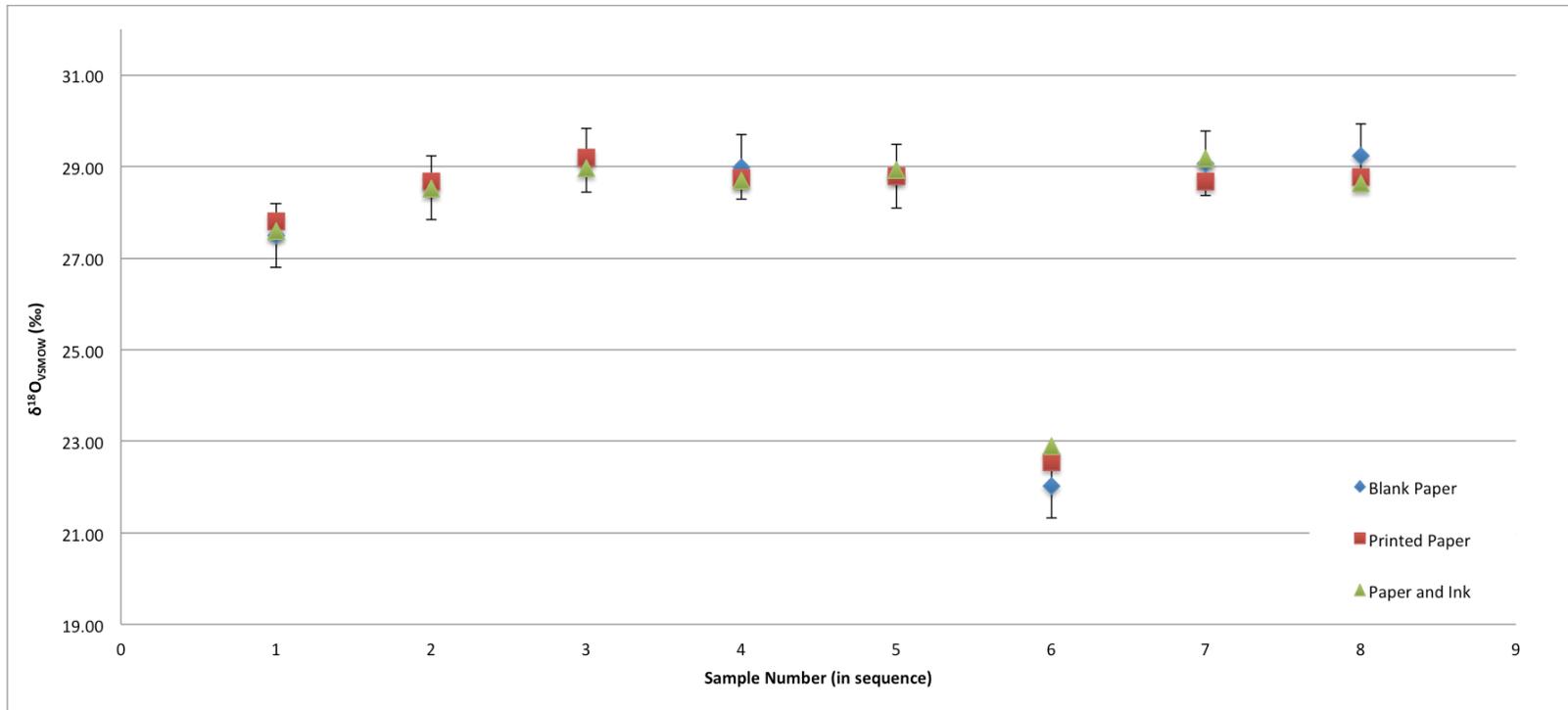


Figure 7.5: Comparison of the $\delta^{18}\text{O}_{\text{VSMOW}}$ values of papers before and after printing using inkjet ink

The oxygen values closely mirror the results seen in the carbon measurements for this experiment. For the toner printed samples, as shown in Figure 7.4, the inclusion of toner has an effect on the majority of the $\delta^{18}\text{O}_{\text{VSMOW}}$ values measured. The effect of the inclusion of toner printing with the paper samples however was not observed to produce a uniform change in the isotopic abundance values measured. While the majority of the oxygen results showed depletion in the oxygen isotopic abundance values, enrichment was observed in 3 samples demonstrating a potential for discrimination that may be higher than observed the carbon results. This requires additional work to confirm however, including the development of a technique for removal of toner from paper.

There was some difference between the blank paper and the printed paper results and this is due to either to the inclusion of extraneous toner particles or the variability of the paper itself. None of the printed paper samples would have been discriminated using the 0.7 ‰ discrimination value.

For the inkjet samples, there was a little to no effect from the printing process, even when the samples were taken from directly on top of the ink. This is due to the fluid nature of the ink and hence the more dilute dispersal of the colourant (dye or pigment) present. These results support the carbon values measured, which showed that no additional sample handling is required for inkjet printed documents.

7.3.3. Usage of Papers – Environmental Effects and Forensic Testing

The papers used and their untreated $\delta^{18}\text{O}_{\text{VSMOW}}$ values are shown in Table 7.4. These values were pooled from the three sheets measured from each brand. The measured values for the environmental samples are shown in Table 7.5. The mean and standard deviation values for the untreated papers in this table were taken from the relevant single sheets.

Paper Brand	Paper Designator	Country of Origin	Untreated Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) \pm 1 Std Dev for 3 sheets sampled
Double A	A	Thailand	23.74 \pm 0.24
Reflex	B	Australia	28.08 \pm 0.57
Olympic	C	Australia	28.11 \pm 0.48
Lazer IT	D	Indonesia	20.58 \pm 0.27
Fuji Xerox	E	China	25.50 \pm 0.31

Table 7.4: Brands of ream used and pooled mean values for environmental and forensic test evaluation

Treatment	Paper Designator	Untreated Paper Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) \pm 1 Std Dev	Treated Paper Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) \pm 1 Std Dev	Difference of Means (‰)
Dirt/Dust	AC	23.84 \pm 0.26	23.51 \pm 0.29	0.33
	BA	28.13 \pm 0.38	28.85 \pm 0.21	0.72
	CA	28.34 \pm 0.16	28.93 \pm 0.23	0.59
	DA	20.62 \pm 0.09	21.25 \pm 0.11	0.63
	EA	25.66 \pm 0.22	26.03 \pm 0.27	0.37
Fire	AB	23.79 \pm 0.14	24.23 \pm 0.36	0.44
	BB	28.52 \pm 0.00	28.55 \pm 0.10	0.03
	AC	28.34 \pm 0.16	28.20 \pm 0.78	0.14
	DB	20.55 \pm 0.26	21.23 \pm 0.32	0.68
	EA	25.66 \pm 0.22	26.64 \pm 0.20	0.98
Handling - High	AA	23.66 \pm 0.34	24.59 \pm 0.28	0.93
	BB	28.52 \pm 0.00	28.90 \pm 0.11	0.38
	CC	27.67 \pm 0.64	29.03 \pm 0.28	1.36
	DC	20.55 \pm 0.47	21.51 \pm 0.07	0.96
	EC	25.34 \pm 0.44	26.53 \pm 0.14	1.19

Table 7.5: $\delta^{18}\text{O}_{\text{VSMOW}}$ measurement values for environmental testing experiment

Both the accumulation of dirt/dust and heavy contact was observed to have an effect on the measured values. The amount of material transferred was inconsistent between test samples, which may have contributed to the variability of the results (and the higher standard deviation values) and this shows that handling of this type needs to be considered during sampling and subsequent interpretation of casework results. Charring of the paper samples via the application of flame also had an effect on the results and the inconsistency seen in the values (i.e. increase in standard deviations) is due to different levels of charring being sampled throughout the replicates.

Table 7.6 contains the results for the samples that were treated for fingerprint detection using 1,2-Indandione-Zn and collection of DNA via tape lifting.

Treatment	Sample Designator	Untreated Paper Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) \pm 1 Std Dev	Treated Paper Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) \pm 1 Std Dev	Difference of Means (‰)
Indandione	AB	23.79 \pm 0.14	24.40 \pm 0.16	0.61
	BC	27.73 \pm 0.78	28.02 \pm 0.53	0.29
	CB	28.33 \pm 0.25	28.37 \pm 0.55	0.04
	DB	20.55 \pm 0.26	21.37 \pm 0.42	0.82
	EB	25.50 \pm 0.28	26.35 \pm 0.33	0.85
Tape lifting	AB	23.79 \pm 0.14	24.36 \pm 0.30	0.57
	BC	27.73 \pm 0.78	28.45 \pm 0.93	0.72
	CB	28.33 \pm 0.25	28.52 \pm 0.34	0.19
	DC	20.55 \pm 0.26	20.47 \pm 0.48	0.08
	EB	25.50 \pm 0.28	26.03 \pm 0.29	0.53

Table 7.6: $\delta^{18}\text{O}_{\text{VSMOW}}$ measurement values for forensic testing experiment

Significant differences were observed between the untreated paper and some of the papers tested for both the indandione and the tape lifting procedures. The differences would have discriminated the samples in three of the ten comparisons made, highlighting that the use of oxygen isotopes for comparison after these treatments would require careful interpretation.

7.4. Conclusions

The source materials and effect of the production process was examined on the $\delta^{18}\text{O}_{\text{VSMOW}}$ abundance values of office papers. Similar to the carbon results, both fractionation and mixing were observed to have occurred. This means that the bulk-measured values for document papers are not directly comparable to predicted/modelled tree cellulose isoscapes.

Acidification was observed to have an effect on the values measured, particularly for Australian papers that are known to contain calcium carbonate fillers. Additional work is required utilising other analytical techniques to further explore these results however and in response, a pilot study using X-Ray Diffraction (XRD) was conducted to identify the actual filler types present in a limited sample of papers. Those results are included in Chapter 9.

As for the carbon experiments, inclusion of toner with the paper samples was observed to have an effect on the $\delta^{18}\text{O}_{\text{VSMOW}}$ values measured and so careful sampling is required for documents printed with these types of processes. Inkjet printing did not have an effect even when sampled from directly on top of the printing. Different to the carbon results and of interest, the environmental experiments showed larger differences between clean and treated paper. This is particularly of concern for the forensically treated samples, whose values were shown to change after testing. This will need to be considered prior to analysis or during interpretation depending on the stage that the examination is undertaken.

Chapter 8

8. Hydrogen Isotopes – Preliminary Study

8.1. Introduction

For forensic purposes, studies utilising stable isotopes generally include a combination of three isotopes (and the addition of other techniques) to maximise discrimination. While homogeneity assessments of the extent presented in this work have not been cited in the literature, the use of multiple isotopes is generally described as both necessary and positive for discrimination of samples.

The use of hydrogen isotopes is challenging due to hydrogen exchange between the surface of molecules containing proteins or other exchangeable compounds, particularly those containing $-\text{COOH}$, $-\text{NH}_2$, $-\text{OH}$ or $-\text{SH}$ functional groups (Carter, 2011) with atmospheric moisture. In cellulose, this becomes an issue through the exchange of hydrogen from the $-\text{OH}$ groups which are exposed on the sides of glucose monomers within the polymer chain. In cellulose fibres and hence paper sheets, we can expect a proportion of these $-\text{OH}$ groups to be cross-linked to other cellulose polymers via hydrogen bonding, reducing the number of groups available for exchange. In raw cotton fibres, the molar exchange fraction of H in cellulose was calculated to be 0.05, with experiments confirming the fraction to be 0.046 ± 0.009 (Meier-Augenstein et al., 2014). While exchange experiments have not been undertaken directly on document papers (and will not be conducted here), we can expect a similar fraction to be available for exchange in document paper cellulose.

A number of studies have been undertaken, particularly for the measurement of hair, to try to counteract or quantify the effects of ‘exchangeable hydrogen’ (Bowen et al., 2005, Coplen and Qi, 2012, Meier-Augenstein et al., 2011). Although each method has demonstrated repeatability within a laboratory, a number of the variables known to affect exchangeable hydrogen including sample size and handling are still under investigation across a range of materials (Chesson, 2009).

Despite the publication of a number of methods, there is yet to be consensus or a standard methodology adopted between laboratories. This is most clearly demonstrated in the results of the Forensic IRMS (FIRMS) Network proficiency test scheme (<http://www.forensic-isotopes.org/pts-results.html>). Performance monitoring across laboratories and through the measurement of a number of different materials has demonstrated the variability of $\delta^2\text{H}$ values obtained between laboratories. Given the same materials measured for $\delta^{13}\text{C}$, $\delta^{18}\text{O}$ or $\delta^{15}\text{N}$ produced comparable values, the difference for $\delta^2\text{H}$ can be attributed to variability in sample handling methodology prior to IRMS. Taking one material as an example, hemp was measured by 11 laboratories with a robust standard deviation of $\delta^2\text{H}_{\text{VSMOW}}$ 10.663 ‰ and a range of 42 ‰ reported. Both of these values are well outside of the expected variability both for the material and the instrument, demonstrating that differences in protocols have affected the measured results obtained.

Given that inter-laboratory comparability is yet to be achieved, a decision was made at the start and again during the course of this project to conduct a select amount of $\delta^2\text{H}$ work here as a scoping study only. This would focus not on producing internationally comparable results but instead on ensuring that the methodology and thus results are fit for purpose in providing an indication of the utility of hydrogen isotopes for the comparison of document paper. In the event that the preliminary hydrogen results obtained are positive, collaboration with the international community would continue so that future results are comparable outside of the AFP laboratory.

This decision should also be viewed in the context of the broader aims of this research – to have a validated method for the casework comparison of document papers. Given the ongoing development and discussion internationally on the control of exchangeable hydrogen, this could not have been assured if hydrogen was included as a core component of the study. Validation of a suitable technique for control of hydrogen exchange has been flagged as future work required within the AFP laboratory.

This chapter outlines a short study of a limited selection of 25 papers from the background population, to examine whether the addition of hydrogen isotopes may assist in producing additional discrimination to the $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ values previously

measured. From the outset of these experiments, it was known that these results would be of limited value: as the laboratory methodology for the measurement of the $\delta^2\text{H}$ of organic samples has not yet been fully validated; the sample set was limited; and the inter/intra brand variability of document papers had not been defined. The experiment was conducted however for the purpose of informing future work on $\delta^2\text{H}$ for the comparison of document papers, and to provide some indication of whether hydrogen isotopes can be used for discrimination when carbon and oxygen isotopes have been unsuccessful.

8.2. Materials and Methods

8.2.1. Standards and Samples

International standards NBS-22 (Oil) and IAEA-CH-7 (Polyethylene) were used to correct the unknown values. A polyethylene glycol (PEG) standard material provided by another forensic IRMS laboratory was used as a quality assurance material. Further in-house calibration of this material is required for future work to assure reproducibility over time however the measured values obtained in these analytical sequences did agree with the value provided ($\delta^2\text{H}_{\text{VSMOW}} -34.4 \text{ ‰}$).

25 paper samples were selected for measurement from the 125-paper background sample set measured in the carbon and oxygen studies. These papers were selected as samples that had low or no discrimination with other similar papers using carbon or oxygen isotopes and if discrimination was called, the difference was only using a single isotope and was close to the defined benchmark/discrimination value. The samples selected are shown in Table 14.3, with sample numbers retained from previous tables to maintain comparability between sets of results.

As each of the paper samples had been stored in the AFP laboratory for over 12 months prior to analysis, and the analyses were conducted at one point in time across two analytical sequences run on consecutive days, homogeneity of exchangeable hydrogen was assumed. As discussed in the introduction, comparability internationally and between the AFP and other laboratories for these results is not being claimed.

Given the principle of identical treatment, it can be assumed that any magnitude of the differences between values would not change significantly should a protocol for exchange of hydrogen be undertaken in future. Hence, what is important to note in these experiments is not the actual value of the results obtained (although still traceable to the VSMOW international referencing scale) but the differences observed between them.

Triplicate samples were taken for each paper (1 sample each from 3 pages) and after being punched and placed into silver capsules, the samples were dried for 4 hours in a 60°C oven to remove absorbed water. The capsules were then crimped and placed in a helium-purged autosampler for analysis. Again, noting that the sample replicates are lower than previously recommended, this level of sampling was deemed to be acceptable for the purposes here.

8.2.2. Instrumentation and Equipment

The instrumentation, equipment and method used to measure and correct the hydrogen isotopic abundance values are similar to those utilised for the oxygen isotopes in Chapter 5.

The following differences in the instrumental set-up were made to adjust for the measurement of hydrogen isotopes:

- TC/EA furnace temperature set to 1450°C
- GC oven temperature set to 20°C
- Reference peak placement at 10, 50, 200 and 240 seconds (for 20 second duration) with a total run length of 420 seconds, reference gas pulses set at approx. 1.5 psi to scale the peak at 2500mV
- Peak elution occurred at approx. 90 seconds with a width of approx. 30 seconds.

A typical chromatogram for a paper sample is shown in Figure 8.1. A long sample run time was selected to ensure that the oxygen in the sample had eluted and passed through the detector prior to the next sample dropping. If not enough time was allowed between samples, this isobaric interference was shown to overlap with the mass 2

sample peak of the next sample in the sequence, particularly in samples that contained a high proportion of oxygen, as is the case for paper samples.

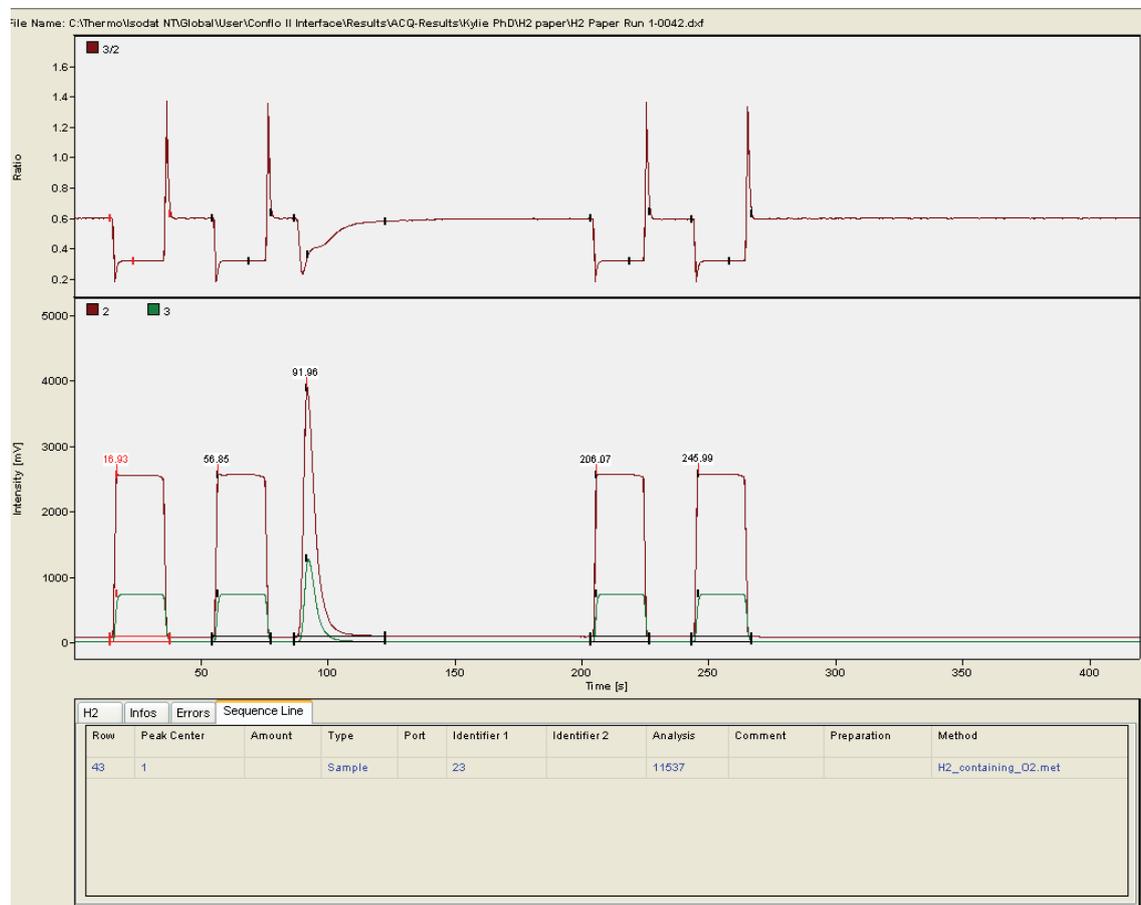


Figure 8.1: Typical chromatogram for H₂ measurement of document papers.

A Genius ME5 (Sartorius, Goettingen, Germany) analytical balance was used to weigh the international standards NBS-22 (oil) and IAEA-CH-7 (polyethylene), both with published values on the VSMOW scale, and the laboratory quality control material PEG. The sample sizes used to equate to a peak height of approx. 3500mV were 100 μ g for NBS-22 and 250 μ g for IAEA-CH-7 and PEG. Each sample was weighed to the target weight \pm 10 μ g. Paper samples were prepared using a combination of a 1.2mm and 2mm Harris Uni-core micro-punches (Proscitech, Queensland, Australia), to a sample size of 250 μ g. Samples were placed into 3.3 x 4 mm silver capsules for solids (Thermo Fisher Scientific, Sydney, Australia) and stored in a Perspex desiccator with self-indicating silica gel prior to oven drying and measurement.

8.3. Results and Discussion

8.3.1. Discrimination within the sample set

The 25 papers measured were selected based on their lack of discrimination using carbon and oxygen isotopes. The hydrogen isotopic abundance values were measured within the range $\delta^2\text{H}_{\text{VSMOW}}$ -11 ‰ to -46 ‰. This range is wider than expected, though still relatively narrow on the hydrogen scale. The sample values are plotted below in Figure 8.2 and Figure 8.3, with Figure 8.3 plotting paper values by region of origin. Details of the samples including brand are included in Table 14.13.

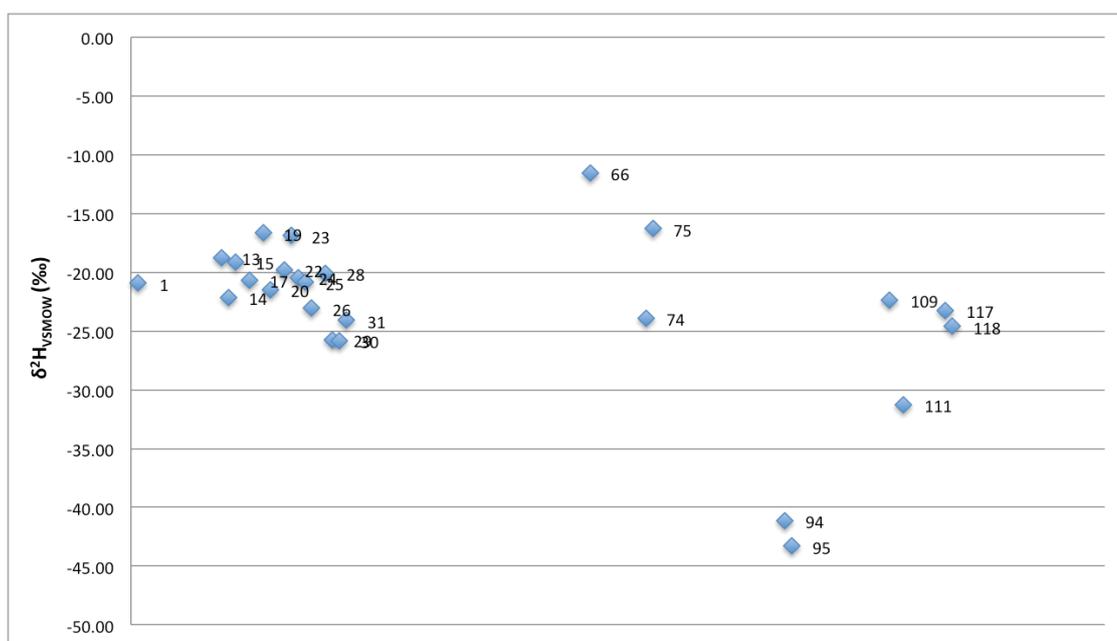


Figure 8.2: Mean $\delta^2\text{H}_{\text{VSMOW}}$ (‰) values of 25 papers collected from Australia and New Zealand

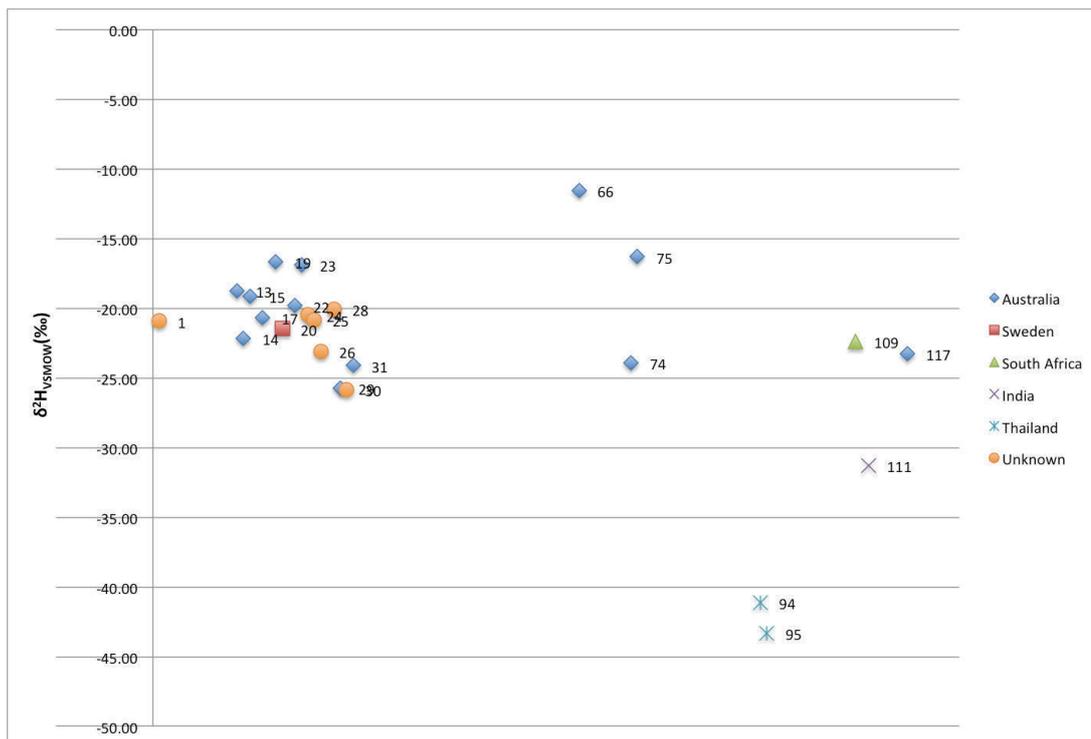


Figure 8.3: Mean $\delta^2\text{H}_{\text{VSMOW}}$ (‰) values of 25 papers plotted by region of origin

The results for the 25 paper samples show that even though there is a range of 35 ‰ for the values, the majority of the values are in a much tighter band between -16 and -26 ‰. As seen in Figure 8.3, regionally there is not as much of an opportunity to determine whether region has an effect due to the limited sample size and provenance of the sample group. From this sample set however, the samples produced in Thailand were measured within the -40 to -45 ‰ range while all other samples are intermingled.

The interpretation of region effects from this sample set is also limited in that it contains a number of papers of unknown origin. These samples were sent from document examination laboratories within Australia and New Zealand as part of the usage (printing) study and hence origin information (ream packaging) may not have been available during collection.

As there is no homogeneity study to accompany this data for discrimination, an estimate of a suitable benchmark/discrimination value is difficult to determine for this limited sample set. A range of benchmark values were calculated for use under these limited circumstances and pair-wise comparisons undertaken for these 25 papers. The

range of discrimination powers for the 300 comparison pairs is seen in Table 8.1. The comparison table is included electronically within Appendix 5.

Potential Benchmark Value	Discrimination Power
5 ‰	0.44
10 ‰	0.22
15 ‰	0.15
20 ‰	0.10

Table 8.1: Discrimination powers obtained from pair comparisons of 25 samples measured for $\delta^2\text{H}_{\text{VSMOW}}$ values

From these values and the standard deviations in Table 14.13, and taking into account that these papers were selected due to being difficult to discriminate with the carbon and oxygen isotopes, a reasonable estimate for a benchmark value would be 15 ‰. Based on the standard deviation values, the use of 15 ‰ as a benchmark value would be a conservative estimate, as it is larger than the largest standard deviation measured whilst still allowing reasonable discrimination of the sample set.

Upon closer inspection however, the majority of the discriminated pairs for the 15 ‰ benchmark/discrimination value can be attributed to discriminating the two samples from Thailand (samples 94 and 95) that were measured in the -45 ‰ region. If these two samples are taken out of the sample set, the discrimination power for the 15 ‰ benchmark value is reduced to 0.003, or only one sample pair discriminated from 300 comparison pairs. The two Thailand samples themselves are not discriminated from each other using $\delta^2\text{H}$ values. Comparing this to the carbon and oxygen results, this pair of papers was discriminated from each other and from a large number of the other papers using the $\delta^{13}\text{C}$ values. The $\delta^{18}\text{O}$ measurements were observed to discriminate each of these papers to some others, but not in discriminating each other demonstrating that the $\delta^2\text{H}$ comparison would have been of assistance during interpretation.

Overall, the results from these 25 papers show that the hydrogen values are of limited value in providing discrimination when carbon and oxygen isotopes have previously

been used. The use of hydrogen measurements may assist in providing a stronger basis for discrimination however, particularly when a comparison pair has only been discriminated using a single isotope.

8.4. Conclusion

In the context presented here, noting the small and selective sample set measured and the limitations of the method used to measure the values, it appears that adding hydrogen isotopes to a paper examination protocol that already includes carbon and oxygen is unlikely to create additional unique discrimination. Discrimination was observed between some previously discriminated samples however, indicating that a wider sample set would lead to higher discrimination factors than observed here.

These results indicate that hydrogen isotopic values may assist in strengthening opinions, particularly where differentiation is difficult to call due to the difference in values sitting close to the benchmark discrimination values or due to discrimination only being observed using a single isotope.

An extensive amount of work would be required prior to use in casework however including method validation, sample handling, homogeneity and an examination of the effect of usage on the $\delta^2\text{H}$ values. The protocol validated would also need to account for exchangeable hydrogen and be shown to produce comparable results between laboratories through an inter-laboratory comparison exercise or through the FIRMS proficiency-testing scheme.

Chapter 9

9. Other Paper Examination Techniques

9.1. Introduction

There are a number of examination techniques that are more easily accessible, require minimal training and are more cost effective than IRMS for the comparison of document papers. These techniques have been used in document examination casework in the past, though not routinely. For laboratories that choose to offer paper examination comparisons, these are generally the techniques that are used, which have been adapted from quality assurance laboratories established within paper manufacturing facilities. It is not clear from the literature however, how discriminating and therefore useful these techniques actually are.

The most basic examinations involve the examination and comparison of the physical properties of paper. Initially, this would include comparing two papers using ultraviolet (UV) light, transmitted light and white (ambient) light. UV light at a range of wavelengths (254, 312 and 365nm) is used to examine the presence and intensity of optical brighteners included in the paper furnish. These optical brighteners are also visible, in combination with the dyes present in the paper, when viewed and compared using ambient light. Transmitted light is used to examine the paper fibre arrangement and opacity (density) of the sheet.

Other examinations include the measurement of the physical size of the sheet and the calculation of sheet grammage by weighing the sheet and calculating the grams per m² (gsm). Using the grammage, the apparent sheet density (kg/m³) can be calculated after measurement of the thickness of the sheet using calibrated callipers or a weighted micrometre (Murphy, 2009). The term apparent sheet density is used to denote the measured/calculated sheet density of a single sheet. Overall, the benefit of these types of physical examinations is that they are non-destructive to the document and can be

performed using equipment that is more readily accessible than analytical chemistry techniques.

Thin Layer Chromatography (TLC) can be used to separate the components of a mixture by testing their affinities between a stationary and mobile phase. The stationary phase used is a silica material that is coated onto an aluminium backing. The mobile phase (eluent) will vary depending on the type of sample being run but for inks and paper it is generally a mixture of ethanol, ethyl acetate and water in a ratio of 2:1:1. While commonly used for writing inks such as ballpoint inks (LaPorte et al., 2006, Roux et al., 1999, Trejos et al., 2010, Wilson et al., 2004) and printing inks (Bell et al., 2013, Poon et al., 2005), TLC also has the potential to differentiate papers by comparing the optical brightener and dye composition extracted.

The addition of other analytical techniques should also be considered, where they measure or characterise specific components of the paper of interest. While a number of techniques are able to achieve this, such as Laser Ablation Inductively Coupled Plasma Mass Spectrometry (LA-ICP-MS), ideally the techniques should also be relatively simple for sampling, have only limited destruction of the sample and a high discrimination. In addition, any technique selected should have gone through an extensive validation process to inform interpretation.

As IRMS analysis generally targets and measures the cellulosic content (through gross mass) with a smaller contribution from filler material, X-Ray Diffraction (XRD) analysis of the papers was selected as an additional technique to identify the type of filler present in paper samples. This technique has been used previously, with some success (Foner and Adan, 1983) though has not been used extensively. The use of XRD will also assist in providing context to the acidification results included in Chapters 5 and 8.

The purpose of this chapter is to test and define the use of a number of paper examination techniques. Following this, a wider examination protocol that is complimentary to and includes IRMS examinations can then be formed.

For discrimination and interpretation, the ability to use multiple techniques adds weight or significance to the results of difference or similarity obtained. In addition, having the ability to utilise alternative or non-destructive methods in an examination protocol allows for greater flexibility to be built into a protocol, to allow for different casework questions or scenarios where preliminary results may be required at short notice.

9.2. Materials and Methods

9.2.1. Standards and Samples

As the examination of paper is a comparative process and a number of the techniques used in this chapter are not quantitative, the papers measured in this chapter were compared against sample number 71. This paper was chosen as being representative of one of the most common brands of high quality Australian made paper (Reflex).

The paper samples were selected from the background population measured in the isotopic studies in Chapters 3 and 5 and the sample reference numbers have been retained from those studies. In total, 60 paper samples from the background population were measured using the techniques in this chapter. This is excluding TLC, of which 46 out of the 60-sample set were measured for comparison. The brand and collection information for these paper samples is detailed in Table 14.3.

9.2.2. Comparison using UV, Transmitted and Ambient Light

A Foster and Freeman Video Spectral Comparator 6000 High Resolution (VSC6000HR) instrument (Worcestershire, UK) was used to provide a range of light sources including:

- UV light at 256nm, 312nm and 365nm.
- Transmitted white light
- Ambient white light

Each paper sample was placed next to the sample 71 page and examined under the three types of light with the room darkened and the instrument sides closed to

eliminate interference. Where differences were observed, the results were photographed using the inbuilt camera.

9.2.3. Measurement of Grammage and Apparent Sheet Density

A Vibra AJ 420CE balance (Wedderburn Scales, Sydney, Australia) was used to weigh the paper samples. The balance was zeroed with two paper clips on the pan, and each sheet rolled and secured with the paper clips prior to measurement to prevent the sheet folding over the sides of the pan and resting on the bench during measurement. The balance was covered by a Perspex windshield that was closed while collecting the weight values. To measure the thickness of the sheet, each paper sample was placed under the foot of a model D0011 Digital Micrometer (IDM Instruments, Victoria, Australia) that had been calibrated to measure paper thickness.

Three sheets from each of the brands in Table 14.14 were weighed three times and their thickness measured in ten separate locations using the calibrated micrometer. For each sheet, grammage was calculated using the average mass (g) and the standard A4 sheet size (27.9 x 21.5cm). After measurement of the sheet thickness, the grammage was converted to kg per sheet by dividing by 1000, then multiplied with the average sheet thickness (in μm) to determine the apparent sheet density (kg/m^3). Ideally, ten sheets would be used to more accurately determine the sheet density for a given ream. As this is unrealistic for casework situations, three sheets were measured and averaged for each ream as an intermediate between the ideal (ten sheets) and the worst-case reality (one sheet, likely to be common in forensic casework).

As outlined in other chapters, Australian Standards (1998) stipulate that papers for comparison should be measured after adjustment of the absorbed water within the paper sheets. The natural hygroscopicity for papers varies based on where the paper was produced however this can be adjusted by storing the paper in a low humidity environment before increasing to a higher humidity to equilibrate samples prior to measurement. For these tests this adjustment was not conducted as the papers measured had been held in the same laboratory conditions for at least 12 months prior to measurement, minimizing the effects that would have been seen if the papers were recently sourced from different locations/retailers.

As there is no described method for discrimination of samples using grammage and sheet density, additional work was required in an attempt to define a sample discrimination protocol. To determine the homogeneity of a single brand, seven pages from six reams were measured. Additionally, the same seven brands and reams measured as part of the between ream homogeneity experiments conducted in Chapters 3 and 5 were measured in duplicate (i.e. two sheets from each ream were measured). While ideally a more extensive set of homogeneity experiments would have been undertaken, the experiments described here were assessed as being fit for purpose and were able to inform discrimination.

As the results of the grammage and apparent sheet density tests are quantitative, comparison was undertaken at two levels – each sample was compared to paper sample 71 and pair-wise comparisons of the entire background population, where each sample measured was cross compared and discriminated based on the benchmark values defined.

9.2.4. Thin Layer Chromatography

A 1mm micropunch (ProSciTech, Sydney, Australia) was used to take four punches from the paper samples in Table 11.3. The samples were placed into Micro-Hemacrit Capillary Tubes with a 1.1-1.2mm internal diameter (Kimble Chase Scientific, ProSciTech, Sydney, Australia) that had been sealed at one end using an open flame, and covered with approximately 3.5 μ L of undiluted pyridine (Sigma Aldrich, Australia). The top of the tube was then sealed with an open flame and the sample left to extract until use.

Sample tubes were broken and the eluent spotted on to Merck KGaA Silica Gel 60F₂₅₄ Aluminum backed thin layer chromatography plates (Rowe Scientific, Sydney, Australia) using a 0.63mm Perkin Elmer Syringe (product number N6101390, Rowe Scientific, Sydney, Australia) with at least 1cm spacing between samples. The plates were placed in an oven to dry at 60°C for 15 minutes then placed in a glass tank containing 1.5cm of mobile phase of ethyl acetate, ethanol and water in a ratio of 2:1:1. The plate was placed into the tank with the lid on and left to run until the mobile phase reached approximately 1cm from the top of the plate. Once removed, the plate

was left on a bench to dry and was imaged using UV light at 254nm using the Video Spectral Comparator (VSC6000HR).

Visual comparison in combination with the calculation of retention factors was used to discriminate samples. To calculate the retention factor (R_f) - defined as the distance travelled by the separated peak divided by the distance travelled by the mobile phase, the line tool in Adobe Photoshop CC (trial version 14.0 x 64) was used to measure the distance of the peaks and the final position of the mobile phase. Care was taken to measure the centre point of the sample start position to the middle of the eluted peak.

A difference was confirmed based on peak placement and/or the absence of well-defined additional peaks in either the sample or the standard (paper sample 71). For some plates, edge effects were observed that affected the first or last sample run on the plate. This altered the placement of samples and had to be taken into account during discrimination. An example of this is shown in Figure 9.1.

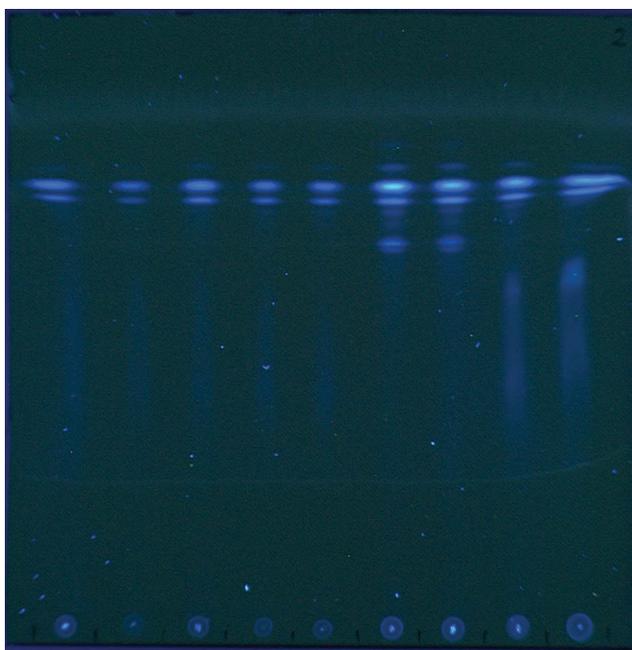


Figure 9.1: Image of developed TLC plate, showing edge effect on the sample run in the far right position.

9.2.5. X-Ray Diffraction Analysis

An XRD unit was not available in the AFP laboratory and as such, external assistance was sought for the measurement of a selection of paper samples by XRD.

14 paper samples were selected from the 125 paper background sample set and given to Sietronics Pty Ltd (Mitchell, ACT, Australia) with a request to identify the filler types contained in the samples, using a semi-destructive technique only (i.e. with no sample digestion/chemical processing though the samples were able to be physically removed from the sheet). The 14 samples measured are shown in Table 9.3.

A Bruker-AXS D4 XRD (Bruker AXS Corporation, VIC, Australia) with copper radiation at 40kV and 30mA was used to identify the fillers. An approx. 2cm² sub-sample was taken from each page and adhered to a low background holder using a small amount of Bostick[®] Clear GluStick[™]. Both the low background holder and the GluStick were observed to not interfere with the XRD traces of the paper samples.

A range of 1.3 to 70°2θ with a 0.02 degree step and a 2 second per step count time was used during measurements. A graphite monochromator was used in the diffracted beam. Sample results were searched against the Bruker Diffrac^{plus} Search/Match software and the ICDD PDF-2 database (2006).

9.3. Results and discussion

9.3.1. Ambient, UV and Transmitted Light

60 background paper samples were compared to paper sample 71 using UV, transmitted and ambient light. Where differences were noted, an image was taken using the inbuilt camera of the VSC6000HR for review and a decision on whether to call a difference. The images of the papers discriminated are included in electronically in Appendix 5.

The discrimination of papers using these light techniques was found to be quite challenging. This was especially the case when considering the transmitted light

images, which can be used to observe the opacity of the sheet or the arrangement of fibres due to the paper making process. The primary challenge when considering these types of visual comparisons is ensuring that the comparison made is consistent and accurate, particularly as comparison complexity increases. An example of a simple and a more complex comparison using transmitted light is shown in Figure 9.2 and Figure 9.3.



Figure 9.2: Example of a simple discrimination task using transmitted light between sample 71 (left) and sample 94 (right).



Figure 9.3: Example of a more complex discrimination task using transmitted light between sample 71 (left) and sample 85 (right).

To observe consistency within a brand, images were taken from one page from seven reams using each of the lighting techniques. Undertaking these comparisons, it became apparent that identification of a difference was easier when more than one sheet (with the same opacity pattern) is available so that the examiner is able to gain an understanding of what is normal within a sample group. In casework, this implies that it would be easier to use these techniques where a question is being asked about the substitution of a page in a multiple page document rather than in one to one comparison situations. To overcome this, techniques such as the ones presented by Berger et al. (2009, 2012) could be used to reduce the subjective nature of the opinion.

Discrimination was also easier when more than one light examination showed a difference e.g. if there was a difference observed in the transmitted light images in addition to the UV light examinations. With respect to UV light examinations however, it has been shown to be inappropriate to call discrimination based solely on differences in UV reaction (i.e. a difference in fluorescence – either colour or intensity), due to the potential for inhomogeneity within a single ream or brand (Green, 2012). This can occur when papers have been cut and packaged in the same ream but may have been prepared in different manufacturing batches that may have been produced with different constituents.

As part of this project, the reams stored within the AFP laboratory were screened using UV and the observations of Green (2012) were confirmed with at least one ream identified that contained differences in the UV reaction of the sheets within the same package. Also of note, Brunelle and Reed (1984) describe that caution should be taken when comparing different intensities of UV fluorescence due to dulling of intensity as the paper ages or is exposed to other types of environmental weathering.

Ambient light examination was found to be the easiest of the light examination techniques to discriminate samples. In general, the magnitude of the difference was more readily apparent and discrimination was repeated using the other two light techniques. Caution should be used in the same way as for the UV light examinations however, if a visual difference was the only reaction observed (Green, 2012).

Due to the difficulties surrounding UV and ambient light, the results from all three light techniques were used in combination for discrimination. The results of these tests are shown in Table 9.1. The discrimination factor for these light techniques was calculated at 0.34 when each sample was compared against sample 71. Transmitted light was responsible for the majority of the discriminations made using light examination techniques.

9.3.2. Grammage and Apparent Sheet Density

The measurement of grammage and apparent sheet density originated in paper quality assurance mills. As a consequence there are no published values that can be used for forensic discrimination, as quality assurance mill laboratories are testing whether the paper produced is within specified tolerance ranges. The most applicable information was published by Murphy (2009), who discusses that in typical paper manufacturing plants a tolerance of 2% of the stated grammage is allowed for production (e.g. for 80gsm papers this would be 1.6gsm).

The mean grammage and apparent sheet densities measured and calculated for duplicate sheets in seven reams from seven brands are detailed in Table 14.14. As these were only duplicate measurements, the standard deviation values have been omitted. Included in this table are the mean grammage and apparent sheet densities for seven sheets measured from six reams from the Reflex branded reams used in the isotopic homogeneity testing previously undertaken. The grammage and apparent sheet density values were plotted for each ream in Figure 9.4. The variability associated with each brand is highlighted in Figure 9.5.

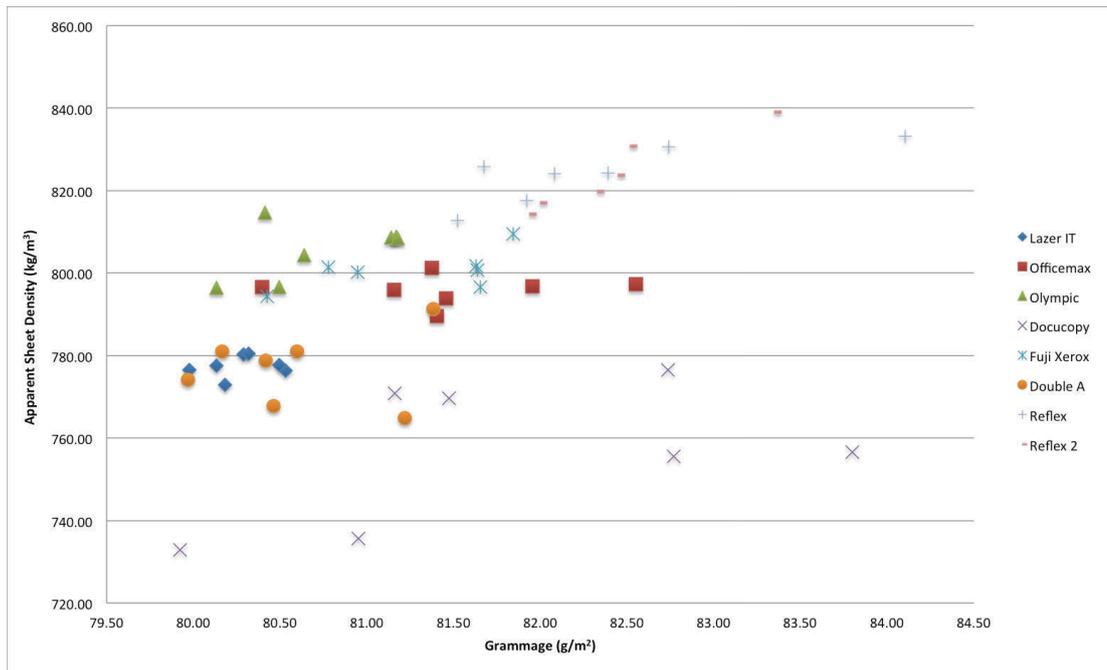


Figure 9.4: Grammage vs. Apparent Sheet Density for seven paper brands

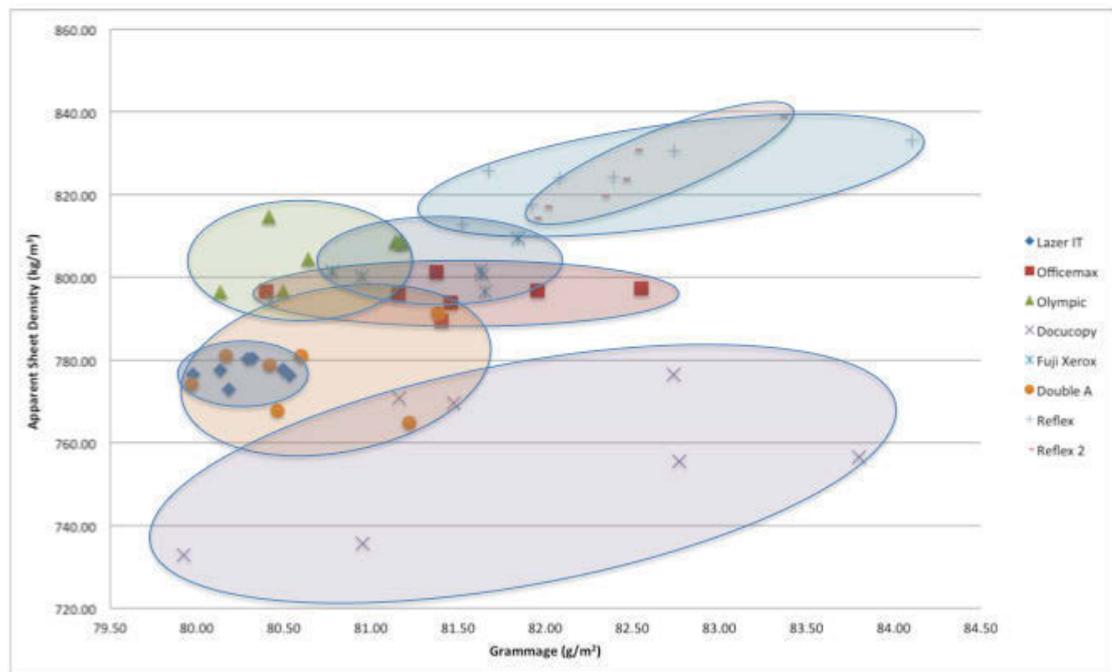


Figure 9.5: Grammage vs. Apparent Sheet Density for seven paper brands, with brands highlighted in coloured circles.

Figure 9.4 shows that there is potential for discrimination of paper samples using these techniques, if the values vary significantly enough. As the range observed for the grammage results is narrow (between 80 and 84gsm) the calculated grammage values may not be able to be used in isolation i.e. without further measurement of the thickness and calculation of the apparent sheet density. Additionally, Figure 9.5 shows that the grammage range expected within a single brand can vary naturally. An example of this is the Officemax brand that contains reams that vary from 80gsm to 82.5gsm. This observation is reinforced in Figure 9.6 and Figure 9.7, which shows the single page measurements for the six 'Reflex 2' reams plotted to show single ream variation.

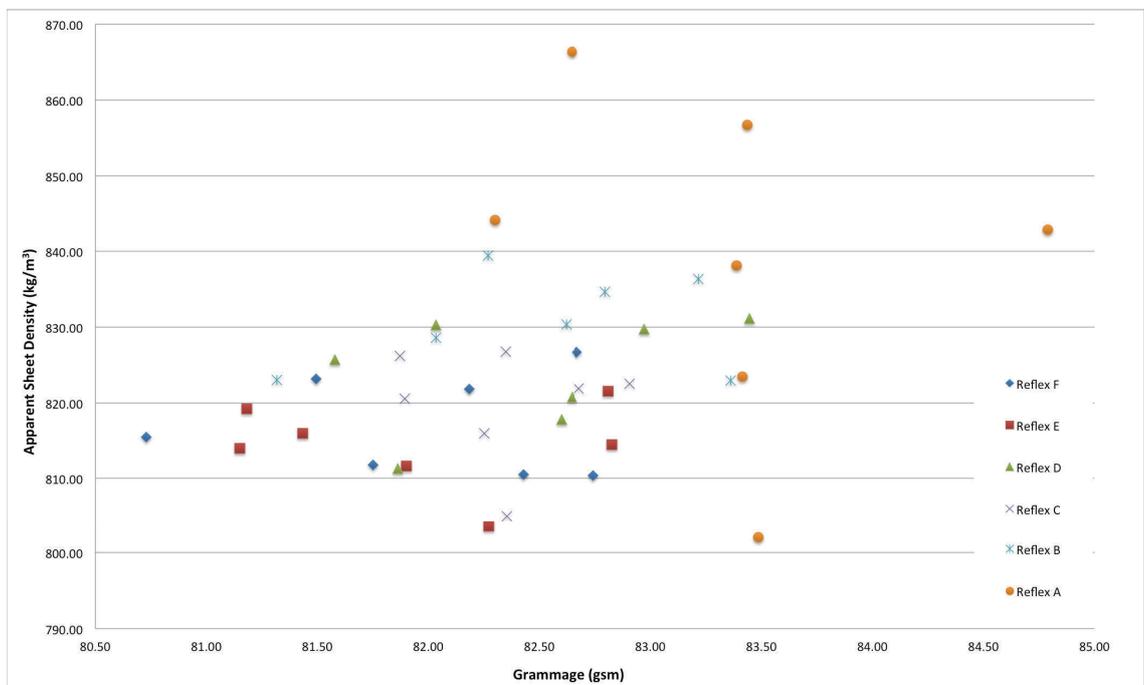


Figure 9.6: Grammage and Apparent Sheet Density for seven replicates of papers measured from six reams of the same brand (Reflex)

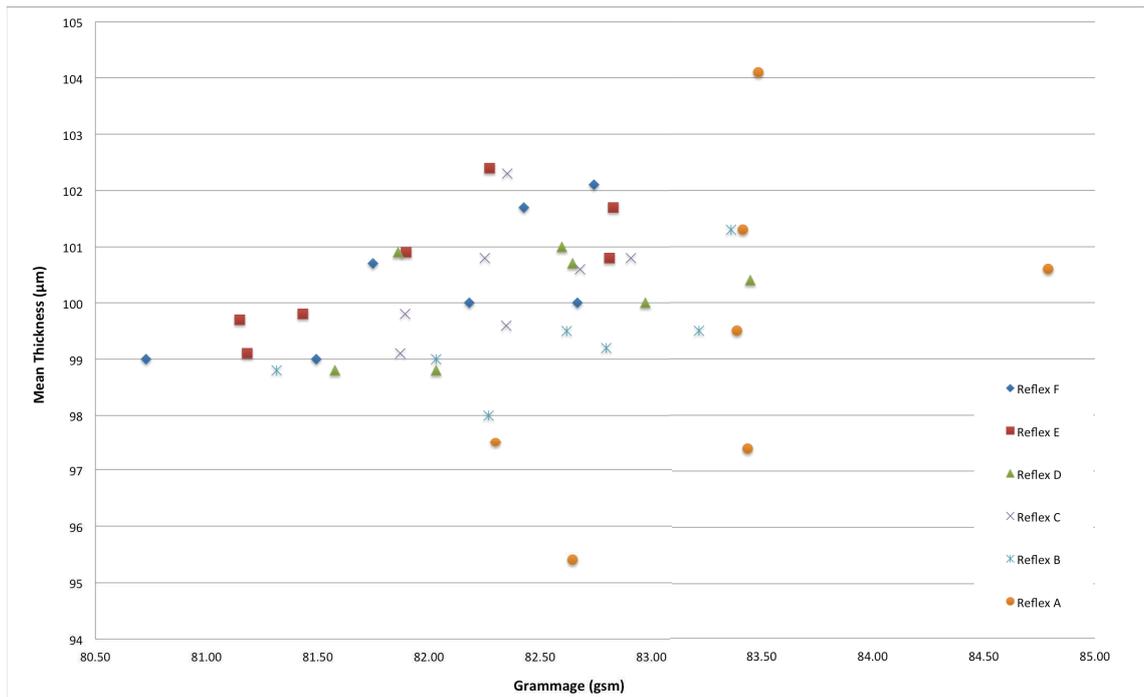


Figure 9.7: Grammage and mean measured thickness for seven replicates of papers measured from six reams of the same brand (Reflex)

The largest range in values for papers from the same ream is just over 2gsm. This does not factor in the wide spread of values seen in the Reflex A ream, which has a range closer to 2.5gsm. Comparing these results to the 2gsm discrimination value proposed, it would be more technically correct to use the 2.5gsm range as a discrimination benchmark as it gives a more conservative estimate of the homogeneity within a single ream, even if the Reflex A ream was designated as an outlier. While this will reduce the discrimination power of grammage as a technique, similarly to the IRMS results, it will avoid false discrimination of samples.

The 35kg/m^3 discrimination/benchmark value is observed to be robust for apparent sheet density when compared to these single ream results, which have a range value below 20kg/m^3 on average. Consideration of a lower benchmark value may be warranted based on these results and should be considered on a case by case basis during interpretation when other measurement results are being collated with apparent sheet density.

Comparing these results with those obtained in the isotopic studies, the same trends regarding inhomogeneity are repeated. In particular, the results show that some brands are more variable (e.g. Docucopy) than others (e.g. Reflex). As this data presents a preliminary indication of the homogeneity of a single ream, the data measured for the seven replicates from six reams (Reflex 2) study will be used to define a suitable benchmark value for discrimination.

The measurements obtained for the 60 background population samples are detailed in Table 14.15, which includes the difference between each sample and the sample 71 paper, with the cell shaded in green for papers that would be discriminated using the benchmark values proposed. A plot of the comparison between sample 71 and the rest of the population group is included as Figure 9.8, with the error bars on sample 71 representing the 2.5gsm and 35 kg/m³ values proposed for discrimination.

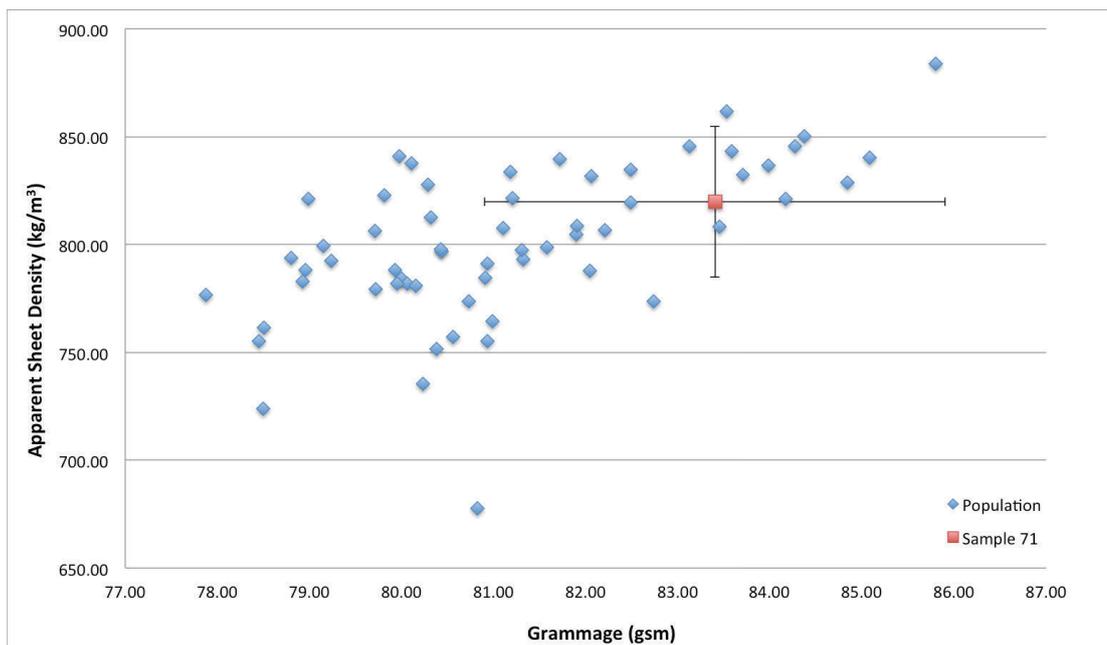


Figure 9.8: Comparison of grammage and apparent sheet density values between sample 71 and the selected background population.

The discrimination power when compared to the sample 71 paper was calculated to be 0.48 for grammage and 0.35 for apparent sheet density. Interestingly, only 15 comparisons were found to be different in both the grammage and apparent sheet density values. 34 papers in total were discriminated from the sample 71 paper. This is 57% of samples.

When the data set of 60 samples was cross compared, with a total of 2346 comparisons made, the discrimination powers were 0.33 and 0.46 for grammage and apparent sheet density respectively. The comparison tables for these techniques are included electronically within Appendix 5.

9.3.3. Thin Layer Chromatography

Table 14.16 presents the detailed results of the comparison of papers using TLC. It should again be highlighted that the comparison of results using this technique is limited to within plate comparisons and as such each ‘unknown’ paper was directly compared to the sample 71 paper. The discrimination summary is included in Table 9.1.

The images of the plates are included electronically in Appendix 5. There were some plates that were difficult to interpret due to the edge effects seen, which may have reduced the number of comparisons discriminated. While it was possible to re-run some of the samples affected, given the low discrimination power of the technique overall, a decision was made that this was not likely to significantly alter the results and the analyses were not repeated. The discrimination power for TLC was calculated to be 0.35.

9.3.4. Collation and Comparison of Techniques

Table 9.1 shows the results of discrimination of the 60 database samples when compared to the sample 71 paper, using the four different techniques. For the TLC results, 46 of the same 60 samples were measured. Where a sample was not available for cross comparison, it has been marked with an N/A.

Sample Number	Light Exams	Grammage	Apparent Sheet Density	TLC
43		x		
46		x		
47				
48		x		
49		x		
50				
53	x	x		N/A
54			x	x
56	x	x		
66		x	x	N/A
67				N/A
68				
69		x		x
72			x	
73				
74				
75				
76				
77			x	
79			x	
80				N/A
81				N/A
82	x			N/A
83		x	x	N/A
84			x	N/A
85				x
86		x		N/A
87				
88	x	x	x	
89		x	x	x
90		x		x
91		x	x	N/A
92	x			N/A
93	x	x		x
94	x	x		

Sample Number	Light Exams	Grammage	Apparent Sheet Density	TLC
95	x	x	x	
96	x	x	x	
97	x		x	
98	x		x	
99	x			N/A
101	x		x	
102	x	x	x	N/A
103	x			x
104		x		x
105	x	x	x	x
106		x	x	x
107	x	x	x	x
108	x	x	x	
109	x	x	x	N/A
112	x			
113	x			
114	x			x
115		x	x	x
117				
118				
120		x	x	
121		x		
122				x
124		x		x
125		x	x	
# Disc.	22	29	21	17
Disc. Power	0.37	0.48	0.35	0.33

Table 9.1: Discrimination of samples against paper sample 71, using a range of examination techniques, with the x denoting discrimination using that technique.

Further to the individual interpretation and comparison results presented in Table 9.1, a structure for interpretation of the results in a broader context needs to be defined. This is particularly the case for subjective measurements such as the light examination

and TLC results. A summary of the discrimination powers observed when one or more techniques showed discrimination is shown in Table 9.2.

Discrimination observed using:	Discrimination Power
One Technique Only	0.28
Two Techniques Only	0.28
Three or Four Techniques	0.16

Table 9.2: Discrimination of samples against paper sample 71

In casework circumstances, non-destructive techniques would be conducted first in preference to destructive techniques. Given the interpretation challenges with the light examination techniques and the known issues published with intra ream differences in UV light exams, a differentiation should not be called unless substantiated by another technique. This is particularly the case for grammage and apparent sheet density given the wide range of variation observed naturally.

Likewise, unless the difference is significant (e.g. the presence of a strongly fluorescent and well defined band absent from the comparison sample), discrimination using TLC should likely only be used when at least one non-destructive technique also shows a difference.

Using these conditions for comparison, the discrimination power for the non-destructive examination types (light examinations, grammage and apparent sheet density) was calculated to be 0.22. Including the TLC results, this discrimination power was extended to 0.25. As this is a limited sample comparison of 60 papers to one specimen paper (sample 71), discrimination within a wider population is expected to be higher than these values.

9.3.5. X-Ray Diffraction

Table 9.3 details the samples measured and the types of filler materials detected after measurement using XRD analysis. All samples were observed to contain cellulose and calcite (CaCO_3) and hence these compounds were not included in the table.

Sample Identifier	Brand	Country of Origin	Talc $\text{Mg}_3(\text{OH})_2\text{Si}_4\text{O}_{10}$	Dolomite $\text{CaMg}(\text{CO}_3)_2$	Aragonite CaCO_3	Quartz SiO_2
43	"ditto"	Australia	X			
66	Olympic	Australia	X			
75	Reflex	Australia				
77	Reflex	Australia	X			
79	Reflex	Australia	X			
82	Paper One	Indonesia				
83	Paper One	Indonesia		X		
88	Lazer IT	Indonesia		X		
89	Lazer IT	Indonesia	X			X
96	Double A	Thailand	X			
103	Yes Bronze	China	X	X		
109	Officemax	South Africa				
111	Bilt Matrix	India	X		X	
121	HP	Brazil	X			

Table 9.3: Filler content for 14 papers analysed by XRD

The XRD results, although only used on a limited sample size, showed that the technique would be a valuable addition to characterise and discriminate paper samples. If available, this analysis should be done prior to IRMS to ensure that papers with the same chemical composition are being compared. Surprisingly, there are also within brand differences that can be used for discrimination. This indicates either that the composition of papers is changing over time or that different manufacturing facilities that supply the same brand are utilising different source materials and paper process constituents.

As discussed in previous chapters, measuring bulk papers using IRMS is not ideal but the results can be trusted, as different filler types and amounts will lead to different isotopic abundance values. For those papers that are cannot be discriminated using the benchmark values, the papers can be reported as unable to be distinguished, or can undergo further processing using acidification to remove calcium carbonate filler material prior to re-measurement, to directly compare the cellulosic content of the sample.

9.4. Conclusions

A number of examination techniques, ranging from simple physical to chemical techniques, were tested on 60 background paper samples with varying results.

For physical examinations that are non-destructive, transmitted light examinations were observed to be the most discriminating however the results obtained were susceptible to subjective interpretation by the examiner. In contrast, the paper grammage and apparent sheet density results were less prone to examiner bias and allowed for quantitative values to be recorded and compared. These techniques required wider benchmark values to be defined however, due to the natural variations observed within reams. Using non-destructive techniques, a discrimination power of 0.22 was achieved for the sample population measured.

Chemical separation using TLC to compare the paper samples' optical brighteners and dyes was shown to be a complementary technique to the non-destructive techniques used. The addition of TLC resulted in an increase of the discrimination power to 0.25. Given the low overall addition to discrimination, this technique is viewed as complementary rather than additive and (as discussed in the following chapter) its use is at the examiners discretion.

XRD was shown to be a useful analytical technique to determine the filler content of a small sample of papers. This technique would be best suited prior to IRMS analysis to identify and compare the types of crystalline fillers used in the paper furnish. Further

work including validation, homogeneity studies and environmental effects will be required prior to use in forensic casework.

A comparison of these techniques to discrimination using IRMS in a broader casework examination protocol will be examined in the following chapter.

Chapter 10

10. Examination Protocol, Interpretation of Results and Reporting

10.1. Introduction

The initial purpose of this chapter is to examine the discrimination of paper samples when a number of techniques are combined. This will first focus on the results of the carbon and oxygen abundance values measured, including the combination of results from the non-destructive and chemical examination techniques discussed in Chapter 9. The primary aim is to define an examination protocol for casework comparison of papers.

Based on the examination protocol defined, a procedure will be proposed to ensure that the factors affecting the interpretation of results are clear. Although not used here, the factors discussed in the interpretation procedure could later be adapted into a Bayesian network or likelihood ratio calculations (Biedermann et al., 2012, Buckleton et al., 2006, Carter et al., 2014, Taroni et al., 2012) with additional experimentation to extend the conclusions of this work to use a numerically defined scale. The scale to be proposed here will be Bayesian in nature and will utilise two alternate propositions and express the findings in terms of “weights” in support of one of the propositions versus the other, but will not numerically define the opinion level to be provided through the use of a likelihood ratio.

As part of the discussion, Appendix 4 includes a proposed report appendix to explain the use of these examination techniques, including IRMS and isotopic discrimination, in the format commonly used for court reporting results of this nature. This technical appendix aims to summarize the techniques and methodology used for measurement in a way that lay persons and the court can understand.

The work in this and the preceding chapter ensures that IRMS is not being developed in isolation but is being included as a part of a wider examination protocol that

discriminates document papers via a combination of techniques. The chapter following this will test the proposed protocol in a range of casework scenarios, through the use of blind trials.

10.2. Combining Techniques

10.2.1. Using Carbon and Oxygen Isotopes for Discrimination

With discrimination powers of 68% and 82% for carbon and oxygen respectively a comparison was undertaken to determine if the results could be optimized (or increased) using a combination of two isotopes. Figure 10.1 shows the spread of values obtained for oxygen and carbon isotopes from the 123-sample background population. The error bars in this figure represent the 0.5 ‰ carbon and 0.7 ‰ oxygen benchmark discrimination values previously defined.

To examine any trends relating to the region of production, the samples have been plotted by production location in Figure 10.2 and Figure 10.3. The focus of this examination is to identify whether the effects observed in the individual plots, particularly with respect to the Australian samples, are consistent.

To determine the effectiveness of using two isotopes in combination, a comparison of the tables utilised in Chapters 3 and 5 was undertaken, with discriminated samples absent from the oxygen discrimination chart (but present in the carbon chart) added to the results.

Using carbon and oxygen isotopic abundance (either singularly or both isotopes), 88% of the samples in 7503 comparisons could be discriminated. This discrimination factor is larger than either single isotope and demonstrates the advantages to using multivariate analyses for the comparison of papers. Aside from discrimination powers, confidence in the results obtained can also be gained through the discrimination or inclusion of samples using multiple techniques (or isotopes).

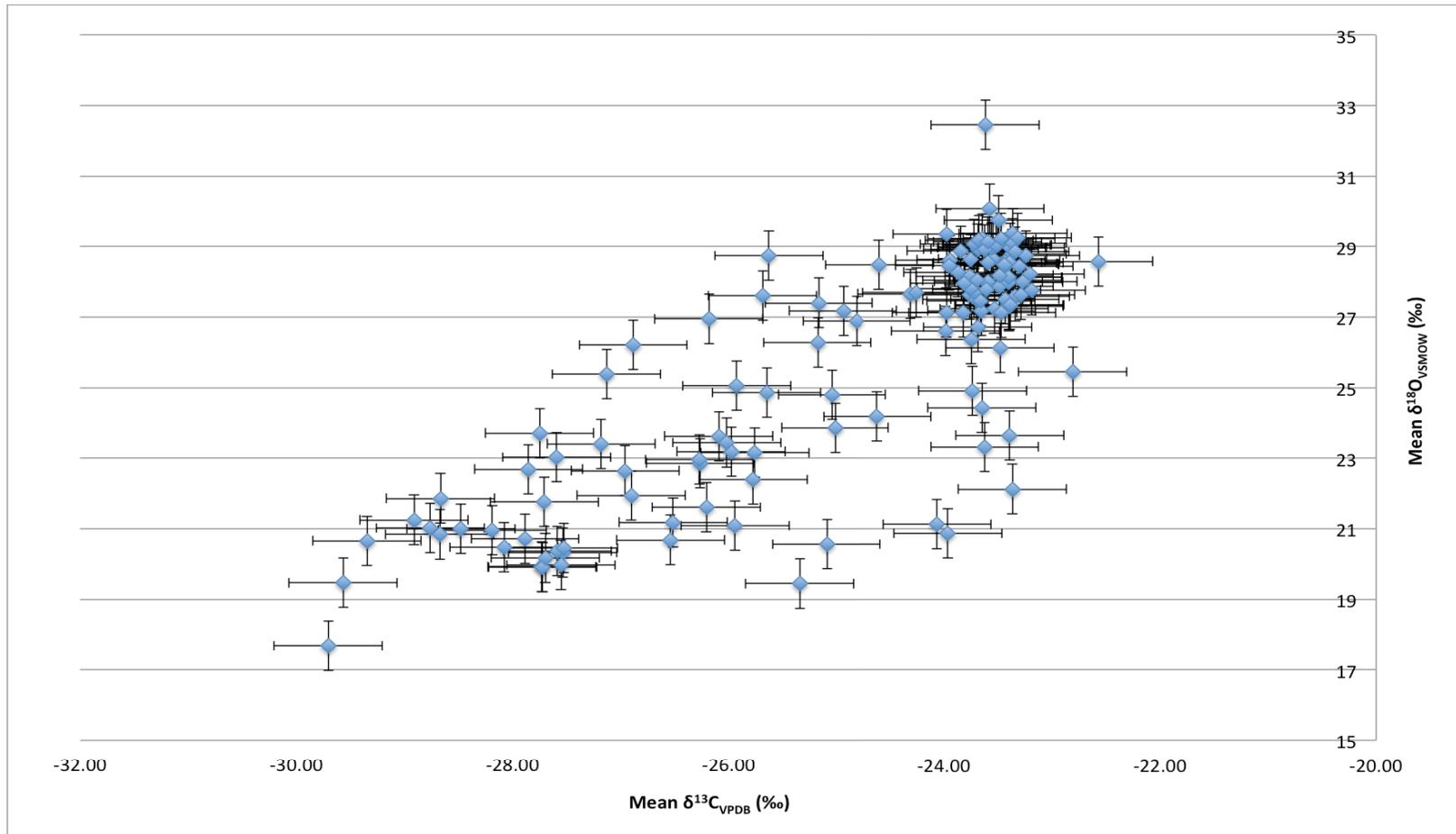


Figure 10.1: Background population samples measured for carbon and oxygen isotopic values. Error bars denote the 0.5 and 0.7‰ benchmark discrimination powers previously defined

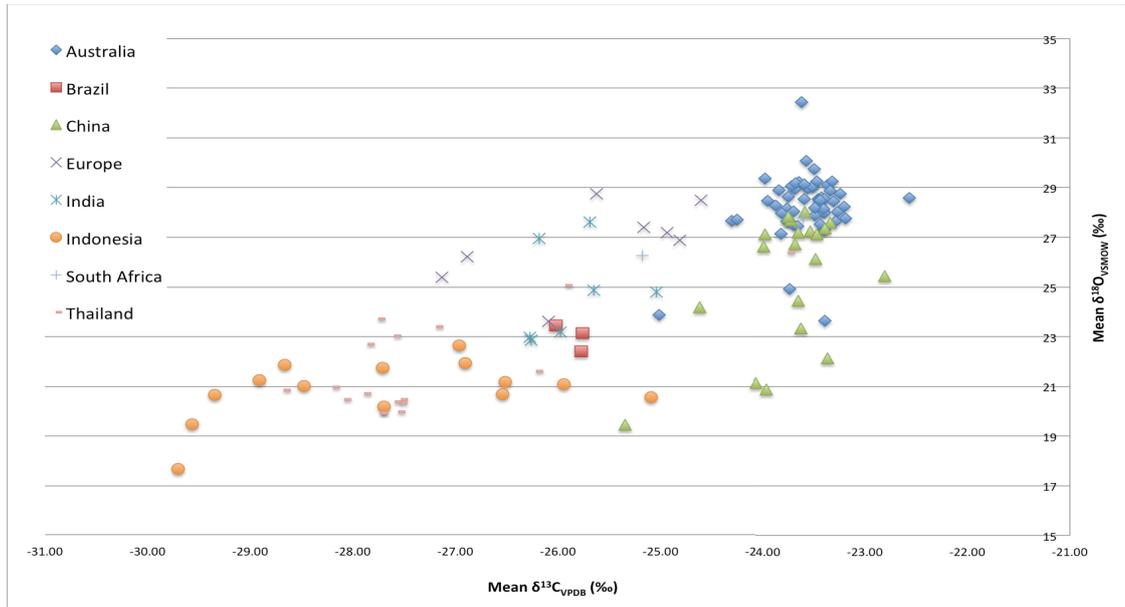


Figure 10.2: Background population samples measured for carbon and oxygen isotopic values, plotted by region of origin

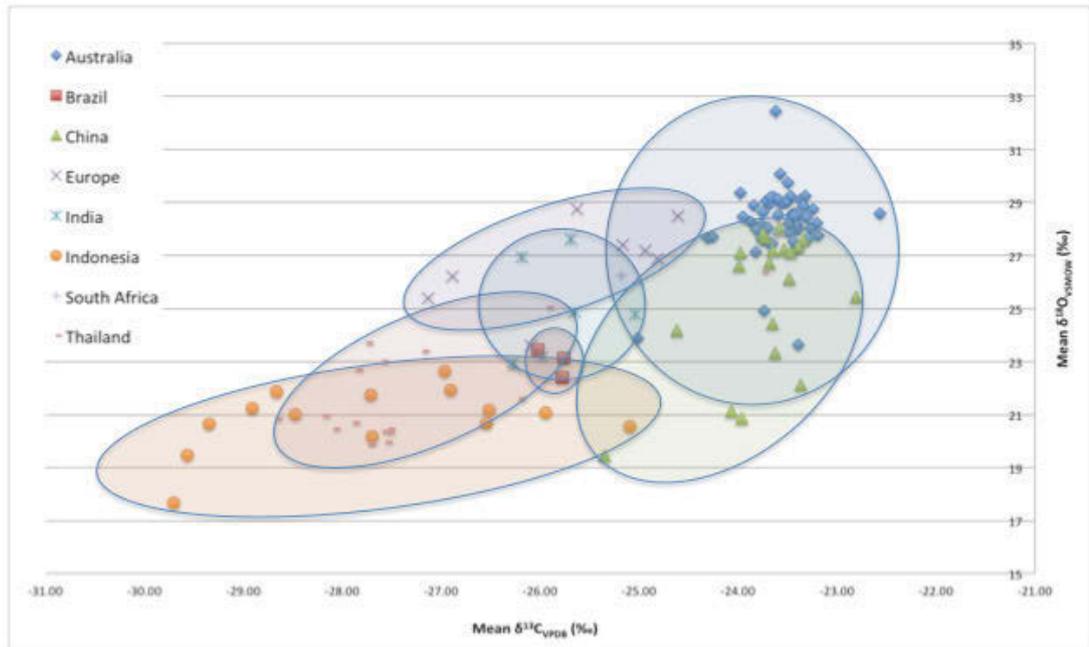


Figure 10.3: Background population samples measured for carbon and oxygen isotopic values, with regions highlighted

Although the region of origin plots Figure 10.2 and Figure 10.3 show that the majority of regions overlap in at least some way, there are a number of regions that sit distinctly apart and could be used to, in the very least, inform which regions a casework sample did not originate from. In particular, Australian and Indonesian produced samples are distinct from one another, as are the Australian, Thailand and Brazilian regions. It should be highlighted again however that a wider international sample study should be undertaken if papers from different countries are to be compared to this sample set. The potential for crossover with other international regions of production is shown in the European samples – which are somewhat distinct but overlap with the Australian region.

From the plot of the papers separated by region, it becomes clear that the Australian samples in particular are still comingled and may not be being discriminated. While in some sense this is expected given the lower variability in source material, it is still worth examining more closely to determine if discrimination is possible.

38 samples (30% of the sample set) of Australian origin were removed from the larger background sample set and compared to determine their discrimination using both carbon and oxygen isotopic abundance values. A plot of this sample set is shown in Figure 10.4, with error bars representing the 0.5 ‰ (carbon) and 0.7 ‰ (oxygen) benchmark discrimination values.

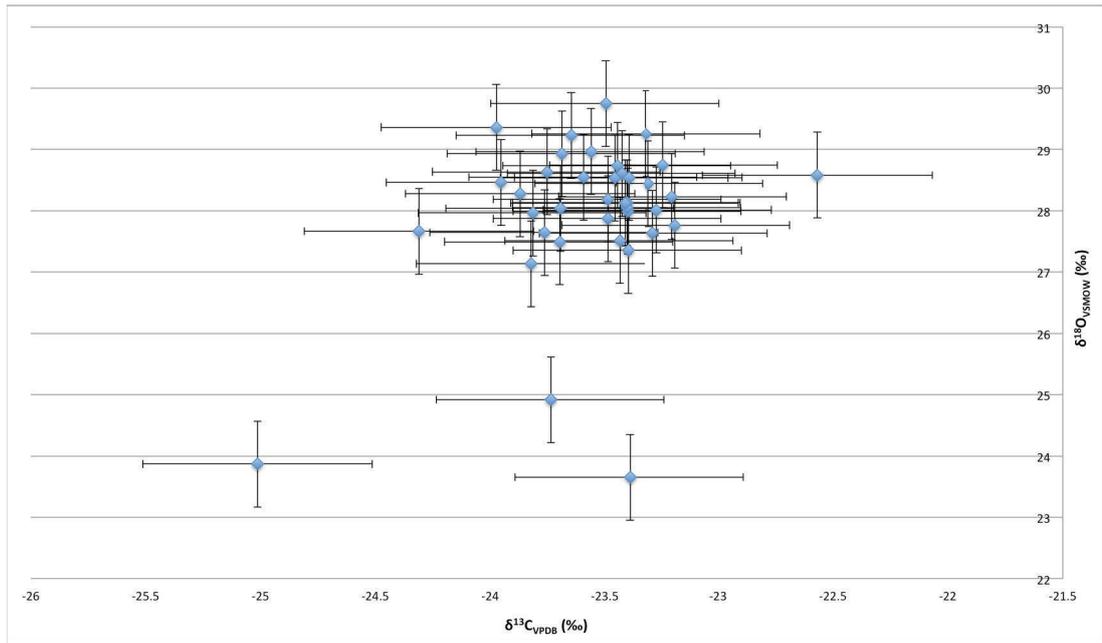


Figure 10.4: Australian produced samples measured for carbon and oxygen isotopic abundance values.

Figure 10.4 displays the close association and overlapping of these samples, suggesting that discrimination may be difficult. While some samples are distinctly different from the main group, there is a likelihood of discrimination for some samples given that the main sample group is sitting in a range of approx. 2 ‰ for carbon and 3.5 ‰ for oxygen.

The 38 Australian produced samples were compared to each other and the discrimination powers from single and combined isotopic values were calculated. The results of the 703 total comparisons made are shown in Table 10.1. The carbon and oxygen (total) values are the number of discriminated pairs in total for each isotope. The carbon and oxygen (singular) values are those comparison pairs that were only observed to be different using one isotope (either carbon or oxygen, not both). The combination value shown is where differences were observed in both the oxygen and carbon isotopic abundance values measured.

Isotope	Number of Sample Pairs Discriminated	Discrimination Power
Carbon (total)	168	0.24
Oxygen (total)	355	0.50
Carbon only (not Oxygen)	73	0.10
Oxygen only (not Carbon)	261	0.37
Both Oxygen and Carbon	95	0.14
Total Comparison Pairs Discriminated (703 comparisons – Carbon only + Oxygen only + Carbon and Oxygen)	429	0.61

Table 10.1: Discrimination of Australian samples using oxygen and carbon isotopes

The results of this comparison show that discrimination is possible, even when the samples have originated from the same production location. This indicates that there are still batch related differences that can be used for discrimination. This result confirms the differences seen for some brands in the homogeneity experiments conducted on the multiple reams from the same brand, sold at the same time. The oxygen isotope abundance values in particular show more variability and potential for discrimination. This indicates that in a practical examination protocol, oxygen isotopes should be measured before carbon.

Relating these results to the context of the paper production process experiments included in Chapter 6, it is hypothesized that the greater variability observed in the oxygen isotopic abundance values may be due to the first fractionation stage during pulping, converting the wood chip into cellulose fibres. Between the two measurement points, a difference of almost 4 ‰ was observed between the raw wood chip and the separated cellulose fibres. As this process utilises heat, pressure, chemicals and large amount of water to wash the refined/extracted cellulose fibres, it is likely that the process input (i.e. the chemicals and the water) used in this process would vary over time, changing the fractionation observed and the cellulose values of the produced paper. This is in addition to the source variation expected from the wood chip itself.

10.2.2. Using IRMS Results with Other Examination Techniques

As the 60 paper sample set measured using non-destructive and TLC results were compared to the sample 71 paper, the same 60 papers were removed from the larger background population and compared to sample 71 for carbon and oxygen isotopic discrimination.

Table 10.2 summarizes the results of the comparisons. Each marked column denotes a called difference. In the column representing the combination of carbon and oxygen or grammage and density, 'x' denotes a carbon or grammage called difference, 'y' denotes an oxygen or sheet density called difference and 'xy' denotes a called difference using both isotopes or both grammage and apparent sheet density. The results marked with a '*' are those sample pairs that have been discriminated with physical or light examination techniques that were not found to be different isotopically.

Sample Number	Light Exams	Grammage (x) Sheet Density (y) or Both (xy)	TLC	Carbon (x) + Oxygen (y) or Both (xy)
43		x*		
46		x		y
47				xy
48		x		
49		x*		
50			N/A	xy
53			x	y
54	x	y		xy
56		x	N/A	xy
66	x*	xy*	N/A	
67				
68			x	x
69		x		y
72				y
73				y
74				
75				
76				
77		y		x
79		y*	N/A	
80			N/A	x
81			N/A	
82			N/A	xy
83	x	x	N/A	xy
84			x	xy
85			N/A	xy
86		x*		
87				xy
88		xy	x	xy
89	x	xy	x	xy
90		x	N/A	xy
91		xy	N/A	xy
92			x	xy
93	x	x		xy
94	x	x		xy
95	x	xy		xy
96	x	xy		xy
97	x	y		xy
98	x	y	N/A	xy
99	x			xy
101	x	y	N/A	xy
102	x	xy	x	y

Sample Number	Light Exams	Grammage (x) Sheet Density (y) or Both (xy)	TLC	Carbon (x) + Oxygen (y) or Both (xy)
103	x		x	xy
104	x	x	x	xy
105		xy	x	xy
106	x	xy	x	xy
107		xy		y
108	x	xy	N/A	y
109	x	xy		xy
112	x			xy
113	x		x	xy
114	x		x	y
115	x	xy		xy
117				y
118				y
120		xy*		
121		x	x	xy
122			x	xy
124		x		xy
125		xy		xy
Total # Discriminated	21	35	15	48
Discrimination Power (DP)	0.35	0.58	0.25	0.80

Table 10.2: Comparison of techniques used to discriminate 60 samples of 80gsm paper.

Figure 10.5 collates the results in Table 10.2 into a flow chart that builds on the Discrimination Powers (DP) for each type of examination. There were 6 pairs that were discriminated using light and physical examination techniques that were not discriminated using carbon or oxygen isotopes. This has extended the total discrimination of this 60-sample set when compared to paper sample 71 to 54 pairs, a discrimination power of 0.9.

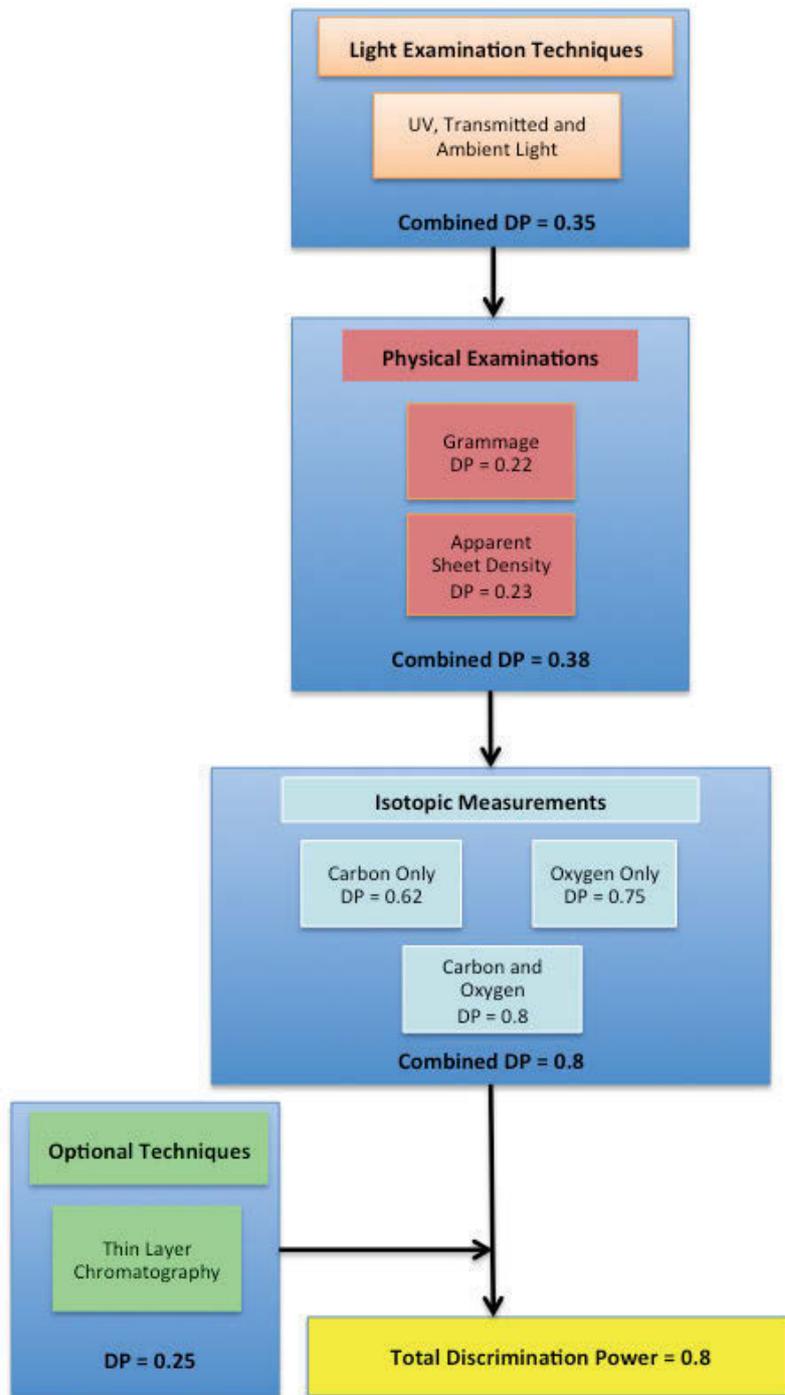


Figure 10.5: Summary of discrimination powers for range of techniques for comparison of 60-sample paper set with paper sample 71

The discrimination powers for the 60-sample set were calculated to be 0.6 and 0.75 for carbon and oxygen isotopes respectively. When combined, the two isotopes had a total discrimination power of 0.8 (48 of 60 samples compared to sample 71 were discriminated). The discrimination total contained 3 carbon only pairs, 11 oxygen only pairs and 34 pairs that were discriminated using both carbon and oxygen. These results again show that oxygen has a higher discrimination power than carbon but that in combination, the two isotopes can achieve a reasonably high discrimination, even within a limited sample set. The similarity of discrimination values also shows that this limited test set is representative for the wider population.

The grammage and apparent sheet density results were shown to discriminate six sample pairs that were not discriminated using IRMS. No additional pairs were discriminated using TLC. One of the additional pairs that were called different was also found to be different using light examination techniques. This pair of samples (papers 71 and 66) have been produced in Australia under different brand names and as such, have likely been produced to different grades using the same source materials.

Comparing these samples to samples that were produced at the same Australian mill around the same date (samples 67 and 74, 75 and 76) we see that discrimination was not called for these samples using isotopic measurements either. These results indicate that it may be possible for the two papers to be isotopically indistinguishable but have physical attributes that are different. These results suggest that it is worthwhile conducting these examinations prior to IRMS for discrimination purposes or as an additive result to strengthen the findings of other examination types however results of this kind are likely to lead to a challenging situation during interpretation.

For the light examinations, the same pair of samples (papers 71 and 66) that were discriminated using grammage were also discriminated using the light techniques. 72% (21 samples) of the isotopically discriminated samples were also discriminated using the light techniques.

When compared to the isotopic discrimination values, TLC was not found to add any discriminated pairs unique to the ones discriminated isotopically. Given that both TLC and IRMS are destructive techniques, there does not appear to be a significant benefit

in undertaking TLC prior to IRMS measurement. There would be benefit however in utilising TLC analysis alongside IRMS to strengthen the opinion able to be given at the completion of the examination. This would assist in cases where for example, the carbon isotopic abundance value was observed to be different but oxygen was not.

10.2.3. Proposed Examination Protocol

The purpose of defining an examination protocol is to outline an examination procedure that can be followed from evidence submission through to reporting. The interpretation stage in this protocol defines the questions that should be considered so that discrimination can be made. Provision also needs to be made for examination of papers using non-destructive techniques for cases where permission to undertake destructive examinations is not granted.

Based on the comparison results discussed above, the following flow chart in Figure 10.6 is proposed for the forensic examination and comparison of paper. While the experimental results show that IRMS is responsible for the majority of discrimination, inclusion of other paper examination techniques has been demonstrated to be beneficial. The turn around time for other paper examination techniques is also likely to be faster than IRMS, and more readily able to be provided as a preliminary indication of difference within the 48 hours generally required for intelligence level reporting.

While it was shown to be useful in discriminating papers based on filler content, XRD is not being included here as the technique is not currently available in the AFP laboratory. If it becomes available in future, a verification study will be conducted and the technique added to the standard procedure.

While a top down flowchart is provided in Figure 10.6, a concurrent examination protocol is possible if samples are removed from the bulk papers prior to humidification processes. While some correction is required for the weight measurements taken for grammage and apparent sheet density, this would allow all examinations to proceed within a shorter time period. It is also feasible to remove and

store samples from papers for IRMS and TLC, in case examination is required in future.

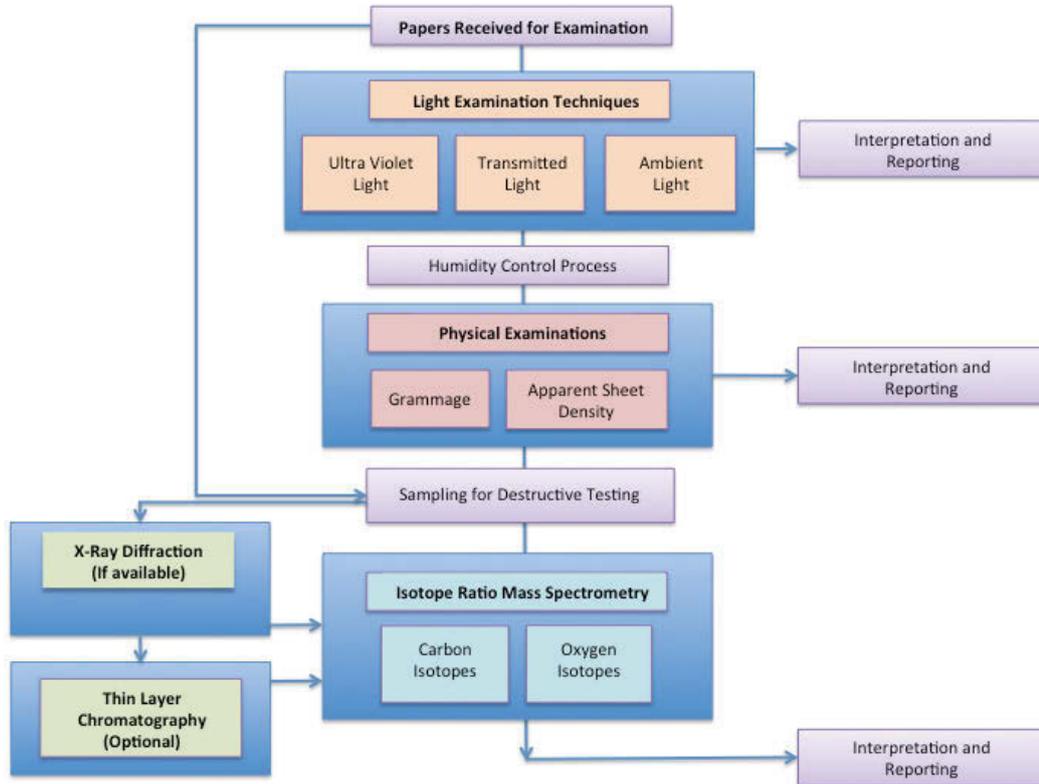


Figure 10.6: Proposed examination flow chart using a range of examination techniques

Within the proposed examination protocol, samples are received in the laboratory and the non-destructive techniques are undertaken prior to destructive techniques. A loop is provided however, to bypass the non-destructive techniques in favour of focusing on the more discriminating IRMS technique. TLC is included in this procedure, but as an optional technique to make clear that it is an accessory to IRMS to be used when appropriate. This is the same for XRD.

Provision is made after each grouping of techniques (light, physical or IRMS) for interpretation and reporting.

Where there are turnaround time constraints, permission to destruct is not given, or a more streamlined process is required, a concurrent examination procedure can be used

where samples are removed from the bulk papers prior to light and physical examination techniques. The risk with this procedure is incorrect calculation of grammage and subsequent calculation of apparent sheet density, due to sample removal. The time saved would primarily be from the humidification control processes required.

10.3. Interpretation of Results

A large number of publications exist that discuss the interpretation of examination results in a forensic context. These documents include quality and standards documents, in particular the recently published Australian Standard on interpretation AS5388.3 Forensic Analysis – Interpretation (2013a) and the European Network of Forensic Science Institutes Guideline for Evaluative Reporting in Forensic Science (2015).

From this document, interpretation is defined as a process whereby the examiner gives consideration to evaluating the information obtained from examinations, according to their ability to explain the issue in question. This can be achieved through the use of a combination of professional judgement or statistics, as long as the recognised limitations of the methods and instrumentation are taken into account.

Evaluation and interpretation is a comparative process, the outcomes of which may be formulated into hypotheses, which are concurrently evaluated in light of the results obtained. At the end of the process an opinion will be given by the expert as to the best supported conclusion (Australia, 2013b).

In European laboratories in particular, interpretation of forensic results follows an approach based on the Bayes Theorem (Armstrong and Hibbert, 2009, Biedermann et al., 2012, Hibbert and Armstrong, 2009, Taroni et al., 2004, ENFSI, 2015), which is defined as:

$$\text{Posterior odds} = \text{Prior odds} \times \text{Likelihood ratio}$$

where the likelihood ratio is an expression of the probability of the evidence observed, in the context of two or more alternate and opposing hypotheses. The two other factors in Bayes Theorem, the posterior odds (an expression of the probability of the two alternate hypotheses, given the evidence) and the prior odds (the probability of the two alternate hypotheses) are generally outside of the scope of most forensic examinations and expert roles.

In general, there are two ways that Bayesian interpretation and reporting is applied to forensic casework. The first is an expression of the likelihood ratio that is calculated into a numerical expression of the probability of the hypothesis. Further to this, some laboratories have been constructing Bayesian Networks – probability webs that describe the likelihood of evidence given a number of factors (including situational, context and investigative factors) - to increase or decrease a likelihood ratio based on a case based assessment of the evidence (Biedermann et al., 2012, Taroni et al., 2004).

In document examination, a number of papers have been published that have aimed to apply a Bayesian framework with numerical calculation of a likelihood ratio to handwriting examination and black toner evidence (Davis et al., 2012, Hepler et al., 2012, Taroni et al., 2012, Biedermann et al., 2011). A number of papers have also used calculation of the likelihood ratios as part of an isotopic interpretation framework (Carter et al., 2014, Farmer et al., 2009b) for tapes and white architectural paints.

While in use internationally, calculation and presentation of evidence using a numerical likelihood ratio is not commonplace for evidence types other than DNA in Australia. Concern over the accessibility and interpretation of the scales when presented to their intended audience – courts and juries – has meant that calculations of likelihood ratios in disciplines that do not naturally deal with numbers or statistics have been slow to change (Martire et al., 2013, Thompson and Newman, 2015).

Given the range of examination types utilised here and the numerous contextual differences that will arise in casework, a combined approach to calculating a likelihood ratio (based solely on the document examination results) is recommended for development in future. Due to the length and detail of this project, and the focus on producing practical results, a decision was made to not include such a model here.

This is an approach that would assist in weighing opinions for reporting however it would never fully replace examiner expertise and decision-making.

While a likelihood ratio calculation for just the IRMS analyses could be undertaken utilising the background database constructed along the same lines as presented by Farmer et al (2013), the collation of multiple examination types would require the use of more extensive modelling, where different examination results are used to determine a single likelihood ratio to describe the probability of two samples of paper coming from the same or different sources. An example of multiple examinations to form a combined likelihood ratio such can be found in (Taroni et al., 2006) which details the use of a Bayesian network to combine not only different results but different conditions related to the evidence (e.g. if the sample had been degraded).

The focus for this project will instead be on identifying the decision and weighting points for opinion formulation based on experience by the examiner. As the data collected here would be easily transformed to a Bayesian model, this is strongly recommended for future development. Although Bayesian interpretation is not included in this project, the research and data here provide a strong evidence-based foundation for the formulation of subjective opinions arising from these forensic paper examinations.

10.3.1. Factors Affecting Interpretation

The factors affecting the interpretation (and hence weighting) of results can be placed into two broad categories – data/analytical factors and casework factors. For each of these categories the factors include:

Data/analytical factors:

- Analytical error – false positives
- Homogeneity/variability of the material
- Number of techniques showing discrimination
- Types of techniques showing discrimination
- Size of difference observed

Case context factors:

- Purpose of document e.g. business, personal
- Type of document – handwritten, printed
- Environmental context (if known) e.g. exposure to heat, white powders
- Age of document – known or assumed e.g. printed date
- The question being posed by the examination e.g. page substitution, comparison with a suspect source, comparison between two or more questioned documents

These factors have been combined into a proposed interpretation workflow, included as Figure 10.7. To accompany the workflow, Table 10.3 discusses each of the factors in greater detail, including the positive and negative impacts of each factor on the confidence of the examiners opinion when calling a difference. The magnitude of the confidence will be used to determine what level on the reporting scale the result is reported at.

10.3.2. Proposed Interpretation Workflow

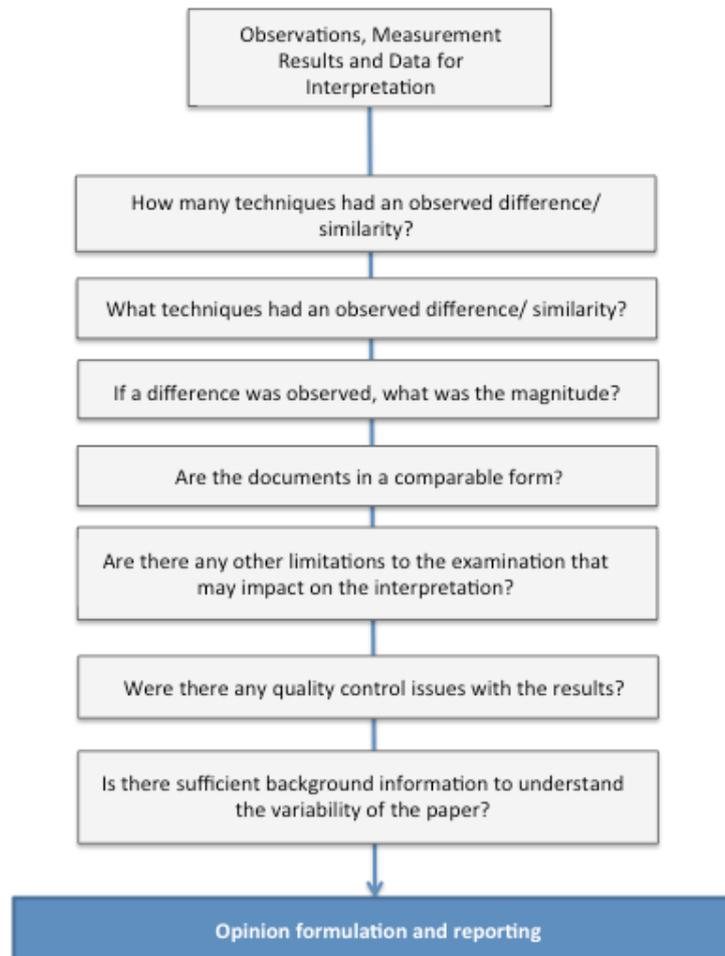


Figure 10.7: Interpretation factors for the reporting of paper examination techniques

Factors to be considered	Discussion
<p>How many techniques had an observed difference/similarity?</p> <p>If yes/no, what techniques?</p>	<p>Confidence of a different source can be gained when differences are observed using IRMS and further increased when a non-destructive technique/s can be added to a chemical one. When only the light and physical examination techniques show a difference, thought should be given as to whether two different grades of paper that have originated from the same source are present.</p> <p>If discrimination cannot be achieved using the techniques, in particular IRMS, then there is support that the samples cannot be differentiated and may have originated from the same source or a source with the same chemical and physical properties.</p>
<p>If a difference was observed, how large is it?</p>	<p>For numerical results, confidence is increased the further away the result is from the benchmark value. For subjective interpretations, this is also the case however the magnitude of the difference is determined at the examiners discretion.</p> <p>For numerical results that are close to the benchmark values, false positives need to be considered. Variability of the material and the method of sampling also needs to be checked and considered for its possible effect at this point. Some differences may still be called when two samples vary within the 95% confidence interval (2 x standard deviations) however a full opinion of difference of source should not be given in these cases.</p> <p>Are the results obtained from the examinations straight forward i.e. are the differences clear and for multiple page cases, are the differences repeatable? If so, there is no negative impact on the confidence level.</p> <p>The confidence level or weighting can be negatively impacted by results that are difficult to interpret (including ones that are close to the benchmark/discrimination values) or inconsistent across multiple pages that are expected to have the same results.</p>

<p>Are the documents in a comparable form?</p>	<p>Documents that are in a comparable form do not negatively effect the confidence. An example of documents with a comparable form are multiple pages that have been produced using the same printing technique, and that have been sampled appropriately.</p> <p>When documents have a limited comparability the confidence in the results of the comparison should be questioned. This is particularly the case when the papers have been significantly affected by a known environmental factor. An example this could be found in casework where a heavily handled piece of paper is requested to be compared to a blank/clean sheet seized from a suspect’s premises.</p>
<p>Are there any other limitations to the examination that may impact on the interpretation of the results?</p>	<p>If no, the confidence level previously reached will not be affected.</p> <p>The more limitations that are associated with the examination, the larger the impact. Limitations could include: forensically examined, environmentally affected or previously damaged documents.</p>
<p>Were there any quality control issues with the results?</p>	<p>Check measurement results, QA samples (for IRMS) and blanks to ensure that the measurements made were robust. Check data correction techniques to ensure correction was undertaken correctly. Check that handling with respect to hygroscopicity was conducted appropriately.</p>
<p>Is there sufficient background information to understand the variability of the paper?</p>	<p>Is the paper 80gsm standard white office paper or a different paper? Does a background population/sample need to be undertaken to confirm the variability of the material?</p>

Table 10.3: Interpretation factors with a discussion of their impacts on confidence.

For examination results that do not have an observable difference, interpretation is in some ways simpler and should focus on ensuring that the examinations conducted and the results produced are of a sufficient quality to ensure that the results are robust. This would be done through quality control processes built in to the examination

procedures and would likely require only a short revision of the methods used and the results obtained, generally through a technical review process.

10.4. Reporting of Results

Similarly to the Australian Standard for interpretation, an Australian Standard has been published to give guidance on the reporting required for forensic results. The standard AS5888.4 Forensic Analysis – Reporting (2013a) aims to standardize the format and content of reports in terms of communicating results and opinions and to provide the reader with the transparency required to properly assess its contents.

The main component that needs to be considered is the form of the reporting scale used and how the interpretation feeds into the expression and formulation of an opinion. The other more general requirements and structure of reporting will be facilitated or provided by the examiners quality assurance system or the jurisdiction they are working within.

With respect to opinion reporting scales, there are two general types in use in forensic reporting – probabilistic scales which express the strength of the opinion against two opposing alternate hypotheses and evaluative scales which typically express opinions in terms of the strength of the association e.g. in handwriting examinations that use an evaluative scale, the strength of the association between the specimen and questioned material.

A verbal probabilistic scale is proposed here, that will have two weighted support levels for values that are discriminated, next to an inconclusive opinion. For the opposing side of the scale, when similarities are observed, a single unweighted statement will be utilised, which will include an additional comment on the limitations related to positively associating source in mass manufactured materials.

This approach is similar in nature to the existing schemes in other trace chemical evidence reporting. This scale can also be adapted and used if numerical likelihood ratios are developed in future however expansion of the similarities side of the scale

would require significant database development and the use of further chemical testing using a greater number of techniques.

A suggested reporting appendix to attach to technical reports is included in Appendix 4.

10.4.1. Proposed Reporting Scale and Report Content

In general, the propositions to be considered will be:

Proposition 1: The specimen and questioned papers originated from the same source

Proposition 2: The specimen and questioned papers originated from a different source

As papers are a manufactured material, even using highly discriminating techniques such as the ones here, the possibility that more than one source has produced a paper of the same physical, optical, chemical and isotopic properties cannot be excluded. Additionally, the size of a source or batch produced in the same facility cannot be defined. Due to this, when no difference is observed between questioned and specimen papers, in general, it is not possible to weigh the proposition in terms of the papers having solely originated from the same source as it is equally likely that another batch (hence source) produced a paper of the same properties.

The weight of the evidence should be scaled to formulate one of the following opinions:

- There is support for the proposition that the specimen and questioned papers originated from the same source, or from a source with the same physical, optical, chemical and isotopic properties, rather than a different one.
- The evidence favours neither proposition and the results are inconclusive.
- There is support for the proposition that the specimen and questioned papers may have originated from different sources, rather than from the same one.

- There is strong support for the proposition that the specimen and questioned papers originated from different sources, rather than from the same one.

10.5. Conclusions

A comparison between the discrimination powers for carbon and oxygen isotopes, and for all other examination types was conducted. Oxygen isotopes were observed to be the most discriminating of the techniques and with carbon, resulted in a total discrimination of 89% of the 123-paper sample background population.

Although their discrimination powers were observed to be lower, the physical and light examination techniques were shown to be complimentary and are proposed to be used prior to IRMS measurement, or as a substitute when permission for destructive sampling is not granted. TLC comparisons were not found to add new information to the IRMS results and hence are seen as an optional technique for use in cases where the other results obtained may be ambiguous or when IRMS is not available. Where available, XRD should be added to identify the types of filler materials in the samples, prior to IRMS.

An examination protocol that commences with non-destructive techniques followed by the destructive chemical techniques is proposed for the examination of papers in casework. To sit alongside this, an interpretation and reporting framework has been proposed for the expression of results.

Future work should focus on the transformation of the data collected and the interpretation framework into a Bayesian framework using likelihood ratios to better inform the confidence statements utilised in reporting.

Chapter 11

11. Blind Trial of Examination Protocol

11.1. Introduction

While the experiments detailed in the previous chapters have demonstrated discrimination between paper samples and an examination protocol based on these results proposed, samples that mimic casework scenarios have yet to be tested. The aim of this chapter is to conduct a blind trial to test the proposed protocol in a manner that removes bias and is the closest comparison available to future casework applications. This is an essential final step to ensure that the procedure developed is ready for accreditation and application in an operational context.

The trials to be tested include:

- Single page comparisons, likely to be requested in cases such as threatening or anonymous correspondence where common source is required to be established or in cases where multiple pages have been produced, which purport to have been produced over a time range but may have been produced at one time.
- Questioned to specimen comparisons where single page questioned documents are to be compared to papers collected from one or more suspect reams, likely to have been seized from printers.
- Cases of suspected page substitutions in documents such as contracts and wills. As modern typed documents are typically easy to reproduce, page substitutions can difficult to detect if the substitution is undertaken with care.

The overall aim of conducting these blind trials is to determine if the examination, interpretation and reporting protocols are appropriate and effective for the casework examination of papers.

11.2. Method and Materials

11.2.1. Organisation of Blind Trials

As these trials involved the examination of physical materials, a single blind trial was appropriate, where the each of the trials involved testing and interpretation by one researcher without prior knowledge of the expected result. A colleague that has previously not been involved with this project collated the trials. The papers were selected from the 125 paper background samples and all of the papers were selected and combined randomly with no reference to prior results. The trial samples were marked with a sample number and placed into plastic sheet protectors marked A to H.

8 tests were prepared that varied amongst the casework examination types as follows:

- 5 x single page (1:1) comparisons, with papers marked as samples 1-10, designated tests A-E
- 1 x single toner printed page marked as paper sample 11, to be compared to a single specimen ream with three pages marked as paper sample 12, designated test F
- 1 x single toner printed page marked as paper sample 13, to be compared to three specimen reams, containing three pages in each specimen, marked as paper samples 14, 15 and 16, designated test G
- 1 x seven page toner printed document that may have had a page substituted, with papers marked as samples 17 – 23, designated test H

To add complexity and to more closely mimic casework situations, the questioned paper samples within tests F and G and the page substitution trial test H contained documents that had been produced using a toner printing process. Toner was selected over an inkjet printing process as it was shown to have a greater effect on the isotopic abundance values measured if included when the paper was sampled. Toner printing was introduced onto the pages in a typical document style (full page of text with single spacing in paragraphs), across the entirety of the page. An image of one of the printed documents is shown in Figure 11.1 below.

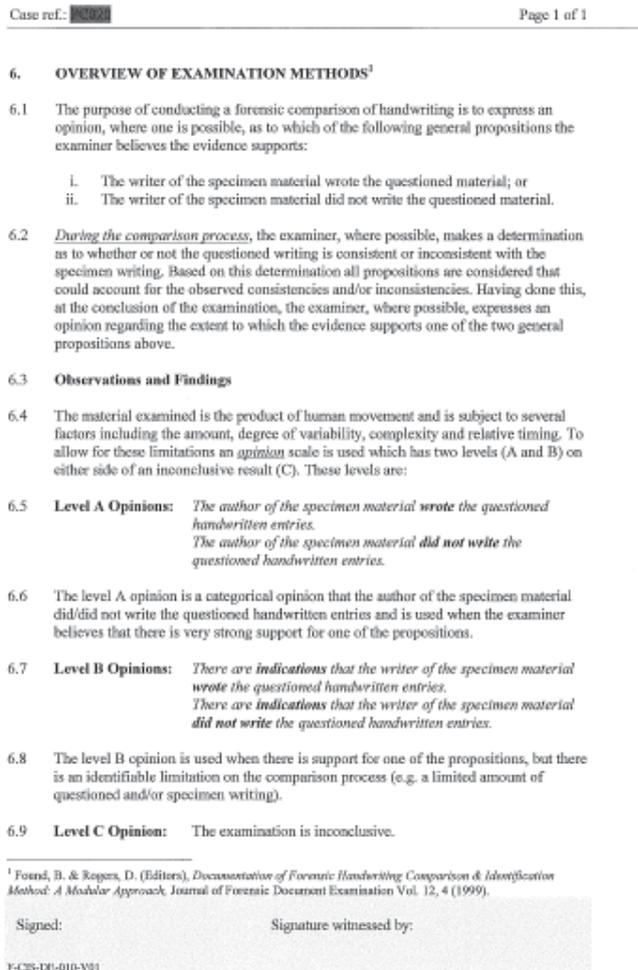


Figure 11.1: Image of paper sample 11 from test F showing toner printing on page.

The tests to be undertaken (as per the proposed examination flowchart) include:

- Light examination techniques – UV, transmitted and ambient light
- Physical examination techniques – grammage and apparent sheet density
- Measurement of carbon and oxygen isotopes using IRMS

A decision was made not to include TLC comparisons in this blind trial as it does not form part of the core examination protocol being proposed. All analytical methods and sampling protocols utilised, excluding the carbon IRMS method, are based on the methods presented in previous chapters.

As per previous experiments, no paper humidity control process will be utilised as it is assumed that the paper moisture content would have equilibrated during storage time.

For casework, humidification process (as discussed in prior chapters) would need to be followed.

For the trials that include 3 sheets of specimen paper for the suspect reams/printers, the samples were treated in the following manner:

- For the light examination techniques, each of the 3 sheets was cross-compared to ensure homogeneity. If homogeneity was observed one sheet was selected for comparison against the questioned paper
- For the grammage and apparent sheet density, each of the 3 sheets was measured separately and a combined mean value calculated for comparison to the questioned sheet
- For the IRMS analysis, triplicate samples were taken from each of the 3 sheets (9 replicates in total) and a mean value calculated for comparison to the questioned sheet. When no difference was observed between the sheets, these were combined with a pooled mean and standard deviation value obtained for all sheets.

For the printed papers a 15X hand magnifier was used to identify regions that were free from extraneous toner before the samples were removed from the sheet.

11.2.2. Change to Carbon Method

Between the completion of the carbon isotope work presented in Chapters 2-4 and the blind trials, the autosampler attached to the Flash EATM 1112 was upgraded to a Costech Zero Blank autosampler (Thermo Finnigan, Sydney, Australia). The EA method was modified to account for the change in operation of the autosampler, with the oxygen pulse extended from 3 to 10 seconds at 250mL/min and the sample dropped at 5 seconds. While this is likely to be an excess of oxygen, the sample combustion was observed to be efficient using these settings. Production of nitrous oxide (N₂O) was considered (which is known to cause isobaric interference) but based on prior experiments and the low proportion of nitrogen in paper, the likelihood of this type of interference was considered to be low. No interference was observed in the chromatography. The EA temperatures, including the GC column, and packing materials were unchanged.

11.3. Results and Discussion

The results of each test undertaken will be presented here separately, with interpretation and a final conclusion based on the results included. The observations and notes made will be in a similar format to casework. Differences are called based on observations (for the light examination techniques) or the benchmark values defined (for physical examinations and isotopic data). For ease of reference, the benchmark values are 0.5 ‰ for $\delta^{13}\text{C}_{\text{VPDB}}$, 0.7 ‰ for $\delta^{18}\text{O}_{\text{VSMOW}}$, 2.5gsm for grammage and 35kg/m^3 for apparent sheet density.

An interpretation table based on the one included in Chapter 10 will be used for each test to aid in the weighting of the final subjective opinion. While this will not be used directly for casework, a discussion of the key aspects presented here should be included in casework notes.

11.3.1. Trial A

Some visible differences were observed between paper 1 and paper 2 in ambient light. This included a slight colour difference (paper 1 appears to be more yellow than sample 2). The colour difference in particular is visible in the transmitted light image. Lines from the production process are visible under UV light at 365nm in paper 2. No lines are observed on paper 1. These results are shown in Figure 11.2 below.

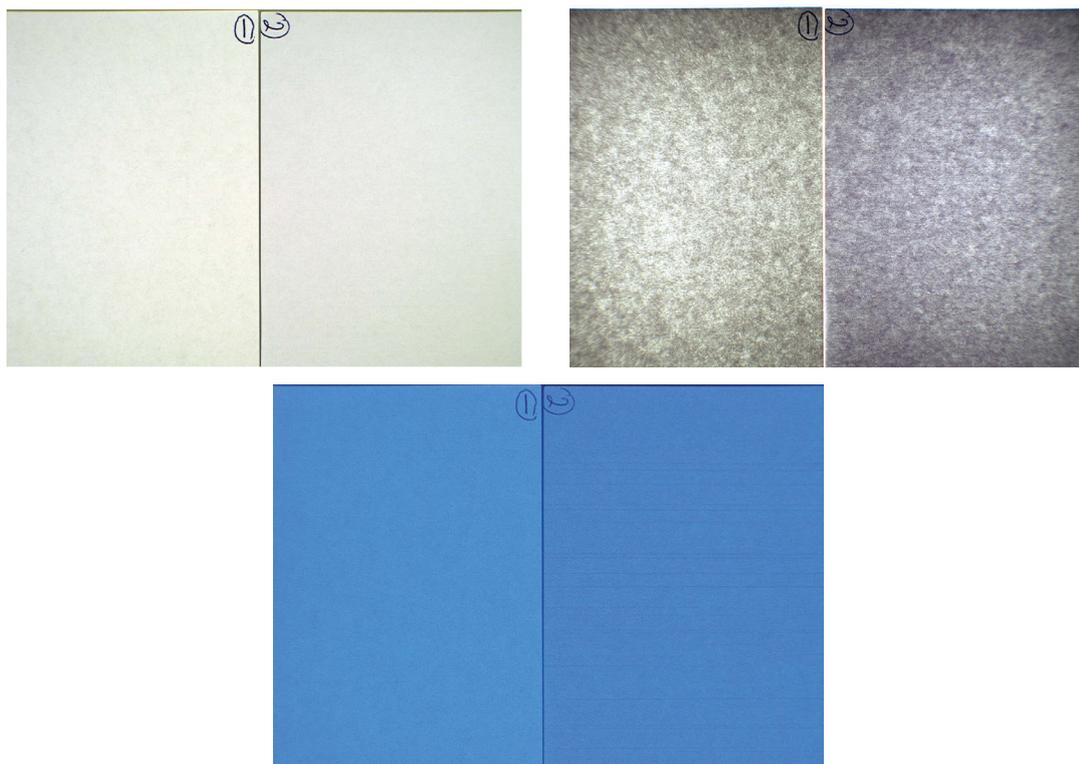


Figure 11.2: Light examination including ambient, transmitted and UV (at 256nm) for trial A

Table 11.1 contains the data obtained for the two samples:

	Paper 1	Paper 2	Difference
Grammage (gsm)	83.25	78.852	4.39
Apparent Sheet Density (kg/m³)	803.53	786.16	17.36
Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev } (\text{‰})$	-23.61 ± 0.08	-28.87 ± 0.05	5.27
Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev } (\text{‰})$	29.40 ± 0.21	21.96 ± 0.27	7.44

Table 11.1: Results of grammage, apparent sheet density and isotopic measurements for blind trial A

The interpretation workflow for these results is contained in Table 11.2:

Factors to be considered	Discussion
<p>Did the results of any techniques have an observed difference/ similarity?</p> <p>If yes, what technique/s?</p> <p>If yes, how large was the difference?</p>	<p>Differences were observed in four techniques - light examination, grammage and both carbon and oxygen IRMS.</p> <p>The apparent sheet density could not be differentiated using the benchmark value however the two sample values were not close.</p> <p>Differences in construction (colour and opacity) were observed using the light examination techniques. Grammage between the sheets was significantly different – more than double the benchmark value.</p> <p>Significant differences in the isotopic abundance values of the papers were observed: 5‰ for carbon, 7‰ for oxygen.</p> <p>While the apparent sheet density did not show a difference, this does not lower the confidence given the difference in the apparent sheet density and the magnitude of the difference in the IRMS results.</p>
<p>Are the documents in a comparable form?</p>	<p>Yes</p>
<p>Are there any other identifiable limitations to the examination that may impact on the interpretation of the results?</p>	<p>No limitations were identified.</p>

Table 11.2: Interpretation considerations for blind trial A

Final Opinion:

Based on these examination results, in my opinion there is strong support for the proposition that the two paper samples originated from different sources, rather than from the same one.

11.3.2. Trial B

No overt differences were observed between paper 3 and paper 4 in ambient, transmitted or UV light. There may be a slight colour difference, with paper 4 appearing more yellow than paper 3. These results are shown in Figure 11.3 below.

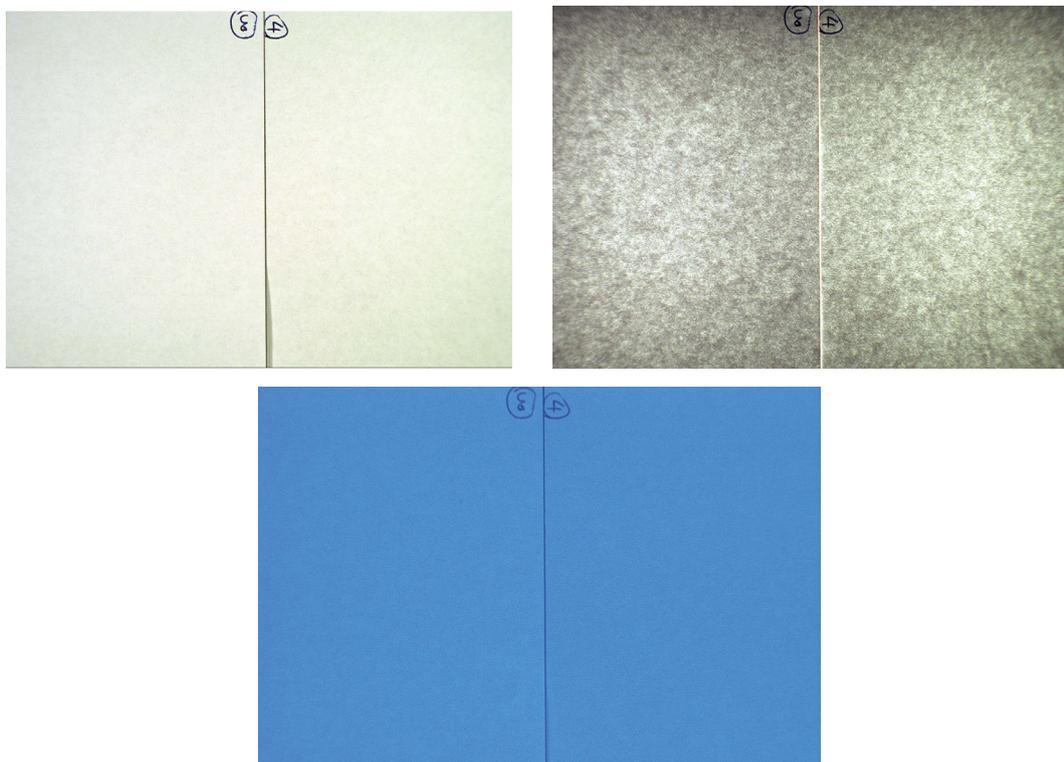


Figure 11.3: Light examination including ambient, transmitted and UV (at 256nm) for trial B

Table 11.3 contains the data obtained for the two samples:

	Paper 3	Paper 4	Difference
Grammage (gsm)	79.97	77.31	2.66
Apparent Sheet Density (kg/m ³)	811.89	796.98	14.88
Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev } (\text{‰})$	-23.56 ± 0.07	-23.51 ± 0.02	0.05
Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev } (\text{‰})$	29.29 ± 0.09	29.51 ± 0.02	0.22

Table 11.3: Results of grammage, apparent sheet density and isotopic measurements for blind trial B

The interpretation workflow for these results is presented in Table 11.4:

Factors to be considered	Discussion
<p>Did the results of any techniques have an observed difference/similarity?</p> <p>If yes, what technique/s?</p> <p>If yes, how large was the difference?</p>	<p>A difference larger than the benchmark value was observed in the grammage results. The difference observed was close to the benchmark value.</p> <p>No difference was observed using light examination techniques, apparent sheet density or IRMS. The difference between the carbon and the oxygen isotopic results is very small and within the instrument measurement uncertainty.</p>
<p>Are the documents in a comparable form?</p>	<p>Yes</p>
<p>Are there any other identifiable limitations to the examination that may impact on the interpretation of the results?</p>	<p>No limitations were identified.</p>

Table 11.4: Interpretation considerations for blind trial B

Final Opinion:

Based on these examination results, in my opinion there is support for the proposition that the specimen and questioned papers originated from the same source, or from a source with the same physical, optical and isotopic properties, rather than from different ones.

11.3.3. Trial C

No overt differences were observed between paper 5 and paper 6 in ambient, transmitted or UV light. Using transmitted and UV light, similarities were observed in the paper fibre dispersion of the sheets. No colour differences were observed. These results are shown in Figure 11.4 below.

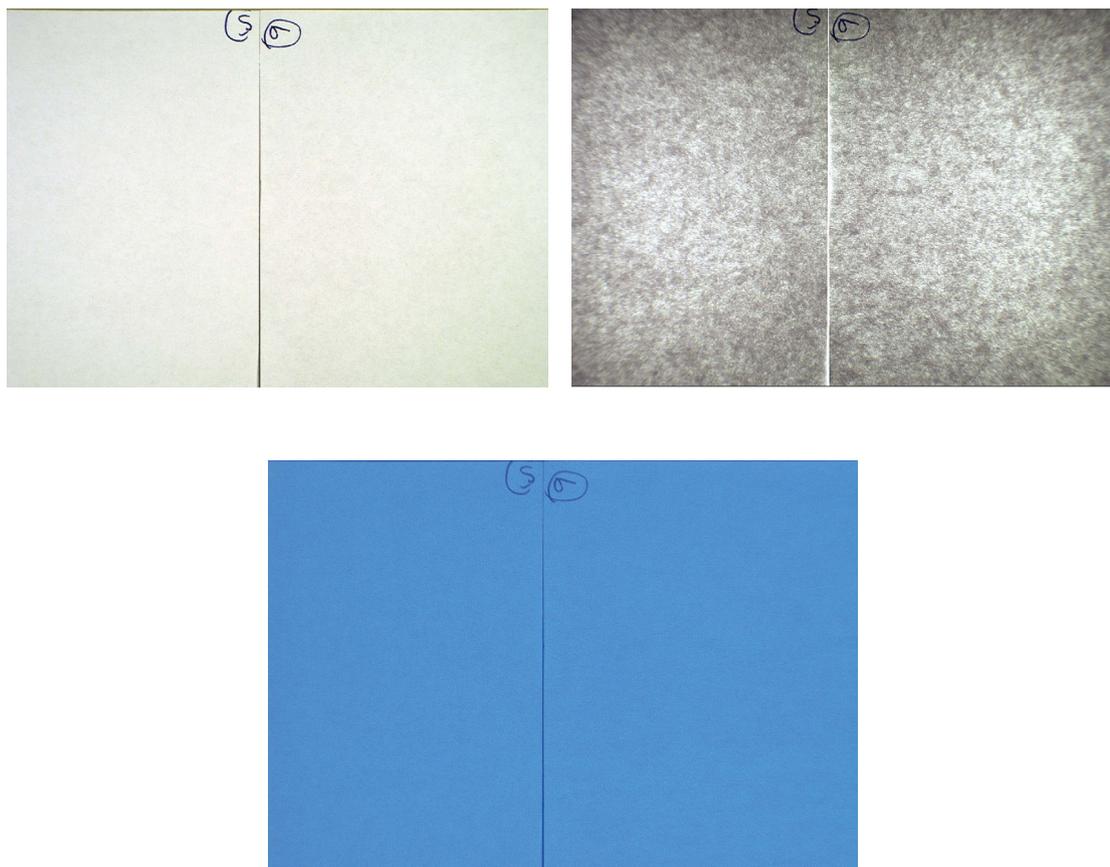


Figure 11.4: Light examination including ambient, transmitted and UV (at 256nm) for trial C

Table 11.5 contains the data obtained for the two samples:

	Paper 5	Paper 6	Difference
Grammage (gsm)	84.98	83.12	1.86
Apparent Sheet Density (kg/m ³)	857.59	853.41	4.189
Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev } (\text{‰})$	-23.71 ± 0.09	-23.67 ± 0.11	0.03
Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev } (\text{‰})$	29.57 ± 0.22	29.50 ± 0.02	0.07

Table 11.5: Results of grammage, apparent sheet density and isotopic measurements for blind trial C

The interpretation workflow for these results is presented in Table 11.6:

Factors to be considered	Discussion
<p>Did the results of any techniques have an observed difference/similarity?</p> <p>If yes, what technique/s?</p> <p>If yes, how large was the difference?</p>	<p>No differences were observed in any of the tests. Observed similarities were strengthened by the low standard deviation values calculated.</p> <p>The observed differences are very small, as are the standard deviations of sample measurements. For the IRMS results in particular, this indicates good analytical precision.</p>
<p>Are the documents in a comparable form?</p>	<p>Yes</p>
<p>Are there any other identifiable limitations to the examination that may impact on the interpretation of the results?</p>	<p>No limitations were identified.</p>

Table 11.6: Interpretation considerations for blind trial C

Final Opinion:

Based on these examination results, in my opinion there is support for the proposition that the specimen and questioned papers originated from the same source, or from a source with the same physical, optical and isotopic properties, rather than from different ones.

11.3.4. Trial D

Differences were observed between paper 7 and paper 8 using transmitted light. Differences include the dispersion of the fibres throughout the paper and in the texture on the surface of the paper (particularly when looking at paper 8). Similarities and no overt differences were observed using white and UV light. Given the mixture of results, the observations made with transmitted light should be used with caution. These results are shown in Figure 11.5 below.

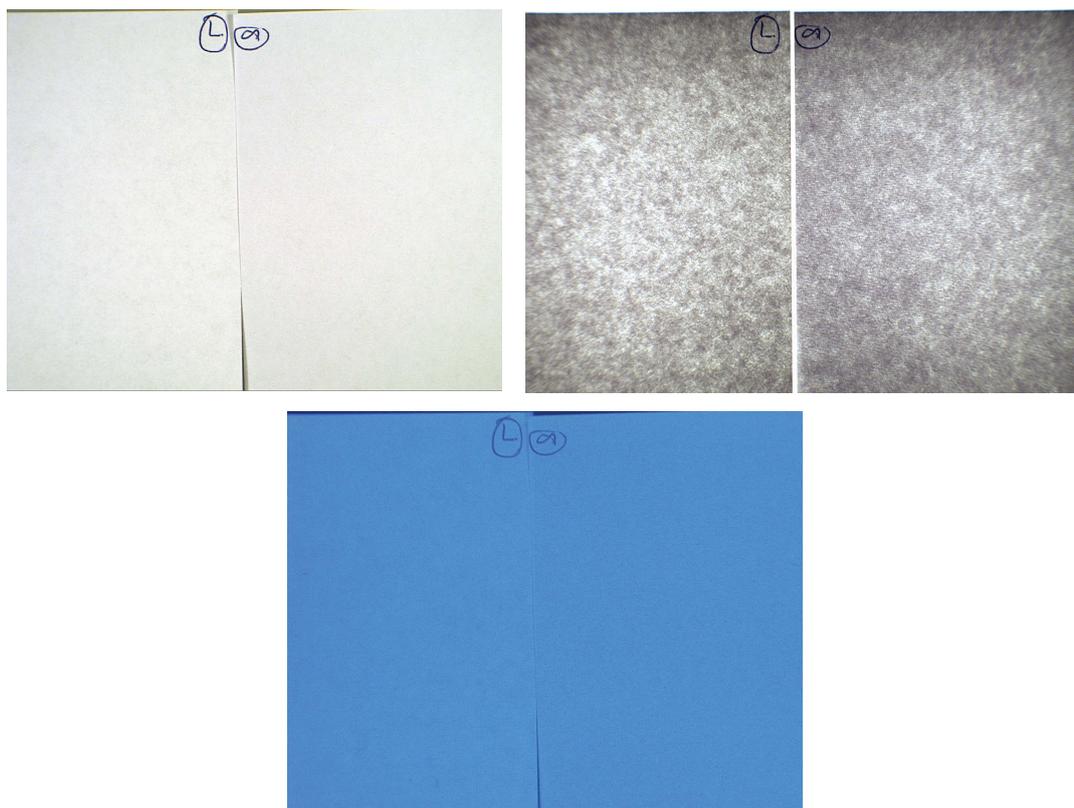


Figure 11.5: Light examination including ambient, transmitted and UV (at 256nm) for trial D

Table 11.7 contains data obtained for the two samples:

	Paper 7	Paper 8	Difference
Grammage (gsm)	83.81	83.11	0.69
Apparent Sheet Density (kg/m ³)	843.12	805.34	37.78
Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev } (\text{‰})$	-23.72 ± 0.03	-26.45 ± 0.04	2.74
Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev } (\text{‰})$	28.75 ± 1.10	23.08 ± 0.12	5.66

Table 11.7: Results of grammage, apparent sheet density and isotopic measurements for blind trial D

The interpretation workflow for these results is presented in Table 11.8:

Factors to be considered	Discussion
<p>Did the results of any techniques have an observed difference/similarity?</p> <p>If yes, what technique/s?</p> <p>If yes, how large was the difference?</p>	<p>Three techniques, including both isotopes, discriminated the samples. Differences were observed in the apparent sheet density. Large differences were observed in the carbon and oxygen isotopic abundance values (2.7‰ for carbon, 5.6‰ for oxygen) lending weight to the proposition of different source.</p> <p>Grammage, ambient and UV light were observed to be similar</p>
<p>Are the documents in a comparable form?</p>	<p>Yes</p>
<p>Are there any other identifiable limitations to the examination that may impact on the interpretation of the results?</p>	<p>No limitations were identified.</p>

Table 11.8: Interpretation considerations for blind trial D

Final Opinion:

Based on these examination results, in my opinion there is strong support for the proposition that the two paper samples originated from different sources, rather than from the same one.

11.3.5. Trial E

Similarities and no differences were observed between paper 9 and paper 10 when observed using ambient, transmitted and UV light. These results are shown in Figure 11.6 below.

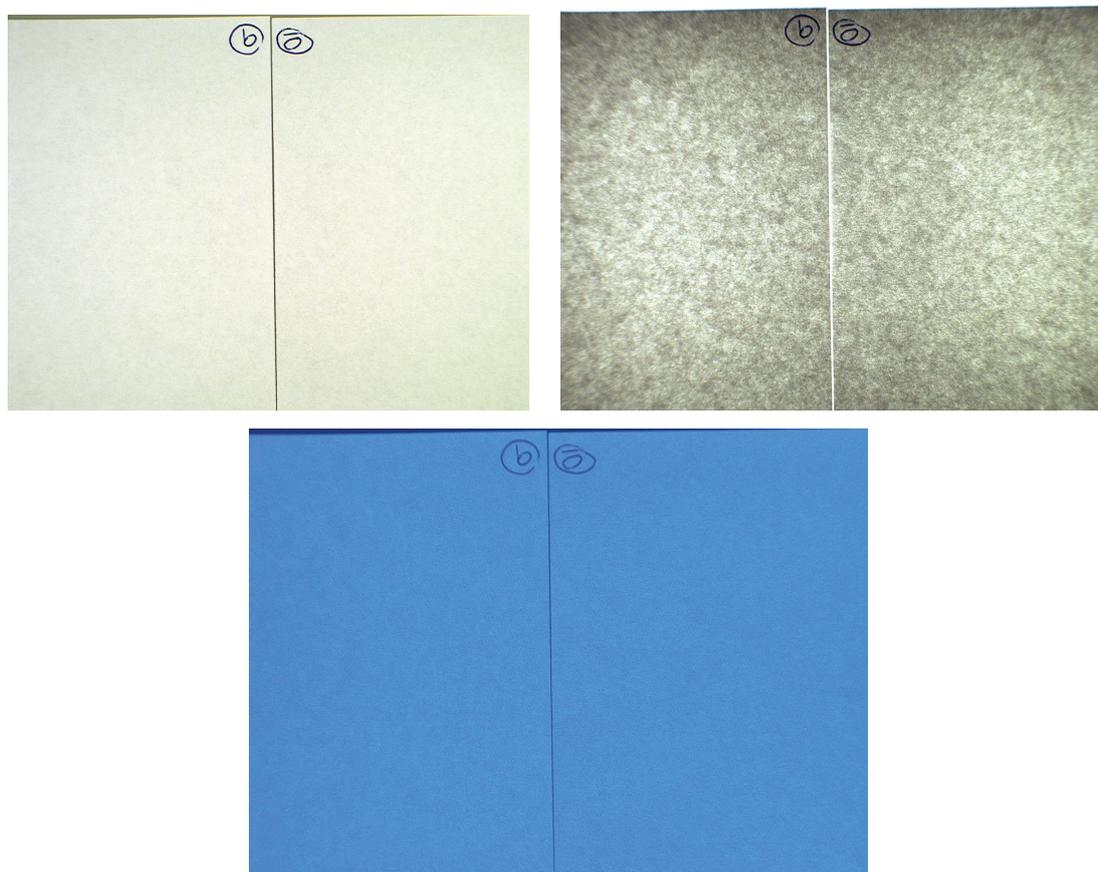


Figure 11.6: Light examination including ambient, transmitted and UV (at 256nm) for trial E

Table 11.9 contains the data obtained for the two samples:

	Paper 9	Paper 10	Difference
Grammage (gsm)	79.06	77.64	1.42
Apparent Sheet Density (kg/m ³)	790.56	778.72	11.83
Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev } (\text{‰})$	-23.56 ± 0.07	-23.50 ± 0.04	0.06
Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev } (\text{‰})$	29.33 ± 0.35	29.54 ± 0.12	0.20

Table 11.9: Results of grammage, apparent sheet density and isotopic measurements for blind trial

E

The interpretation workflow for these results is presented in Table 11.10:

Factors to be considered	Discussion
<p>Did the results of any techniques have an observed difference/similarity?</p> <p>If yes, what technique/s?</p> <p>If yes, how large was the difference?</p>	<p>Nil differences were observed. Similarities were observed for all techniques, in particular for the IRMS results which are very closely grouped.</p> <p>Similarities were observed between paper 7 and paper 8 using both carbon and oxygen isotopes. The values would not be discriminated using the benchmark values and the measurement standard deviations indicate good analytical precision.</p>
<p>Are the documents in a comparable form?</p>	<p>Yes</p>
<p>Are there any other identifiable limitations to the examination that may impact on the interpretation of the results?</p>	<p>No limitations were identified.</p>

Table 11.10: Interpretation considerations for blind trial E

Final Opinion:

Based on the observations and results of the tests conducted, in my opinion, there is support for the proposition that the specimen and questioned papers originated from the same source, or from a source with the same physical, optical, chemical and isotopic properties, rather than from a different one.

11.3.6. Trial F

As three pages were provided for paper 12 for comparison against paper 11, paper 12 was cross-compared initially to ensure that it was consistent. For the light examination techniques, and the images presented here, once consistency was ensured a single page (12a) was selected for direct comparison against paper 11.

No difference was observed using transmitted light between papers 11 and 12. Small differences were observed when viewing the papers with ambient light (paper 11 appeared slightly more yellow) and UV (paper 11 is duller than paper 12). Given the toner printing on paper 11, which may have changed the look/colour of the sheet and the issues with discrimination using UV alone, a clear difference was not seen using light examination techniques. These results are shown in Figure 11.7 below.



Figure 11.7: Light examination including ambient, transmitted and UV (at 256nm) for trial F

Table 11.1 contains the data was obtained for the two samples:

	Paper 11	Paper 12	Difference
Grammage (gsm)	79.76	82.01	2.24
Apparent Sheet Density (kg/m³)	790.54	779.21	11.34
Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev}$ (‰)	-28.22 \pm 0.06	-29.13 \pm 0.16	0.91
Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev}$ (‰)	22.66 \pm 0.15	21.68 \pm 0.18	0.98

Table 11.11: Results of grammage, apparent sheet density and isotopic measurements for blind trial F

The interpretation workflow for these results is presented in Table 11.12:

Factors to be considered	Discussion
<p>Did the results of any techniques have an observed difference/similarity?</p> <p>If yes, what technique/s?</p> <p>If yes, how large was the difference?</p>	<p>Differences were observed using both carbon and oxygen isotopic values. No difference was observed when comparing the papers using grammage and apparent sheet density.</p> <p>The IRMS results show a difference between the papers, confirming the observations made with the light examination results. The differences are not large but are larger than both the benchmark values.</p>
<p>Are the documents in a comparable form?</p>	<p>Yes</p>
<p>Are there any other identifiable limitations to the examination that may impact on the interpretation of the results?</p>	<p>No limitations were identified.</p>

Table 11.12: Interpretation considerations for blind trial F

Final Opinion:

Based on the observations and results of the tests conducted, in my opinion, there is support for the proposition that the specimen and questioned papers originated from different sources, rather than from the same one.

11.3.7. Trial G

Three specimen/suspect paper samples, with three pages per specimen, were compared to toner printed paper 13.

Paper 14 was observed to be different to paper 13 when compared with ambient and transmitted light. No other overt differences were observed when comparing paper 13 with the specimen papers 15 and 16. These results are shown in Figure 11.8 below.

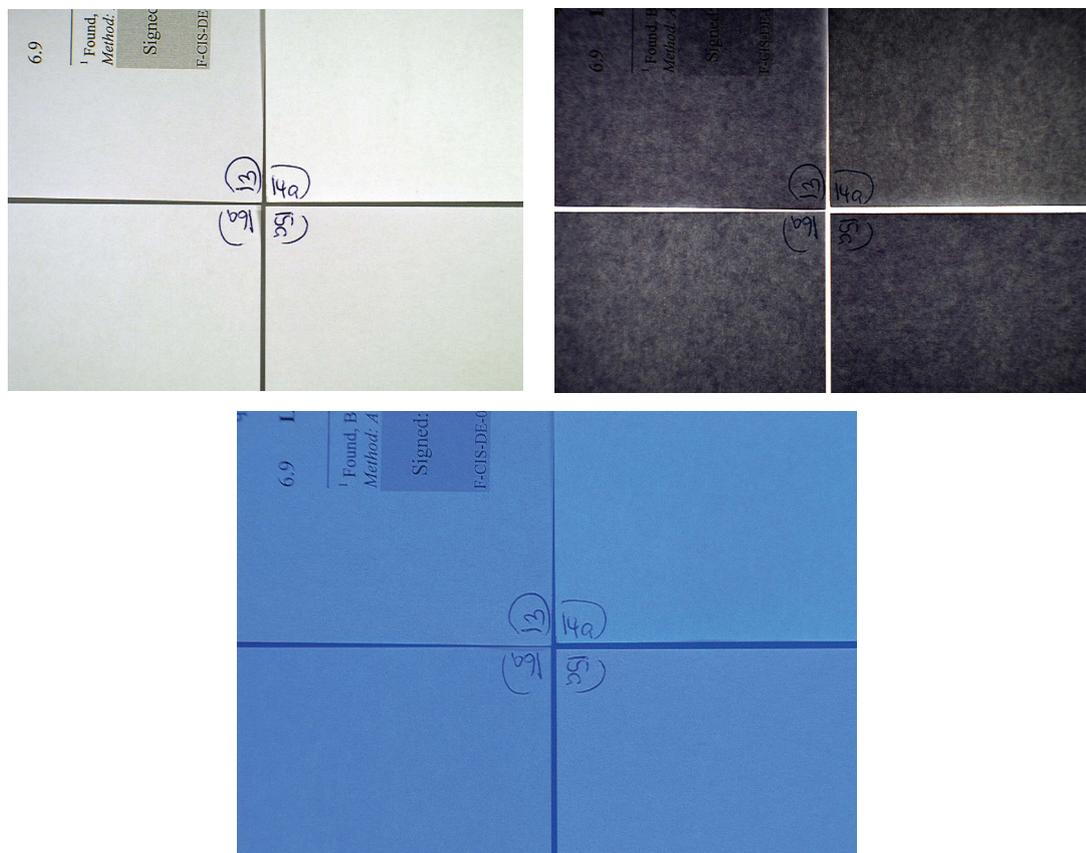


Figure 11.8: Light examination including ambient, transmitted and UV (at 256nm) for trial G

Table 11.13 contains the data that was obtained for the samples:

	Paper 13 (Q)	Paper 14	Paper 15	Paper 16
Grammage (gsm)	79.32	80.65	79.57	79.59
Difference		1.34	0.26	0.28
Apparent Sheet Density (kg/m³)	796.36	766.73	787.91	779.21
Difference		29.62	8.45	35.17
Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev } (\text{‰})$	-27.83 ± 0.18	-23.59 ± 0.25	-27.92 ± 0.04	-26.83 ± 0.05
Difference		4.24	0.09	1.00
Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev } (\text{‰})$	22.67 ± 0.21	26.80 ± 0.15	22.74 ± 0.12	22.39 ± 0.18
Difference		4.13	0.07	0.29

Table 11.13: Results of grammage, apparent sheet density and isotopic measurements for trial G

The interpretation workflow for these results is presented in Table 11.14:

Did the results of any techniques have an observed difference/similarity?	Differences were observed between Paper 13 and Paper 14 using light examination, grammage and both carbon and oxygen isotopes. The differences are large, particularly for the isotopic measurements.
If yes, what technique/s?	No differences were observed between Paper 13 and Paper 15.
If yes, how large was the difference?	Differences were observed between Paper 13 and Paper 15 using apparent sheet density and carbon isotopic abundance values. No other differences were observed.
Are the documents in a comparable form?	Large differences in the isotopic abundance values were observed between paper 13 and paper 14. For paper 16, the predominant difference observed was the carbon isotopic abundance values. The difference was well above the conservative benchmark value and was strengthened by the differences also observed in the apparent sheet density.
	Paper 13 has been toner printed however samples were taken appropriately to account for this.

Is the comparison simple or complex?	Complex, due to mixed IRMS results for paper 16
Are there any other identifiable limitations to the examination that may impact on the interpretation of the results?	No limitations were identified.

Table 11.14: Interpretation considerations for blind trial G

Final Opinion:

Based on the observations and results of the tests conducted, in my opinion, there is strong support for the proposition that questioned paper 13 originated from different sources than the specimen papers 14 and 16, rather than from the same one.

Based on the observations and results of the tests conducted, in my opinion, there is support for the proposition that the questioned paper 13 originated from the same source as the specimen paper 15, or from a source with the same physical, optical, chemical and isotopic properties.

11.3.8. Trial H

In this trial, the examiner was asked to compare seven sheets of paper (papers 17 to 23) to determine if a page substitution had occurred. This may commonly be encountered in casework as a page substitution or alteration to a contract for example. The papers were examined using transmitted light in a single layer, to ensure there was no influence from the papers that may be underneath. For the ambient and UV light techniques, the results were not observed to be different when viewed as single layers or overlaid and as such, the seven pages were stacked and are presented as a single image here.

Page 22 was observed to be different using ambient and UV light but not transmitted light. All other pages were observed to be similar in colour and in their reaction to transmitted and UV light.

These results are shown in Figure 11.9 below.

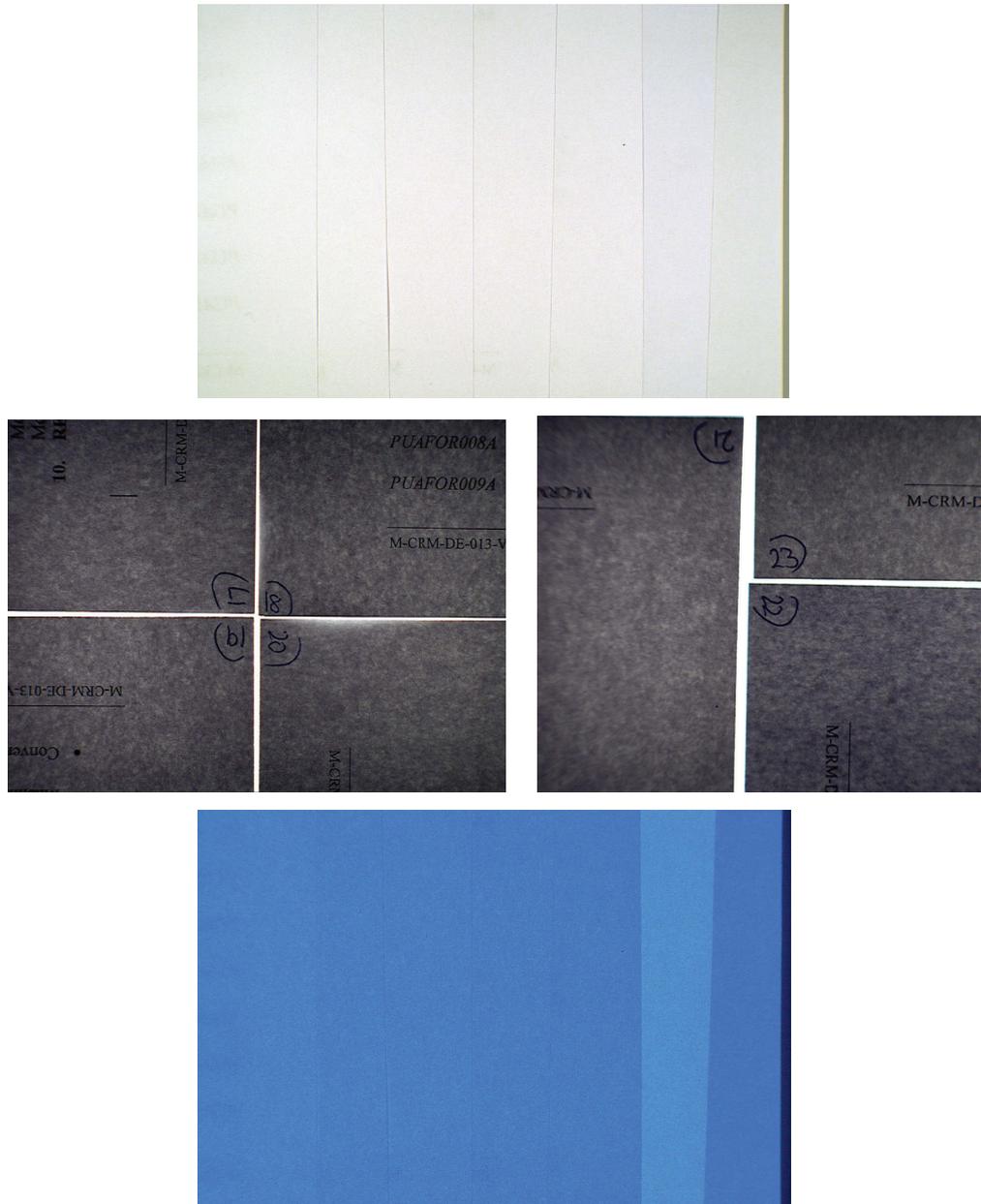


Figure 11.9: Light examination including ambient, transmitted and UV (at 256nm) for trial H

Table 11.15 contains the data obtained for the seven pages, with isotopic data plotted in Figure 11.10:

	Paper 17	Paper 18	Paper 19	Paper 20	Paper 21	Paper 22	Paper 23
Grammage (gsm)	80.55	80.54	80.61	80.78	81.06	79.32	80.91
Apparent Sheet Density (kg/m³)	780.54	784.23	784.15	776.74	777.97	780.02	786.29
Mean $\delta^{13}\text{C}_{\text{VPDB}}$ ± 1 Std Dev (‰)	-20.90 \pm 0.07	-20.87 \pm 0.14	-20.73 \pm 0.24	-20.98 \pm 0.04	-20.48 \pm 0.02	-25.91 \pm 0.07	-20.79 \pm 0.25
Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ ± 1 Std Dev (‰)	27.03 \pm 0.09	27.35 \pm 0.14	27.31 \pm 0.14	27.29 \pm 0.12	27.45 \pm 0.08	22.19 \pm 0.18	27.12 \pm 0.09

Table 11.15: Results of grammage, apparent sheet density and isotopic measurements for trial H

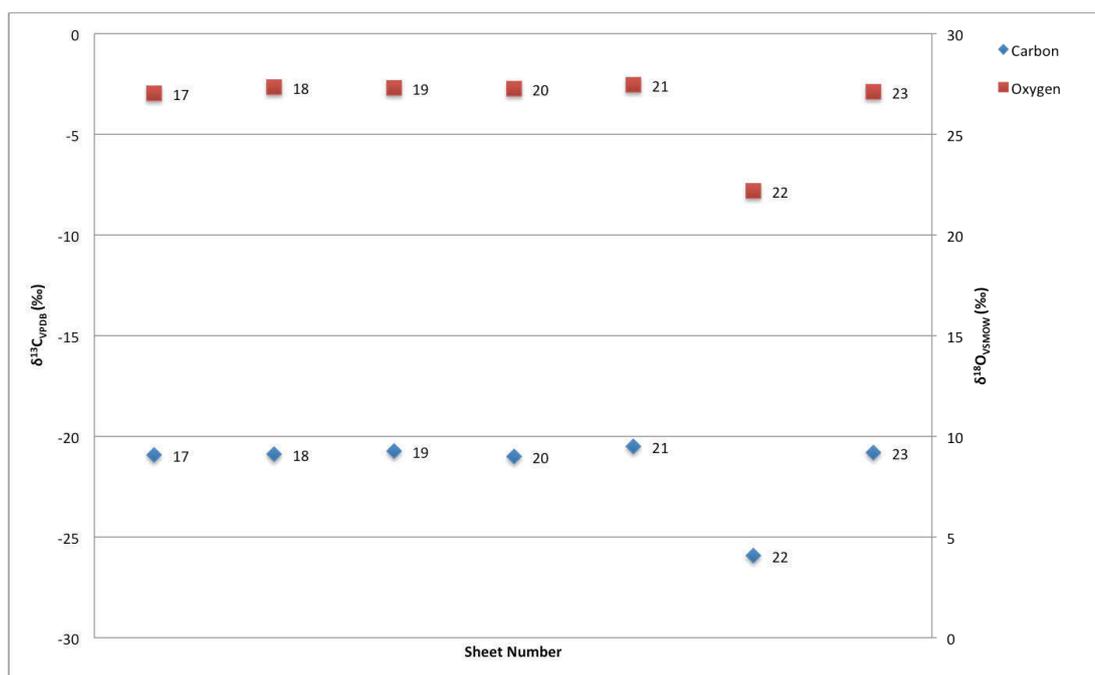


Figure 11.10: Carbon and oxygen isotopic results for seven sheets in blind trial H

The interpretation workflow for these results is presented in Table 11.16:

Factors to be considered	Discussion
<p>Did the results of any techniques have an observed difference/similarity?</p> <p>If yes, what technique/s?</p> <p>If yes, how large was the difference?</p>	<p>Differences were observed between paper 22 and the remaining pages using light examination and IRMS. The differences observed for the isotopic abundance values of paper 22 when compared to the other sheets are significant. This is most easily observed in Figure 11.10.</p> <p>No differences were observed between the light examination techniques, physical measurements and isotopic abundance values for papers 17-21 and paper 23.</p>
<p>Are the documents in a comparable form?</p>	<p>Yes, all documents were toner printed.</p>
<p>Is the comparison simple or complex?</p>	<p>Complex, as multiple pages are required to be compared.</p>
<p>Are there any other identifiable limitations to the examination that may impact on the interpretation of the results?</p>	<p>No limitations were identified.</p>

Table 11.16: Interpretation considerations for blind trial H

Final Opinion:

Based on the observations and results of the tests conducted, in my opinion, there is strong support for the proposition that paper 22 has been substituted from the document and has originated from a different source to the other papers (17-21 and 23).

11.4. Comparison of Trial Findings to Known Results

Table 11.17 contains the results of the blind trials:

	Trial Outcome	Trial Paper Number	Paper Brand	Origin	Correctly called?
A	Papers are different	1	Spilman	Australia	Yes
		2	Victory High White	Indonesia	
B	Papers are different variants of the same brand	3	Post Office Brand	Australia	Unknown
		4	Post Office Economy	Australia	
C	Papers are from same ream	5	Reflex	Australia	Yes
		6	Reflex	Australia	
D	Papers are different	7	Reflex	Australia	Yes
		8	Coles	India	
E	Papers are from the same ream	9	Post Office Economy	Australia	Yes
		10	Post Office Economy	Australia	
F	Papers are from the same brand, but different reams	11	Paper One	Indonesia	Yes
		12	Paper One	Indonesia	
G	Paper 13 has originated from the same ream as Paper 15 and is different to Papers 14 and 16	13	Kodak	Indonesia	Yes
		14	Officemax	South Africa	
		15	Kodak	Indonesia	
		16	Woolworths Essentials	Indonesia	
H	Paper 22 has originated from a different ream and brand to Papers 17-21 and 23.	17-21 and 23	Fuji Xerox Performer	Unknown	Yes
		22	Victory High White	Indonesia	

Table 11.17: Collated results of blind trials A-H

In the majority, the trial results were correctly called. In trial B, two papers that were produced in the same country but packaged as different brands were presented for discrimination. The results obtained for all tests indicated that the papers were similar and hence could not be discriminated. Very simply, this would indicate that the papers have originated from the same source, or a source that produced papers with the same optical, physical and isotopic properties. For Australian papers in particular (which make up 30% of the background population sampled in this project) this is expected and should not be treated as an incorrect result.

Some quality assurance issues were encountered with trial G that resulted in the carbon analytical sequence being repeated. There were a number of reasons for these issues, which may have included:

- Chromatography issues with the instrument due to exhaustion of the internal packings or incomplete combustion;
- Correction issues with the international standards (imprecision of the calibration line) due to chromatography issues; or
- Incorrect benchmark/discrimination value assigned for carbon isotopes, that resulted in a false negative being created.

The initial carbon isotopic values obtained created a type 2 error (a false negative/discrimination) between the questioned paper 13 and the specimen paper 15 in trial G. Given the wording on the inclusion side of the opinion scale - that the findings indicate support either the same or a similar source of the questioned and specimen papers - discriminating papers that have originated from the same source is the most significant error that could occur in casework. Given this, the cause of this error needs to be determined to prevent this from occurring in casework in future.

Through critical evaluation of the analytical sequence, issues with the analytical precision of the first sequence were identified. The international cellulose quality control material utilised was incorrectly scale normalised, indicating issues with the measured values obtained for the international standards used for calibration (polyethylene and sucrose).

The source of the problem was traced to combustion issues with the Flash EA, which were evident in the chromatography and that particularly affected the last sets of polyethylene and sucrose international standards run at the end of the sequence. This sequence highlighted the need to measure and monitor quality assurance materials to ensure that calibration and discrimination issues are not created that result in serious errors during interpretation.

11.5. Conclusions

A series of blind trials were conducted to test whether the paper examination procedure, including light examination, physical and isotopic measurement of papers could be used to correctly discriminate samples. The tests ranged from single sheet comparisons to more complex comparisons with multiple specimen pages and printed questioned sheets.

The examination and interpretation procedure, which included the use of IRMS, was shown to be successful in discriminating paper samples from different sources. Through quality assurance checks, an analytical sequence was repeated before results were reported, demonstrating the importance of monitoring in producing quality results.

Overall the techniques, examination protocol, interpretation and reporting frameworks proposed for the forensic comparison and discrimination of document papers was observed to be precise, robust and fit for purpose in casework.

Chapter 12

12. Conclusions and Future Work

This project has shown that IRMS is a valuable tool in the examination and comparison of document papers. By measuring the carbon and oxygen isotopic values of a background population of papers collected from Australia and New Zealand, variability between paper sources was demonstrated and a combined discrimination of 89% of samples was achieved.

The analytical methods have been developed, optimized and validated to international quality control standards so that they can be utilised as part of casework in a laboratory accredited under ISO17025 (International Organisation for Standardisation, 2005).

The carbon method validation was shown to produce robust and reproducible results with a measurement uncertainty of 0.26 ‰, well within the discrimination benchmark value defined. The oxygen method, while accurate within an analytical sequence, did suffer from some reproducibility issues that will require further work to rectify. Overall however, the oxygen method produced results were robust when quality controls were closely monitored.

For both the carbon and oxygen isotopes of paper, the within ream homogeneity of paper was examined, first by looking to determine the natural variability of seven different brands. Experiments on single reams, using a number of different sampling techniques, were valuable in highlighting a significant risk was posed from the relative inhomogeneity of a single ream of paper. This was most clearly demonstrated when discrimination could be called between different sampling groups, using a 95% population discrimination. Use of discrimination at this level would mean that the technique would only be of use in casework for associating pieces of paper from the same sheet e.g. in cases where the piece of paper had been torn.

To ensure that discrimination within the same ream is avoided, benchmark values were determined for both isotopes, based on the largest range of variation observed for a single ream from seven brands. The proposed benchmark values were tested during subsequent intra-ream variation studies where 7 sheets from 7 reams and 7 brands were measured to determine the variability of reams produced from the same brand at the same time (i.e. the batch variability).

These benchmark values, 0.5 ‰ for carbon and 0.7 ‰ for oxygen, were found to be robust in avoiding discrimination of samples from the same source whilst still providing sufficient discrimination in the broader population i.e. from between different brands. Some effect was observed for the oxygen isotopes with respect to comparison of values across different analytical runs and an extended benchmark value of 1 ‰ was defined for samples that needed to be compared between analytical sequences.

Further work has been conducted to ensure that IRMS was ready for casework application and interpretation. These experiments aimed to determine the effect of usage and forensic testing on the isotopic abundance values measured. Toner printed documents were found to pose a risk to comparability, as did documents that had been heavily handled or exposed to charring. Forensic testing, including DNA tapelifting and fingerprint treatment, did not have an effect on the carbon isotopic values measured but did shift the oxygen values.

A short scoping study for hydrogen isotopes was conducted which, for the limited sample set selected, was not observed to add discrimination beyond that achieved using carbon and oxygen.

Other, traditional paper examination techniques were used on a sub set of 60 papers from the background population, to extend the work beyond just IRMS examinations. While some issues were encountered for subjective discriminations such as those required in the light examinations, discrimination at a lower level (between 25 and 50%) was observed for the sample set for each examination type.

Finally, an examination protocol, based on a combination of techniques was constructed with an interpretation and reporting framework to sit alongside the analytical results obtained.

A range of future work has been identified from this project. These recommendations include:

1. Based on the results seen here, for other analytical techniques that are attempting to measure and compare document papers using elemental abundance techniques, the homogeneity of the 'source' material needs to be explored in detail. This is especially true of the work being undertaken using techniques such as laser ablation inductively coupled plasma mass spectrometry and laser induced breakdown spectroscopy. Additional to this, work also needs to be undertaken to examine the effects of usage and handling on the results of these techniques, particularly when claiming that the techniques are useful for forensic casework.
2. New calibration materials are required for the correction of organic oxygen isotopic values. This will assist in the reduction of reproducibility issues between sequences for oxygen measurements.
3. Further work, as required based on forensic casework, is needed to understand the effects of usage on the isotopic abundance values of papers. This could be done proactively, or as new situations arise in casework.
4. Additional work should be undertaken to further develop and explore the use of hydrogen isotopes for the comparison of document papers including an extended background population and homogeneity experiments.
5. The examination protocol and interpretation framework should be considered for development into a Bayesian network to provide a numerical estimate of the strength of the results rather than the subjective interpretation presented here.

13. References

2001. Robust statistics: a method of coping with outliers. *Analytical Methods Committee Technical Brief*, No 6.
2007. Conflo IV Operating Manual http://es.ucsc.edu/~silab/ThermoManuals/Conflo/1224730_ConFlo_IV_OM_Rev_A_NV.pdf: Thermo Fisher Scientific.
2010. *Inter-Laboratory Comparison Results (ILC-6)* [Online]. http://www.forensic-isotopes.org/assets/FIRMS_ILC_6.pdf: Forensic Isotope Ratio Mass Spectrometry Available: http://www.forensic-isotopes.org/assets/FIRMS_ILC_6.pdf [Accessed 29 March 2013].
- ADAMS, J. 2011. Analysis of printing and writing papers by using direct analysis in real time mass spectrometry. *International Journal of Mass Spectrometry*, 301, 109-126.
- AL-HOSNEY, H. & GRASSIAN, V. 2005. Water, sulfur dioxide and nitric acid adsorption on calcium carbonate: A transmission and ATR-FTIR study. *Phys. Chem. Chem. Phys.*, 7, 1266-1276.
- ANDRASKO, J. 1996. Microreflectance FTIR techniques applied to materials encountered in forensic examination of documents. *Journal of Forensic Sciences*, 41, 812-823.
- ARMSTRONG, N. & HIBBERT, D. 2009. An introduction to Bayesian methods for analyzing chemistry data Part 1: An introduction to Bayesian theory and methods. *Chemometrics and Intelligent Laboratory Systems*.
- AUSTRALIA, S. 1998. Methods of test for pulp and paper - Standard atmosphere for testing paper and board and procedure for monitoring the atmosphere. <http://infostore.saiglobal.com/store/details.aspx?ProductID=357120>.
- AUSTRALIA, S. 2004. *1301.405s-2004 Methods of test for pulp and paper - Grammage of non-creped paper and board* [Online]. <http://infostore.saiglobal.com/store/Details.aspx?ProductID=357077>. [Accessed 29 March 2013].
- AUSTRALIA, S. 2013a. AS 5388.3-2013 Forensic Analysis - Reporting. SAI Global.
- AUSTRALIA, S. 2013b. AS 5388.3-2103 Forensic Analysis-Interpretation. SAI Global.
- BAECHLER, S., RIBAUX, O. & MARGOT, P. 2012. 2012 Student Paper: Toward a Novel Forensic Intelligence Model: Systematic Profiling of False Identity Documents. *Forensic Science Policy & Management: An International Journal*, 3, 70-84.
- BARNES, C. J. & ALLISON, G. B. 1983. The distribution of deuterium and ^{18}O in dry soils : 1. Theory. *Journal of Hydrology*, 60, 141-156.
- BARRIE, A., DAVIES, J., PARK, A. & WORKMAN, C. 1989. Continuous-flow stable isotope analysis for biologists. *Spectroscopy*, 4, 42-52.
- BARROS, V. G., OLIVEIRA, T. M. N., SPANDRE, R., ZUPPI, G. M., RAPAGLIA, J. & VAZ, C. 2013. Lead, nitrogen and carbon stable isotopes in the sediments of Babitonga Bay, Brazil: An oil spill case. *Journal of Environmental Hydrology*, 21, 1-10.
- BELL, S. E. J., STEWART, S. P., HO, Y. C., CRAYTHORNE, B. W. & SPEERS, S. J. 2013. Comparison of the discriminating power of Raman and surface-enhanced Raman spectroscopy with established techniques for the examination of liquid and gel inks. *Journal of Raman Spectroscopy*, 44, 509-517.
-

- BENSON, S. 2009. Introduction of Isotope Ratio Mass Spectrometry (IRMS) for the Forensic Analysis of Explosives. *Thesis - University of Technology Sydney*.
- BENSON, S., LENNARD, C., MAYNARD, P. & ROUX, C. 2006. Forensic Applications of Isotope Ratio Mass Spectrometry—A Review. *Forensic Science International*, 157, 1-22.
- BENSON, S. J., LENNARD, C. J., HILL, D. M., MAYNARD, P. & ROUX, C. 2010. Forensic Analysis of Explosives Using Isotope Ratio Mass Spectrometry (IRMS), Part 1: Instrument Validation of the DELTAplusXP IRMS for Bulk Nitrogen Isotope Ratio Measurements. *Journal of Forensic Sciences*, 55, 193-204.
- BERGER, C. 2009. Objective paper structure comparison through processing of transmitted light images. *Forensic Science International*, 192, 1-6.
- BERGER, C. E. H. & RAMOS, D. 2012. Objective paper structure comparison: Assessing comparison algorithms. *Forensic Science International*.
- BIEDERMANN, A., TARONI, F., BOZZA, S. & MAZZELLA, W. D. 2011. Implementing statistical learning methods through Bayesian networks (Part 2): Bayesian evaluations for results of black toner analyses in forensic document examination. *Forensic Science International*, 204, 58-66.
- BIEDERMANN, A., VOISARD, R. & TARONI, F. 2012. Learning about Bayesian networks for forensic interpretation: An example based on the 'the problem of multiple propositions'. *Science & Justice*, 52, 191-198.
- BIERMANN, C. 1996. *Handbook of pulping and papermaking*, Academic Pr.
- BLANCHARD, D., HARRISON, S., SCIENCES, A. A. O. F. & STATES, U. 1978. Trace elemental profiles and ratios determined by instrumental neutron activation analysis for fine paper identification. *J Forensic Sci*, 23, 679-86.
- BODZIAK, W. J. 1998. Edge Characteristics of Commercially Produced Paper Stock. *Journal of the American Society of Questioned Document Examiners*, 1, 57-66.
- BOWEN, G. 2010. Isoscapes: Spatial Pattern in Isotopic Biogeochemistry. *Annual Review of Earth and Planetary Sciences*, 38, 161-187.
- BOWEN, G. J., CHESSON, L., NIELSON, K., CERLING, T. E. & EHLERINGER, J. R. 2005. Treatment methods for the determination of $\delta^{2}\text{H}$ and $\delta^{18}\text{O}$ of hair keratin by continuous-flow isotope-ratio mass spectrometry. *Rapid Communications in Mass Spectrometry*, 19, 2371-2378.
- BOWEN, G. J., LIU, Z., VANDER ZANDEN, H. B., ZHAO, L. & TAKAHASHI, G. 2014. Geographic assignment with stable isotopes in IsoMAP. *Methods in Ecology & Evolution*, 5, 201-206.
- BRAND, W. A. 1996. High precision isotope ratio monitoring techniques in mass spectrometry. *Journal of Mass Spectrometry*, 31, 225-235.
- BRIEF, A. M. C. T. 2005. *The ZL score - combining your proficiency test results with your own fitness-for-purpose criterion* [Online]. Available: http://www.rsc.org/images/zL-score-proficiency-fitness-criterion-technical-brief-2_tcm18-214844.pdf [Accessed 29 March 2013].
- BRUNELLE, R. & REED, R. 1984. *Forensic examination of ink and paper*, Charles C Thomas.
- BUCKLETON, J. S., TRIGGS, C. M. & CHAMPOD, C. 2006. An extended likelihood ratio framework for interpreting evidence. *Science & Justice*, 46, 69-78.
- CAIN, S. 1984. Laser and Fiber Optic Photographic Analysis of Single-Edge Paper Striations. *Journal of Forensic Sciences*, 29, 9.
- CAMPBELL, N., REECE, J. & MITCHELL, L. 2002. *Biology Jakarta*. Erlangga.
- CAPPA, C. D., HENDRICKS, M. B., DEPAOLO, D. J. & COHEN, R. C. 2003. Isotopic fractionation of water during evaporation. *Journal of Geophysical Research*, 108, 4525.

-
- CARTER, J., HILL, J., DOYLE, S. & LOCK, C. 2009. Results of four inter-laboratory comparisons provided by the Forensic Isotope Ratio Mass Spectrometry (FIRMS) network. *Science & Justice*, 49, 127-137.
- CARTER, J., SLEEMAN, R., HILL, J., IDOINE, F. & TITTERTON, E. 2005. Isotope Ratio Mass Spectrometry as a Tool for Forensic Investigation (examples from recent studies). *Science & Justice*, 45, 141-149.
- CARTER, J. F., BARWICK, V.J. (EDS) 2011. Good Practice Guide for Isotope Ratio Mass Spectrometry. In: FIRMS (ed.). http://www.forensic-isotopes.org/assets/IRMS_Guide_Finalv3.1_Web.pdf.
- CARTER, J. F., DOYLE, S., PHASUMANE, B.-L. & NICDAEID, N. 2014. The role of isotope ratio mass spectrometry as a tool for the comparison of physical evidence. *Science & Justice*.
- CARTER, J. F. & FRY, B. 2013. Ensuring the reliability of stable isotope ratio data - Beyond the principle of identical treatment. *Analytical and Bioanalytical Chemistry*, 405, 2799-2814.
- CASALE, J., EHLERINGER, J., MORELLO, D. R. & LOTT, M. J. 2005. Isotopic Fractionation of Carbon and Nitrogen During the Illicit Processing of Cocaine and Heroin in South America. *Journal of Forensic Sciences*, 50.
- CAUSIN, V., CASAMASSIMA, R., MARRUNCHEDDU, G., LENZONI, G., PELUSO, G. & RIPANI, L. 2011. The discrimination potential of diffuse-reflectance ultraviolet-visible-near infrared spectrophotometry for the forensic analysis of paper. *Forensic Science International*.
- CAUSIN, V., MAREGA, C., MARIGO, A., CASAMASSIMA, R., PELUSO, G. & RIPANI, L. 2010. Forensic differentiation of paper by X-ray diffraction and infrared spectroscopy. *Forensic Science International*, 197, 70-74.
- CERLING, T., HARRIS, J., MACFADDEN, B., LEAKEY, M., QUADE, J., EISENMANN, V. & EHLERINGER, J. 1997. Global vegetation change through the Miocene/Pliocene boundary. *Nature*, 389, 153-158.
- CHESSON, L. A., PODLESIAK, D.W., CERLING, T. E., EHLERINGER, J. R. 2009. Evaluating uncertainty in the calculation of non-exchangable hydrogen fractions within organic materials. *Rapid Commun. Mass Spectrom*, 23, 1275-1280.
- CHESSON, L. A., TIPPLE, B. J., ERKKILA, B. R., CERLING, T. E. & EHLERINGER, J. R. 2010a. B-HIVE: Beeswax hydrogen isotopes as validation of environment. Part I: Bulk honey and honeycomb stable isotope analysis. *Food Chemistry*.
- CHESSON, L. A., VALENZUELA, L. O., BOWEN, G. J., CERLING, T. E. & EHLERINGER, J. R. 2011. Consistent predictable patterns in the hydrogen and oxygen stable isotope ratios of animal proteins consumed by modern humans in the USA. *Rapid Communications in Mass Spectrometry*, 25, 3713-3722.
- CHESSON, L. A., VALENZUELA, L. O., O'GRADY, S. P., CERLING, T. E. & EHLERINGER, J. R. 2010b. Hydrogen and Oxygen Stable Isotope Ratios of Milk in the United States. *Journal of agricultural and food chemistry*, 58, 2358-2363.
- COPLEN, T., HOPPLE, J., BOEHIKE, J., PEISER, H. & RIEDER, S. 2002. Compilation of minimum and maximum isotope ratios of selected elements in naturally occurring terrestrial materials and reagents. United States Geological Survey.
- COPLEN, T. B. 2011. Guidelines and recommended terms for expression of stable-isotope ratio and gas-ratio measurement results. *Rapid Communications in Mass Spectrometry*, 25, 2538-2560.
-

-
- COPLEN, T. B., BRAND, W. A., GEHRE, M., GRÖNING, M., MEIJER, H. A., TOMAN, B. & VERKOUTEREN, R. M. 2006. New Guidelines for $\delta^{13}\text{C}$ Measurements. *Analytical Chemistry*, 78, 2439-2441.
- COPLEN, T. B. & QI, H. 2012. USGS42 and USGS43: Human-hair stable hydrogen and oxygen isotopic reference materials and analytical methods for forensic science and implications for published measurement results. *Forensic Science International*, 214, 135-141.
- CRAIG, H. 1957. Isotopic standards for carbon and oxygen and correction factors for mass-spectrometric analysis of carbon dioxide. *Geochimica et cosmochimica acta*, 12, 133-149.
- CRAIG, H. 1961. Isotopic Variations in Meteoric Waters. *Science*, 133, 1702.
- CRISS, R. 1999. *Principles of stable isotope distribution*, Oxford University Press, USA.
- CROCKETT, T., WEBB, A., BORCHARDT, L. & EASTY, D. 1987. Identification of Polymers in Paper by Pyrolysis Gas Chromatography - Reduction of Paper Interference by Prepyrolysis. *Journal of Chromatography*, 407, 330-339.
- DAEID, N. N., BUCHANAN, H. A. S., SAVAGE, K. A., FRASER, J. G. & CRESSWELL, S. L. 2010. Recent advances in the application of stable isotope ratio analysis in forensic chemistry. *Australian Journal of Chemistry*, 63, 3-7.
- DAÉID, N. N., MEIER-AUGENSTEIN, W. & KEMP, H. F. 2011. Investigating the provenance of un-dyed spun cotton fibre using multi-isotope profiles and chemometric analysis. *Rapid Communications in Mass Spectrometry*, 25, 1812-1816.
- DAVIS, L. J., SAUNDERS, C. P., HEPLER, A. & BUSCAGLIA, J. 2012. Using subsampling to estimate the strength of handwriting evidence via score-based likelihood ratios. *Forensic Science International*, 216, 146-157.
- DAWSON, T. 1993. Hydraulic lift and water use by plants: implications for water balance, performance and plant-plant interactions. *Oecologia*, 95, 565-574.
- DAWSON, T. 1998. Fog in the California redwood forest: ecosystem inputs and use by plants. *Oecologia*, 117, 476-485.
- DAWSON, T. & EHLERINGER, J. 1991. Streamside trees that do not use stream water. *Nature*, 350, 335-337.
- DAWSON, T. & SIEGWOLF, R. 2007. *Stable isotopes as indicators of ecological change*, Academic Press.
- DE GROOT, P. 2004. *Handbook of Stable Isotope Analytical Techniques: Volume 1*, Elsevier Science.
- DE GROOT, P. 2008. *Handbook of Stable Isotope Analytical Techniques Vol II*, Elsevier Science Ltd.
- DEINES, P. 1980. In: Fritz and Fontes, Handbook of Environmental Isotope Geochemistry Vol 1. The Terrestrial Environment. *Elsevier*.
- DENTON, T., SCHMIDT, S., CRITCHLEY, C. & STEWART, G. 2001. Natural abundance of stable carbon and nitrogen isotopes in Cannabis sativa reflects growth conditions. *Aust. J. Plant Physiol*, 28, 1005-1012.
- DESAGE, M., GUILLUY, R. & BRAZIER, H. 1991. Gas chromatography with mass spectrometry or isotope-ratio mass spectrometry in studying the geographical origin of heroin. *Analytica Chimica Acta*, 247, 249-254.
- EHLERINGER, J. 1993. Variation in leaf carbon isotope discrimination in Encelia farinosa: implications for growth, competition, and drought survival. *Oecologia*, 95, 340-346.
- EHLERINGER, J., CASALE, J., LOTT, M. & FORD, V. 2000. Tracing the geographical origin of cocaine. *Nature*, 408, 311-312.
-

- EHLERINGER, J., COOPER, D., LOTT, M. & COOK, C. 1999. Geo-location of heroin and cocaine by stable isotope ratios. *Forensic Science International*, 106, 27-35.
- EHLINGER, J. & DAWSON, T. 1992. Water uptake by plants: perspectives from stable isotope composition. *Plant, Cell and Environment*, 15, 1073-1082.
- ENFSI 2015. ENFSI Guideline for Evaluative Reporting in Forensic Science - Strengthening the Evaluation of Forensic Results across Europe. http://www.enfsi.eu/sites/default/files/afbeeldingen/enfsi_booklet_m1.pdf.
- EURACHEM 1998. The Fitness for Purpose of Analytical Methods; A Laboratory Guide to Method Validation and Related Topics. <http://www.eurachem.org/images/stories/Guides/pdf/valid.pdf>.
- EVETT, I., JACKSON, G., LAMBERT, J. & MCCROSSAN, S. 2000. The impact of the principles of evidence interpretation on the structure and content of statements. *Science & justice: journal of the Forensic Science Society*, 40, 233.
- FARMER, N., CURRAN, J., LUCY, D., NIC DAEID, N. & MEIER-AUGENSTEIN, W. 2009a. Stable isotope profiling of burnt wooden safety matches. *Science & Justice*, 49, 107-113.
- FARMER, N., MEIER-AUGENSTEIN, W. & KALIN, R. 2005. Stable isotope analysis of safety matches using isotope ratio mass spectrometry-a forensic case study. *Rapid Communications in mass spectrometry*, 19, 3182-3186.
- FARMER, N., MEIER-AUGENSTEIN, W. & LUCY, D. 2009b. Stable isotope analysis of white paints and likelihood ratios. *Science & Justice*, 49, 114-119.
- FARQUHAR, G., BARBOUR, M. & HENRY, B. 1998. Interpretation of oxygen isotope composition of leaf material. *Stable Isotopes: Integration of Biological, Ecological, and Geochemical Processes*, 27-48.
- FARQUHAR, G., EHLERINGER, J. & HUBICK, K. 1989. Carbon isotope discrimination and photosynthesis. *Annual Review of Plant Biology*, 40, 503-537.
- FARQUHAR, G., O'LEARY, M. & BERRY, J. 1982. On the relationship between carbon isotope discrimination and the intercellular carbon dioxide concentration in leaves. *Functional Plant Biology*, 9, 121-137.
- FONER, H. & ADAN, N. 1983. The characterization of papers by X-ray diffraction (XRD): measurement of cellulose crystallinity and determination of mineral composition. *Journal of the Forensic Science Society*, 23, 313-321.
- FRANCEY, R. & FARQUHAR, G. 1982. An explanation of $^{13}\text{C}/^{12}\text{C}$ variations in tree rings.
- FRY, B. 2006. *Stable isotope ecology*, Springer Verlag.
- GALIMOV, E., SEVASTYANOV, V., KULBACHEVSKAYA, E. & GOLYAVIN, A. 2005. Isotope ratio mass spectrometry: d^{13}C and d^{15}N analysis for tracing the origin of illicit drugs. *Rapid Communications in mass spectrometry*, 19, 1213-1216.
- GAT, J., MOOK, W. & MEIJER, H. 2001. Environmental isotopes in the hydrological cycle, principles and applications. Vol. II, Atmospheric water. UNESCO/IAEA.
- GAUDINSKI, J. B., DAWSON, T. E., QUIDEAU, S., SCHUUR, E. A. G., RODEN, J. S., TRUMBORE, S. E., SANDQUIST, D. R., OH, S.-W. & WASYLISHEN, R. E. 2005. Comparative Analysis of Cellulose Preparation Techniques for Use with ^{13}C , ^{14}C , and ^{18}O Isotopic Measurements. *Analytical Chemistry*, 77, 7212-7224.
- GEHRE, M. & STRAUCH, G. 2003. High-temperature elemental analysis and pyrolysis techniques for stable isotope analysis. *Rapid Communications in mass spectrometry*, 17, 1497-1503.
- GENTILE, N., SIEGWOLF, R. T., ESSEIVA, P., DOYLE, S., ZOLLINGER, K. & DELÉMONT, O. 2015. Isotope ratio mass spectrometry as a tool for source

- inference in forensic science: A critical review. *Forensic science international*, 251, 139-158.
- GREEN, J. A. 2012. Reliability of Paper Brightness in Authenticating Documents. *Journal of Forensic Sciences*, 57, 1003-1007.
- HÄKKÄNEN, H., HOUNI, J., KASKI, S. & KORPPI-TOMMOLA, J. 2001. Analysis of paper by laser-induced plasma spectroscopy. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 56, 737-742.
- HARRIS, D. 2003. *Quantitative chemical analysis*, WH Freeman.
- HELLIKER, B. & EHLERINGER, J. 2002. Grass blades as tree rings: environmentally induced changes in the oxygen isotope ratio of cellulose along the length of grass blades. *New Phytologist*, 155, 417-424.
- HEPLER, A. B., SAUNDERS, C. P., DAVIS, L. J. & BUSCAGLIA, J. 2012. Score-based likelihood ratios for handwriting evidence. *Forensic Science International*, 219, 129-140.
- HIBBERT, D. & ARMSTRONG, N. 2009. An introduction to Bayesian methods for analyzing chemistry data:: Part II: A review of applications of Bayesian methods in chemistry. *Chemometrics and Intelligent Laboratory Systems*, 97, 211-220.
- HIBBERT, D. & GOODING, J. 2006. *Data analysis for chemistry*, Oxford Univ. Press.
- HILL, S., WATERHOUSE, J., FIELD, E., SWITSUR, V. & AP REES, T. 1995. Rapid recycling of triose phosphates in oak stem tissue. *Plant, Cell & Environment*, 18, 931-936.
- HOEFS, J. 2008. *Stable isotope geochemistry*, Springer.
- HOLLAND, N. 2004. Wire Mark Analysis Using Fast Fourier Transform Processing Techniques in Paper Identification Cases. *Presentation - Australian and New Zealand Forensic Science Society Conference*.
- HURLEY, J., WEST, J. & EHLERINGER, J. 2010a. Stable isotope models to predict geographic origin and cultivation conditions of marijuana. *Science & justice: journal of the Forensic Science Society*, 50, 86.
- HURLEY, J., WEST, J. & EHLERINGER, J. 2010b. Tracing retail cannabis in the United States: Geographic origin and cultivation patterns. *International Journal of Drug Policy*, 21, 222-228.
- IHLE, E. & SCHMIDT, H. 1996. Multielement isotope analysis on drugs of abuse. Possibility for their origin assignment. *Isotopes in Environmental and Health Studies*, 32, 226-228.
- INTERNATIONAL ORGANISATION FOR STANDARDISATION. 2005. *ISO/IEC 17025:2005* [Online]. Available: http://www.iso.org/iso/iso_catalogue/catalogue_tc/caatalogue_detail.htm?csnumber=39883 [Accessed].
- ISO 2005. ISO/IEC 17025:2005 General requirements for the competence of testing and calibration laboratories.
- ISO 2009. ISO Guide 34:2009 General Requirements for the competence of reference materials producers.
- JOINT COMMITTEE FOR GUIDES IN METROLOGY. 2008. *GUM: Guide to the Expression of Uncertainty in Measurement* [Online]. Available: <http://www.bipm.org/en/publications/guides/gum.html> [Accessed].
- JONES, K., BENSON, S. & ROUX, C. 2013a. The forensic analysis of office paper using carbon isotope ratio mass spectrometry , Part 1: Understanding the background population and homogeneity of paper for the comparison and discrimination of samples. *Forensic Science International*, 231, 354-363.
- JONES, K., BENSON, S. & ROUX, C. 2013b. The forensic analysis of office paper using carbon isotope ratio mass spectrometry. Part 3: Characterizing the source

- materials and the effect of production and usage on the $\delta^{13}\text{C}$ values of paper. *Forensic Science International*, 233, 355-364.
- JONES, K., BENSON, S. & ROUX, C. 2013c. The forensic analysis of office paper using carbon isotope ratio mass spectrometry, Part 2: Method development, validation and sample handling. *Forensic Science International*, 231, 364-374.
- JONES, K., BENSON, S. & ROUX, C. 2016. The forensic analysis of office paper using oxygen isotope ratio mass spectrometry. Part 1: Understanding the background population and homogeneity of paper for the comparison and discrimination of samples. *Forensic Science International*, 262, 97-107.
- KEELING, C. D., MOOK, W. I. M. G. & TANS, P. P. 1979. Recent trends in the $^{13}\text{C}/^{12}\text{C}$ ratio of atmospheric carbon dioxide.
- KEELING, C. D., PIPER, S. C., BACASTOW, R. B., WAHLEN, M., WHORF, T. P., HEIMANN, M. , AND MEIJER, H. A. 2001. Exchanges of atmospheric CO_2 and $^{13}\text{CO}_2$ with the terrestrial biosphere and oceans from 1978 to 2000. I. Global aspects. *SIO Reference Series, Scripps Institution of Oceanography*, No. 01-06.
- KIPPAN, H. 2001. *Handbook of print media*, Springer Heidelberg.
- KUMAR, R. 2011. Evaluation of Two Instrumental Methods of Comparing Writing Paper. *Journal of Forensic Sciences*, 56, 514-517.
- LAJTHA, K. & MICHENER, R. 1994. *Stable isotopes in ecology and environmental science*, Wiley-Blackwell.
- LAPORTE, G. M., ARREDONDO, M. D., MCCONNELL, T. S., STEPHENS, J. C., CANTU, A. A. & SHAFFER, D. K. 2006. An evaluation of matching unknown writing inks with the United States International Ink Library. *Journal of Forensic Sciences*, 51, 689-692.
- LEAVITT, S. & LONG, A. 1986. Stable-carbon isotope variability in tree foliage and wood. *Ecology*, 67, 1002-1010.
- LEAVITT, S. W. 2010. Tree-ring C-H-O isotope variability and sampling. *Science of The Total Environment*, 408, 5244-5253.
- LENNARD, C., EL-DEFTAR, M. M. & ROBERTSON, J. 2015. Forensic application of laser-induced breakdown spectroscopy for the discrimination of questioned documents. *Forensic Science International*, 254, 68-79.
- LERMAN, J., DELEENS, E., NATO, A. & MOYSE, A. 1974. Variation in the carbon isotope composition of a plant with Crassulacean acid metabolism. *Plant Physiology*, 53, 581.
- LICHTFOUSE, E. 2000. Compound specific isotope analysis. Application to archaeology, biomedical sciences, biosynthesis, environment, extraterrestrial chemistry, food science, forensic science, humic substances, microbiology, organic geochemistry, soil science and sport. *Rapid Communications in Mass Spectrometry*, 14, 1337-1344.
- LLOYD, R. 1966. Oxygen isotope enrichment of sea water by evaporation. *Geochimica et Cosmochimica Acta*, 30, 801-814.
- MANSO, M., COSTA, M. & CARVALHO, M. 2007. From papyrus to paper: Elemental characterization by X-ray fluorescence spectrometry. *Nuclear Inst. and Methods in Physics Research, A*, 580, 732-734.
- MANSO, M., COSTA, M. & CARVALHO, M. 2008. X-ray fluorescence spectrometry on paper characterization: A case study on XVIII and XIX century documents. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 63, 1320-1323.
- MARCHAND, P. 1989. A Non-Destructive Method for Determining the Grain Direction of Paper. *Canadian Society of Forensic Sciences*, 22, 69-81.
- MARGOT, P. 2011. Forensic science on trial - What is the law of the land? *Australian Journal of Forensic Sciences*, 43, 89-103.

- MARTIRE, K. A., KEMP, R. I., WATKINS, I., SAYLE, M. A. & NEWELL, B. R. 2013. The expression and interpretation of uncertain forensic science evidence: verbal equivalence, evidence strength, and the weak evidence effect. *Law and human behavior*, 37, 197.
- MATHEWS, C., VAN HOLDE, K. & AHERN, K. 2000. *Biochemistry* San Francisco: Benjamin Cummings.
- MAZANY, T., LERMAN, J. & LONG, A. 1980. Carbon-13 in tree-ring cellulose as an indicator of past climates.
- MCCARROLL, D. & LOADER, N. J. 2004. Stable isotopes in tree rings. *Quaternary Science Reviews*, 23, 771-801.
- MCGAW, E., SZYMANSKI, D. & SMITH, R. 2009a. Characterization of undigested particulate material following microwave digestion of recycled document papers. *Journal of Forensic Sciences*, 54, 1171-1175.
- MCGAW, E., SZYMANSKI, D., SMITH, R. & LANSING, E. 2009b. Determination of Trace Elemental Concentrations in Document Papers for Forensic Comparison Using Inductively Coupled Plasma–Mass Spectrometry. *Journal of Forensic Sciences*, 54, 1163-1170.
- MEIER-AUGENSTEIN, W. 2010. *Stable Isotope Forensics: An Introduction to the Forensic Application of Stable Isotope Analysis*, John Wiley & Sons.
- MEIER-AUGENSTEIN, W., CHARTRAND, M. M. G., KEMP, H. F. & ST JEAN, G. 2011. An interlaboratory comparative study into sample preparation for both reproducible and repeatable forensic 2H isotope analysis of human hair by continuous flow isotope ratio mass spectrometry. *Rapid Communications in Mass Spectrometry*, 25, 3331-3338.
- MEIER-AUGENSTEIN, W., KEMP, H. F., SCHENK, E. R. & ALMIRALL, J. R. 2014. Discrimination of unprocessed cotton on the basis of geographic origin using multi-element stable isotope signatures. *Rapid Communications in Mass Spectrometry*, 28, 545-552.
- MIYATA, H., SHINOZAKI, M., NAKAYAMA, T. & ENOMAE, T. 2002. A discrimination method for paper by Fourier transform and cross correlation. *Journal of Forensic Sciences*, 47, 1125-1132.
- MONAHAN, F. J., MOLONEY, A. P., OSORIO, M. T., RØHRLE, F. T., SCHMIDT, O. & BRENNAN, L. 2012. Authentication of grass-fed beef using bovine muscle, hair or urine. *Trends in Food Science & Technology*, 28, 69-76.
- MUCCIO, Z., WÖCKEL, C., AN, Y. & JACKSON, G. P. 2012. Comparison of Bulk and Compound-Specific $\delta^{13}\text{C}$ Isotope Ratio Analyses for the Discrimination Between Cannabis Samples. *Journal of Forensic Sciences*, 57, 757-764.
- MURPHY, J. P. 2009. Paper Analysis. *Wiley Encyclopedia of Forensic Science*.
- NATIONAL ASSOCIATION OF TESTING AUTHORITIES. 2015. *ISO/IEC 17025 Field Application Document Forensic Science - Supplementary Requirements for Accreditation* [Online]. Available: http://www.nata.com.au/nata/phocadownload/publications/Accreditation_criteria/ISO-IEC-17025/Forensic/Forensic-Science-Application-Document.pdf [Accessed 15 July 2015].
- NATIONAL ASSOCIATION OF TESTING AUTHORITIES. April 2009a. *Technical Note 17 - Guidelines for the validation and verification of chemical test methods* [Online]. Available: http://www.nata.asn.au/phocadownload/publications/Guidance_information/tech-notes-information-papers/technical_note_17.pdf [Accessed 29 March 2013].
- NATIONAL ASSOCIATION OF TESTING AUTHORITIES. April 2009b. *Technical Note 33 - Guidelines for estimating and reporting measurement uncertainty of chemical test results* [Online]. Available:

- http://www.nata.asn.au/phocadownload/publications/Guidance_information/tech-notes-information-papers/technical_note_33.pdf [Accessed 29 March 2013].
- NIST. 2011. *Standard Reference Data Program - Benzoic Acid* [Online]. <http://webbook.nist.gov/cgi/cbook.cgi?ID=C65850&Mask=1EFF>: National Institute of Standards and Technology. [Accessed 14 December 2014].
- NMI 2011. National Measurement Institute, Measurement Uncertainty for Chemists Course, Participants Handout.
- NOAA, W. National Oceanic and Atmospheric Administration - Global Monitoring Division.
- ONLINE, S. G. S. 2016. <https://www.saiglobal.com/online/>. [Accessed 9 April 2016].
- PETERS, W., JACOBSON, A. R., SWEENEY, C., ANDREWS, A. E., CONWAY, T. J., MASARIE, K., MILLER, J. B., BRUHWILER, L. M. P., PETRON, G. & HIRSCH, A. I. 2007. An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. *Proceedings of the National Academy of Sciences*, 104, 18925.
- PLATZNER, I., HABFAST, K., WALDER, A. & GOETZ, A. 1997. *Modern Isotope Ratio Mass Spectrometry*, Wiley.
- PODLESAK, D. W., BOWEN, G., O'GRADY, S. P., CERLING, T. E. & EHLERINGER, J. 2012. $\delta^2\text{H}$ and $\delta^{18}\text{O}$ of human body water: A GIS model to distinguish residents from non-residents in the contiguous USA. *Isotopes in Environmental and Health Studies*, 48, 259-279.
- POON, N. L., HO, S. S. H. & LI, C. K. 2005. Differentiation of coloured inks of inkjet printer cartridges by thin layer chromatography and high performance liquid chromatography. *Science and Justice - Journal of the Forensic Science Society*, 45, 187-194.
- PRESTON, T. & OWENS, N. 1983. Interfacing an automatic elemental analyser with an isotope ratio mass spectrometer: the potential for fully automated total nitrogen and nitrogen-15 analysis. *The Analyst*, 108, 971-977.
- QI, H. & COPLEN, T. B. 2011. Investigation of preparation techniques for 2H analysis of keratin materials and a proposed analytical protocol. *Rapid Communications in Mass Spectrometry*, 25, 2209-2222.
- RANDERSON, J., FIELD, C., FUNG, I. & TANS, P. 1999. Increases in early season ecosystem uptake explain recent changes in the seasonal cycle of atmospheric CO_2 at high northern latitudes. *Geophysical Research Letters*, 26, 2765-2768.
- RAYNER, P., ENTING, I., FRANCEY, R. & LANGENFELDS, R. 1999. Reconstructing the recent carbon cycle from atmospheric CO_2 , ^{13}C and O_2/N_2 observations*. *Tellus B*, 51, 213-232.
- REVESZ, K., HAIPING, Q. & COPLEN, T. 2012. Determination of the $\delta^{34}\text{S}$ of Total Sulfur in Solids. In: REVESZ, K. (ed.) *Section C, Stable Isotope-Ratio Methods Book 10, Methods of the Reston Stable Isotope Laboratory*. <http://pubs.usgs.gov/tm/10c4/tm10c4.pdf>: United States Geological Survey.
- RIBAUX, O., MARGOT, P., JULIAN, R. & KELTY, S. 2013. Forensic intelligence.
- ROBERTS, J. 1996a. *The chemistry of paper*, Royal Society of Chemistry.
- ROBERTS, J. C. 1996b. *Paper chemistry*, Kluwer Academic Pub.
- RODEN, J., LIN, G. & EHLERINGER, J. 2000. A mechanistic model for interpretation of hydrogen and oxygen isotope ratios in tree-ring cellulose. *Geochimica et Cosmochimica Acta*, 64, 21-35.
- ROUX, C., NOVOTNY, M., EVANS, I. & LENNARD, C. 1999. A study to investigate the evidential value of blue and black ballpoint pen inks in Australia. *Forensic Science International*, 101, 167-176.
- ROZANSKI, K., SONNTAG, C. & MUNNICH, K. 1982. Factors controlling stable isotope composition of European precipitation. *Tellus*, 34, 142-150.

- ROŽIC, M., ROZMARIC MACEFAT, M. & OREŠCANIN, V. 2005. Elemental analysis of ashes of office papers by EDXRF spectrometry. *Nuclear Inst. and Methods in Physics Research, B*, 229, 117-122.
- SCHAUER, A., LAI, C. T., BOWLING, D. & EHLERINGER, J. 2003. An automated sampler for collection of atmospheric trace gas samples for stable isotope analyses. *Agricultural and Forest Meteorology*, 118, 113-124.
- SCHLESINGER, H. & SETTLE, D. 1971. A large-scale study of paper by neutron activation analysis. *J Forensic Sci*, 16, 309-30.
- SHARP, Z. 2007. *Principles of stable isotope geochemistry*, Pearson Education Upper Saddle River, NJ.
- SIRFER 2010. University of Utah, Stable Isotope Ratio Facility for Environmental Research, Stable Isotopes in Ecology Laboratory Manual.
- SMALLDON, K. & MOFFAT, A. 1973. The calculation of discriminating power for a series of correlated attributes. *Journal of the Forensic Science Society*, 13, 291-295.
- SPENCE, L., BAKER, A. & BYRNE, J. 2000. Characterization of document paper using elemental compositions determined by inductively coupled plasma mass spectrometry. *Journal of Analytical Atomic Spectrometry*, 15, 813-819.
- SPENCE, L., FRANCIS, R. & TINGGI, U. 2002. Comparison of the elemental composition of office document paper: evidence in a homicide case. *Journal of Forensic Sciences*, 47, 648.
- SPINDLER, X., SHIMMON, R., ROUX, C.P. & LENNARD, C.J. 2011. The effect of zinc chloride, humidity and the substrate on the reaction of 1,2-indanedione-zinc with amino acids in latent fingerprint secretions. *Forensic Science International*, 212, 150-157.
- STERNBERG, L., DENIRO, M. & JOHNSON, H. 1984a. Isotope ratios of cellulose from plants having different photosynthetic pathways. *Plant Physiology*, 74, 557.
- STERNBERG, L., DENIRO, M. & TING, I. 1984b. Carbon, hydrogen, and oxygen isotope ratios of cellulose from plants having intermediary photosynthetic modes. *Plant Physiology*, 74, 104.
- SUESS, H. U. 2010. *Pulp Bleaching Today*, Walter de Gruyter.
- TAO, F., LIU, Y. & AN, N. 2010. On the necessity of organic solvent extraction for carbon isotopic analysis of α -cellulose: Implications for environmental reconstructions. *International Journal of Environmental Analytical Chemistry*, 90, 605-619.
- TARONI, F., AITKEN, C., GARBOLINO, P. & BIEDERMANN, A. 2006. *Bayesian Networks and Probabilistic Inference in Forensic Science*, Wiley.
- TARONI, F., BIEDERMANN, A., GARBOLINO, P. & AITKEN, C. 2004. A general approach to Bayesian networks for the interpretation of evidence. *Forensic Science International*, 139, 5-16.
- TARONI, F., MARQUIS, R., SCHMITTBUHL, M., BIEDERMANN, A., THIÉRY, A. & BOZZA, S. 2012. The use of the likelihood ratio for evaluative and investigative purposes in comparative forensic handwriting examination. *Forensic Science International*, 214, 189-194.
- THOMPSON, W. C. & NEWMAN, E. J. 2015. Lay Understanding of Forensic Statistics: Evaluation of Random Match Probabilities, Likelihood Ratios, and Verbal Equivalents. *Law and Human Behavior*.
- TREJOS, T., FLORES, A. & ALMIRALL, J. R. 2010. Micro-spectrochemical analysis of document paper and gel inks by laser ablation inductively coupled plasma mass spectrometry and laser induced breakdown spectroscopy. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 65, 884-895.

- VALENZUELA, L. O., CHESSON, L. A., O'GRADY, S. P., CERLING, T. E. & EHLERINGER, J. R. 2011. Spatial distributions of carbon, nitrogen and sulfur isotope ratios in human hair across the central United States. *Rapid Communications in Mass Spectrometry*, 25, 861-868.
- VAN ES, A., DE KOEIJER, J. & VAN DER PEIJL, G. 2009. Discrimination of Document Paper by XRF, LA-ICP-MS and IRMS using Multivariate Statistical Techniques. *Science & Justice*, 49, 120-126.
- WALLACE - KUNKEL, C. S., LENNARD, C.J., STOILOVIC, M., ROUX, C.P. 2007. Optimisation and Evaluation of 1,2-Indanedione for Use as a Fingerprint Reagent and its Application to Real Samples. *Forensic Science International*, 168, 14-26.
- WALLACE-KUNKEL, C., LENNARD, C., STOILOVIC, M. & ROUX, C. 2007. Optimisation and evaluation of 1,2-indanedione for use as a fingerprint reagent and its application to real samples. *Forensic Science International*, 168, 14-26.
- WAMPLER, T. & LEVY, E. 1986. Pyrolysis GC in the Analysis of Inks and Papers. *LC GC*, 4, 1112-1116.
- WASHALL, J. W. & WAMPLER, T. P. 1989. Analytical pyrolysis of complex multicomponent samples. *Journal of chromatographic science*, 27, 144-148.
- WERNER, R. & BRAND, W. 2001. Referencing strategies and techniques in stable isotope ratio analysis. *Rapid Communications in Mass Spectrometry*, 15, 501-519.
- WERNER, R., BRUCH, B. & BRAND, W. 1999. ConFlo III-An Interface for High Precision d13C and d15N Analysis with an Extended Dynamic Range. *Rapid Communications in mass spectrometry*, 13, 1237-1241.
- WEST, J., HURLEY, J., DUDÁS, F. & EHLERINGER, J. 2009a. The Stable Isotope Ratios of Marijuana. II. Strontium Isotopes Relate to Geographic Origin. *Journal of Forensic Sciences*, 54, 1261-1269.
- WEST, J., HURLEY, J. & EHLERINGER, J. 2009b. Stable Isotope Ratios of Marijuana. I. Carbon and Nitrogen Stable Isotopes Describe Growth Conditions*. *Journal of Forensic Sciences*, 54, 84-89.
- WEST, J. B., BOWEN, G. J. & DAWSON, T. E. 2010. *Isoscapes: understanding movement, pattern, and process on earth through isotope mapping*, Springer Verlag.
- WILSON, J. D., LAPORTE, G. M. & CANTU, A. A. 2004. Differentiation of Black Gel Inks Using Optical and Chemical Techniques. *Journal of Forensic Sciences*, 49, 364-370.
- WOLFGANG, T. F., JOHANNES, R. & KROKER, E. 2001. *Calcium carbonate: from the Cretaceous period into the 21st century*, Springer.
- YINON, J. 2004. *Advances in forensic applications of mass spectrometry*, CRC.
- ZIDERMAN, I. 1981. A Simple Technique for Comparing Paper Samples by Their Performance as a Chromatographic Sorbent (Inverse Paper Chromatography). *Journal of Forensic Sciences*, 26, 387-392.

14. Appendix 1 – Data Tables

Mean Sample Weight (μg)	Mean Instrument Response (mV)	Standard Deviation Instrument Response (mv)	RSD Instrument Response (%)	$\delta^{13}\text{C}$ Mean \pm 1 Std Dev (%)	mV/ μg Sample
Medium Fibre Cellulose					
25.00	407.67	66.88	16.41	-23.76 \pm 0.56	16.31
45.67	779.00	64.12	8.23	-25.11 \pm 0.11	17.06
72.67	1328.67	140.06	10.54	-25.82 \pm 0.16	18.28
105.67	2013.33	199.77	9.92	-26.44 \pm 0.14	19.05
124.67	2369.33	133.13	5.62	-26.58 \pm 0.03	19.00
151.67	3029.33	228.34	7.54	-26.74 \pm 0.04	19.97
174.33	3566.67	75.01	2.10	-26.86 \pm 0.17	20.46
202.00	4250.00	194.66	4.58	-27.49 \pm 0.17	21.04
Cellulose Acetate					
25.00	594.33	143.20	24.09	-31.47 \pm 0.10	23.77
54.00	1266.33	76.00	6.00	-31.88 \pm 0.09	23.45
78.67	1807.00	97.71	5.41	-31.99 \pm 0.01	22.97
102.67	2375.67	75.59	3.18	-32.02 \pm 0.01	23.14
125.67	2930.33	22.85	0.78	-32.01 \pm 0.04	23.32
150.33	3525.00	93.95	2.67	-32.09 \pm 0.04	23.45
179.67	4223.00	61.58	1.46	-32.13 \pm 0.01	23.50
205.67	4786.67	164.81	3.44	-32.07 \pm 0.02	23.27
Alpha Glucose					
31.50	352.00	123.04	34.95	-8.83 \pm 0.44	11.17
49.00	802.67	57.84	7.21	-10.05 \pm 0.57	16.38
69.33	1167.67	132.46	11.34	-10.22 \pm 0.04	16.84
98.33	1719.33	115.77	6.73	-10.62 \pm 0.02	17.49
123.67	2490.00	345.74	13.89	-10.83 \pm 0.15	20.13
153.33	2915.00	32.51	1.12	-10.88 \pm 0.01	19.01
172.67	3407.67	124.64	3.66	-11.00 \pm 0.03	19.74
199.33	3842.33	473.12	12.31	-11.06 \pm 0.05	19.28

Table 14.1: Summarized data for linear range experiments for $\delta^{13}\text{C}$ method validation

Paper Samples		$\delta^{13}\text{C}$ Day 1 (‰)	$\delta^{13}\text{C}$ Day 16 (‰)	Difference (‰)
Paper 1	Mean	-22.15	-23.56	1.41
	St Dev	0.75	0.06	0.69
Paper 2	Mean	-27.69	-28.70	1.01
	St Dev	0.31	0.10	0.21
Paper 3	Mean	-22.88	-23.35	0.47
	St Dev	0.38	0.05	0.33
Paper 4	Mean	-23.64	-24.05	0.41
	St Dev	0.40	0.01	0.39
Paper 5	Mean	-26.05	-26.51	0.46
	St Dev	0.20	0.06	0.14
Cellulose and Glucose Samples				
Mannose	Mean	-22.47	-22.84	0.37
	St Dev	0.04	0.01	0.03
Medium Cellulose	Mean	-26.26	-26.86	0.6
	St Dev	0.50	0.09	0.41
Alpha Glucose	Mean	-11.22	-11.20	0.02
	St Dev	0.80	0.03	0.77
D -Glucose	Mean	-10.68	-10.83	0.15
	St Dev	0.16	0.07	0.09
Cellulose Acetate	Mean	-31.62	-32.06	0.44
	St Dev	0.58	0.04	0.54

Table 14.2: Day 1 vs. Day 16 $\delta^{13}\text{C}_{\text{VPDB}}(\text{‰})$ results for samples exposed to the laboratory atmosphere.

Sample Number	Brand	Country of Origin	Date of Packing (if labelled)	Mean $\delta^{13}\text{C}_{\text{VPDB}}$ ± 1 Std Dev (‰)	Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ ± 1 Std Dev (‰)
1	Unknown	Unknown		-23.95 \pm 0.05	28.87 \pm 0.10
2	Office National	Unknown		-28.76 \pm 0.07	21.02 \pm 0.34
3	Canon high quality paper	China		-23.99 \pm 0.06	26.41 \pm 0.09
4	Office Elements	Thailand		-27.54 \pm 0.09	20.51 \pm 0.07
5	Office Elements	Thailand		-23.75 \pm 0.00	26.54 \pm 0.16
6	Office Elements	Thailand		-27.59 \pm 0.10	20.37 \pm 0.05
7	Office Elements	Thailand		-27.55 \pm 0.11	19.98 \pm 0.13
8	Office Elements	Thailand		-27.73 \pm 0.02	19.66 \pm 0.39
9	Canon high quality paper	China		-23.48 \pm 0.08	25.94 \pm 0.21
10	Office Elements	Thailand		-27.72 \pm 0.11	19.91 \pm 0.16
11	Reflex	Australia		-23.74 \pm 0.07	24.86 \pm 0.16
12	Xerox recycled supreme	Italy		-24.60 \pm 0.04	28.25 \pm 0.33
13	Staples Carbon Neutral	Australia		-23.41 \pm 0.08	28.13 \pm 0.21
14	Staples	Australia		-23.76 \pm 0.08	27.64 \pm 0.01
15	Staples	Australia		-23.70 \pm 0.06	27.65 \pm 0.19
16	Staples Carbon Neutral	Australia		-23.95 \pm 0.13	28.46 \pm 0.51
17	Staples Carbon Neutral	Australia		-23.46 \pm 0.08	28.38 \pm 0.03
18	Xerox	France		-24.94 \pm 0.04	27.18 \pm 0.04
19	Corporate Express	Australia		-23.24 \pm 0.05	28.75 \pm 0.24
20	4CC	Sweden		-25.63 \pm 0.09	28.75 \pm 0.24

Sample Number	Brand	Country of Origin	Date of Packing (if labelled)	Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev}$ (‰)	Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev}$ (‰)
21	Staples	Australia		-23.39 ± 0.16	28.50 ± 0.14
22	Staples	Australia		-23.43 ± 0.12	28.51 ± 0.19
23	Staples	Australia		-23.44 ± 0.11	28.74 ± 0.20
24	Unknown	Unknown		-23.37 ± 0.07	28.80 ± 0.01
25	Unknown	Unknown		-23.55 ± 0.03	28.95 ± 0.09
26	Unknown	Unknown		-23.39 ± 0.02	28.27 ± 0.00
27	Reflex	Australia		-23.40 ± 0.10	28.54 ± 0.14
28	Unknown	Unknown		-23.55 ± 0.02	29.14 ± 0.10
29	Reflex	Australia	04/11/11	-23.56 ± 0.04	28.85 ± 0.07
30	Unknown	Unknown		-23.36 ± 0.10	29.28 ± 0.05
31	Staples	Australia		-23.75 ± 0.17	28.64 ± 0.16
32	Xerox	France		-24.84 ± 0.05	27.44 ± 0.80
33	Reflex	Australia	04/11/11	-23.69 ± 0.14	29.06 ± 0.03
34	Reflex	Australia	6/3/11	-23.65 ± 0.06	29.07 ± 0.03
35	Officemax Multipurpose	China		-23.34 ± 0.07	27.22 ± 0.18
36	EXP800	Thailand		-26.21 ± 0.14	21.61 ± 0.16
37	Excellent Copy Paper	Indonesia		-26.51 ± 0.06	21.18 ± 0.15
38	Value Plus	Unknown	15/10/11	-23.62 ± 0.11	28.93 ± 0.09
39	Staples	Australia	12/02/12	-23.50 ± 0.04	29.05 ± 0.05
40	Staples	Australia	09/01/12	-23.62 ± 0.11	32.45 ± 0.17

Sample Number	Brand	Country of Origin	Date of Packing (if labelled)	Mean $\delta^{13}\text{C}_{\text{VPDB}}$ ± 1 Std Dev (‰)	Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ ± 1 Std Dev (‰)
41	Mondi Colour Copy Paper	Austria		-26.09 \pm 0.09	23.51 \pm 0.02
42	Ditto	Australia	16/06/11	-23.72 \pm 0.09	29.07 \pm 0.12
43	"ditto"	Australia	13/08/09	-23.20 \pm 0.06	28.23 \pm 0.18
44	"ditto"	Australia	14/08/09	-23.31 \pm 0.02	28.44 \pm 0.05
45	"ditto"	Australia	4/03/10	-24.31 \pm 0.04	27.79 \pm 0.15
46	"ditto"	Australia		-23.52 \pm 0.15	29.04 \pm 0.19
47	Ditto	Australia		-23.84 \pm 0.07	28.89 \pm 0.08
48	Post Office	Australia	4/11/10	-23.59 \pm 0.03	28.55 \pm 0.04
49	Australia Post Economy recyclable	Australia	16/09/10	-23.40 \pm 0.10	28.05 \pm 0.47
50	Coles A4 copy paper	India	1/08/10	-26.13 \pm 0.08	23.19 \pm 0.23
51	Coles A4 copy paper	India	30/04/11	-26.26 \pm 0.04	22.86 \pm 0.09
52	Coles	India		-25.19 \pm 0.04	24.80 \pm 0.12
53	Fuji Xerox	China		-23.69 \pm 0.13	26.72 \pm 0.14
54	Fuji Xerox	China		-26.82 \pm 0.04	N/A
55	Fuji Xerox Performer	China		-23.53 \pm 0.15	27.24 \pm 0.52
56	HP	Brazil		-25.76 \pm 0.04	23.15 \pm 0.18
57	Select One	Australia	12/05/11	-23.47 \pm 0.04	29.24 \pm 0.16
58	Juan Jian	China		-23.65 \pm 0.06	24.68 \pm 0.20
59	Fuji Xerox Professional	China		-23.72 \pm 0.11	27.47 \pm 0.22
60	Fuji Xerox Laserprint	Indonesia		-26.96 \pm 0.15	22.65 \pm 0.20

Sample Number	Brand	Country of Origin	Date of Packing (if labelled)	Mean $\delta^{13}\text{C}_{\text{VPDB}}$ ± 1 Std Dev (‰)	Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ ± 1 Std Dev (‰)
61	Fuji Xerox Business	China		-23.75 \pm 0.08	27.91 \pm 0.00
62	Canon Business	China		-23.59 \pm 0.08	28.01 \pm 0.04
63	Canon Premium	Austria		-26.88 \pm 0.13	26.21 \pm 0.06
64	Black and Gold	China		-23.65 \pm 0.06	27.18 \pm 0.11
65	Quality	Thailand	9/01/11	-25.93 \pm 0.06	25.16 \pm 0.14
66	Olympic	Australia	1/04/09	-23.19 \pm 0.14	27.76 \pm 0.17
67	Olympic	Australia	13/01/10	-23.29 \pm 0.08	27.63 \pm 0.11
68	Olympic	Australia	12/06/10	-22.57 \pm 0.19	27.37 \pm 0.00
69	Olympic	Australia	18/08/10	-23.37 \pm 0.04	29.10 \pm 0.18
70	Olympic	Australia	16/09/11	-23.44 \pm 0.19	28.84 \pm 0.07
71	Reflex	Australia	10/09/09	-23.27 \pm 0.02	28.16 \pm 0.18
72	Reflex	Australia	17/03/11	-23.67 \pm 0.07	29.18 \pm 0.12
73	Reflex	Australia	28/11/11	-23.60 \pm 0.02	29.14 \pm 0.19
74	Reflex	Australia	12/10/09	-23.49 \pm 0.06	27.87 \pm 0.07
75	Reflex	Australia	10/08/09	-23.40 \pm 0.08	27.99 \pm 0.19
76	Reflex	Australia	5/08/09	-23.40 \pm 0.33	28.13 \pm 0.15
77	Reflex	Australia	18/03/10	-23.81 \pm 0.06	27.96 \pm 0.09
78	Reflex	Australia	3/02/10	-23.97 \pm 0.15	29.53 \pm 0.13
79	Reflex	Australia	14/02/10	-23.49 \pm 0.08	28.19 \pm 0.21

Sample Number	Brand	Country of Origin	Date of Packing (if labelled)	Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev}$ (‰)	Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev}$ (‰)
80	Reflex	Australia	11/07/10	-23.87 ± 0.17	28.28 ± 0.15
81	Reflex	Australia	6/11/09	-23.69 ± 0.19	28.04 ± 0.20
82	Paper One	Indonesia		-29.57 ± 0.04	19.46 ± 0.24
83	Paper One	Indonesia		-23.45 ± 0.08	20.57 ± 0.09
84	Paper One	Indonesia		-29.70 ± 0.09	19.39 ± 0.00
85	Paper One	Indonesia		-29.35 ± 0.10	20.66 ± 0.03
86	Officeworks	Australia	26/09/09	-23.44 ± 0.04	27.51 ± 0.00
87	Officeworks	Australia		-25.01 ± 0.06	24.34 ± 0.19
88	Lazer IT	Indonesia		-27.50 ± 0.07	20.18 ± 0.00
89	Lazer IT	Indonesia		-26.54 ± 0.04	21.19 ± 0.13
90	Lazer IT	Indonesia		-26.26 ± 0.11	21.09 ± 0.26
91	Lazer IT	Indonesia		-26.90 ± 0.17	21.95 ± 0.08
92	Double A	Thailand	7/07/09	-27.18 ± 0.07	23.41 ± 0.12
93	Double A	Thailand	28/08/09	-27.85 ± 0.11	22.58 ± 0.04
94	Double A	Thailand	18/01/10	-28.08 ± 0.02	20.48 ± 0.17
95	Double A	Thailand	15/06/09	-28.67 ± 0.05	20.84 ± 0.19
96	Double A	Thailand	10/09/09	-27.53 ± 0.14	20.46 ± 0.08
97	Double A	Thailand	28/08/09	-27.89 ± 0.10	20.71 ± 0.02
98	Double A	Thailand		-28.19 ± 0.00	20.96 ± 0.13

Sample Number	Brand	Country of Origin	Date of Packing (if labelled)	Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev} (\text{‰})$	Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev} (\text{‰})$
99	Double A	Thailand	23/12/10	-27.59 \pm 0.08	23.04 \pm 0.04
100	Double A	Thailand	12/10/11	-27.75 \pm 0.05	23.83 \pm 0.16
101	Coles Smart Buy	China		-25.34 \pm 0.09	19.44 \pm 0.25
102	Coles Smart Buy	China		-22.81 \pm 0.17	25.45 \pm 0.02
103	Yes Bronze	China		-24.62 \pm 0.12	24.38 \pm 0.09
104	Kodak	Indonesia		-28.48 \pm 0.12	20.82 \pm 0.15
105	Victory Docucopy	China		-23.96 \pm 0.16	20.87 \pm 0.18
106	Victory Docucopy	China		-24.07 \pm 0.01	21.27 \pm 0.33
107	Victory Docucopy	China		-23.62 \pm 0.34	23.75 \pm 0.33
108	Victory Docucopy	China		-23.37 \pm 0.13	22.12 \pm 0.12
109	Officemax	South Africa		-25.17 \pm 0.09	26.28 \pm 0.29
110	Officemax	South Africa		-23.61 \pm 0.04	27.79 \pm 0.10
111	Bilt Matrix	India	29/09/09	-25.65 \pm 0.06	24.86 \pm 0.17
112	Bilt Matrix	India	25/04/11	-26.27 \pm 0.13	22.97 \pm 0.15
113	Canon	China		-23.98 \pm 0.05	26.96 \pm 0.16
114	Canon	China		-23.47 \pm 0.03	27.12 \pm 0.07
115	Woolworths essential copy paper	Indonesia		-27.71 \pm 0.02	21.76 \pm 0.13
116	Reflex Platnium	Australia	23/03/10	-23.58 \pm 0.02	30.22 \pm 0.22
117	Spilman	Australia	8/11/09	-23.50 \pm 0.04	29.59 \pm 0.08

Sample Number	Brand	Country of Origin	Date of Packing (if labelled)	Mean $\delta^{13}\text{C}_{\text{VPDB}}$ ± 1 Std Dev (‰)	Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ ± 1 Std Dev (‰)
118	Spilman	Australia	31/07/10	-23.32 \pm 0.02	29.39 \pm 0.10
119	Spilman	Australia	26/10/10	-23.34 \pm 0.09	28.86 \pm 0.23
120	Yes Silver	China		-23.39 \pm 0.10	27.35 \pm 0.01
121	HP	Brazil		-25.78 \pm 0.01	23.06 \pm 0.00
122	HP	Brazil		-26.02 \pm 0.07	23.45 \pm 0.02
123	HP	Brazil		-26.02 \pm 0.14	N/A
124	Paper One Premium	Indonesia		-28.67 \pm 0.02	21.70 \pm 0.13
125	Victory High White	Indonesia		-29.03 \pm 0.04	21.01 \pm 0.06

Table 14.3: Paper samples collected as background population samples, and their measured $\delta^{13}\text{C}$ and $\delta^{18}\text{O}_{\text{VSMOW}}$ (‰) abundance values

Sample Number	Printer Manufacturer	Make	Model	Paper Brand	Paper Country of Production
1	Brother	MFC	7340	Unknown	N/A
2	Kyocer	KM	2050	Office National	N/A
3	HP	Laserjet	P1606dn	Canon High Quality paper	China
4	Fuji Xerox	DocuCentre	C2200	Office Elements	Thailand
5	Canon	LBP	3100 B	Office Elements	Thailand
6	Brother	HL	2142	Office Elements	Thailand
7	Toshiba	estudio	255	Office Elements	Thailand
8	HP	Laserjet	P1505n	Office Elements	Thailand
9	Fuji Xerox	DocuPrint	240A	Australian	Australia
10	Fuji Xerox	DocuCentre	336	Australian	Australia
11	Fuji Xerox	Apeos Port IV	C5570	Xerox Recycled supreme	Japan
12	Fuji Xerox	Apeos Port II	C4300	Staples Carbon Neutral	Australia
13	Lexmark	T644		Staples	Australia
14	Ricoh	Aficio	2060	Staples Carbon Neutral	Australia
15	Ricoh	Aficio	2060	Staples Carbon Neutral	Australia
16	HP		2605	4CC	Sweden
17	Ricoh	Aficio	MP C2000	Staples	Australia
18	HP	Laserjet	P4515x	Unknown	Unknown
19	Ricoh	Aficio	MP 3350	Reflex	Australia
20	Ricoh	Aficio	2075	Staples	Australia
21	Fuji Xerox	DocuColor	252	Xerox	France
22	Konica Minolta	Bizhub	C280	Reflex	Australia
23	Lexmark		E450dn	EXP800	Thailand
24	Fuji Xerox	Apeos Port	5070	Staples	Australia
25	Fuji Xerox	Docucentre	336	Staples	Australia
26	Fuji Xerox	DocuPrint	240A	Staples	Australia
27	Lexmark	HFDI	C935	Mondi Colour Copy Paper	Austria
28	Konica Minolta	Bizhub	C280	Officemax Multipurpose	China
29	Konica Minolta	Bizhub	C280	EXP800	Thailand
30	Lexmark		E450dn	EXPGreen 50% Recycled	Australia
31	Konica Minolta	Bizhub	C280	EXPGreen 50% Recycled	Australia
32	HP	Laserjet	P1606dn	Canon High Quality Paper	China
33	HP	Laserjet	P1606dn	Office Elements	Thailand
34	Ricoh	Aficio	SC420DN	Xerox	France

Sample Number	Printer Manufacturer	Make	Model	Paper Brand	Paper Country of Production
35	Ricoh	Aficio	2060	Corporate Express	Australia
36	HP		2605	Staples	Australia
37	Ricoh	Aficio	SP C420DN	Staples	Australia
38	Lexmark		E450dn	Excellent Copy Paper	Indonesia
39	Konica Minolta	Bizhub	C280	Excellent Copy Paper	Indonesia
40	Lexmark	E450dn		Value Plus	unknown
41	Konica Minolta	Bizhub		Value Plus	unknown
42	Ricoh	Aficio	2060	Unknown	Unknown
43	Ricoh	Aficio	SP C430DN	Unknown	Unknown
44	Ricoh	Aficio	2060	Reflex	Australia
45	Lexmark		E450dn	Reflex	Australia
46	Lexmark		E450dn	Officemax Multipurpose	China

Table 14.4: Printer manufacturer and paper information for toner printers tested for their effect on the $\delta^{13}\text{C}$ abundance values of paper.

Sample Number	Blank Paper Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev}$ (‰)	Printed Paper Non-image Area Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev}$ (‰)	Printed Paper Sampled on Image Area Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1 \text{ Std Dev}$ (‰)
1	-23.95 ± 0.05	-24.01 ± 0.05	-24.38 ± 0.01
2	-28.76 ± 0.07	-28.75 ± 0.03	-28.80 ± 0.00
3	-23.99 ± 0.06	-23.94 ± 0.04	-25.16 ± 0.05
4	-27.54 ± 0.09	-27.54 ± 0.08	-27.82 ± 0.06
5	-23.75 ± 0.00	-23.89 ± 0.05	-25.03 ± 0.14
6	-27.59 ± 0.10	-27.55 ± 0.06	-27.59 ± 0.05
7	-27.55 ± 0.11	-27.69 ± 0.07	-27.86 ± 0.06
8	-27.73 ± 0.02	-27.80 ± 0.08	-27.89 ± 0.07
9	-23.82 ± 0.19	-23.79 ± 0.00	-24.16 ± 0.07
10	-23.40 ± 0.05	-23.41 ± 0.12	-24.29 ± 0.06
11	-24.60 ± 0.04	-24.53 ± 0.09	-25.24 ± 0.02
12	-23.41 ± 0.08	-23.46 ± 0.09	-23.91 ± 0.05
13	-23.76 ± 0.08	-23.70 ± 0.09	-24.54 ± 0.00
14	-23.95 ± 0.13	-23.96 ± 0.03	-24.26 ± 0.12
15	-23.46 ± 0.08	-23.40 ± 0.14	-23.84 ± 0.08
16	-25.63 ± 0.09	-25.65 ± 0.09	-25.98 ± 0.02
17	-23.44 ± 0.11	-23.45 ± 0.11	-23.86 ± 0.05
18	-23.55 ± 0.03	-23.63 ± 0.03	-24.28 ± 0.08
19	-23.56 ± 0.04	-23.56 ± 0.06	-23.97 ± 0.11
20	-23.75 ± 0.17	-23.74 ± 0.03	-24.31 ± 0.00
21	-25.17 ± 0.57	-24.89 ± 0.10	-25.20 ± 0.03
22	-23.65 ± 0.06	-23.55 ± 0.04	-23.90 ± 0.08
23	-26.21 ± 0.14	-26.27 ± 0.00	-26.41 ± 0.06
24	-23.50 ± 0.04	-23.58 ± 0.06	-24.03 ± 0.13
25	-23.62 ± 0.11	-23.69 ± 0.15	-24.30 ± 0.05
26	-23.65 ± 0.01	-23.66 ± 0.01	-24.11 ± 0.03
27	-26.09 ± 0.09	-26.03 ± 0.09	-26.27 ± 0.09
28	-23.34 ± 0.07	-23.65 ± 0.12	-23.89 ± 0.00
29	-26.21 ± 0.14	-26.16 ± 0.10	-26.38 ± 0.07
30	-23.76 ± 0.21	-23.85 ± 0.11	-24.35 ± 0.00
31	-23.76 ± 0.21	-23.84 ± 0.19	-24.24 ± 0.08
32	-23.48 ± 0.08	-23.68 ± 0.06	-24.46 ± 0.05
33	-27.72 ± 0.11	-27.87 ± 0.06	-28.04 ± 0.04
34	-24.94 ± 0.04	-24.98 ± 0.08	-25.22 ± 0.07
35	-23.24 ± 0.05	-23.30 ± 0.03	-23.90 ± 0.06
36	-23.39 ± 0.16	-23.51 ± 0.06	-24.31 ± 0.09
37	-23.43 ± 0.12	-23.86 ± 0.76	-23.88 ± 0.05
38	-26.51 ± 0.10	-26.58 ± 0.14	-26.81 ± 0.12

Sample Number	Blank Paper Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1$ Std Dev (‰)	Printed Paper Non-image Area Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1$ Std Dev (‰)	Printed Paper Sampled on Image Area Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1$ Std Dev (‰)
39	-26.51 ± 0.06	-26.54 ± 0.00	-26.64 ± 0.06
40	-23.62 ± 0.11	-23.60 ± 0.11	-24.23 ± 0.16
41	-23.62 ± 0.11	-23.47 ± 0.05	-23.98 ± 0.20
42	-23.39 ± 0.02	-23.57 ± 0.07	-23.86 ± 0.13
43	-23.36 ± 0.10	-23.43 ± 0.02	-23.83 ± 0.02
44	-23.69 ± 0.14	-23.58 ± 0.02	-24.00 ± 0.04
45	-23.54 ± 0.02	-23.55 ± 0.00	-24.09 ± 0.06
46	-23.41 ± 0.08	-23.48 ± 0.08	-24.05 ± 0.09
39	-26.51 ± 0.06	-26.54 ± 0.00	-26.64 ± 0.06
40	-23.62 ± 0.11	-23.60 ± 0.11	-24.23 ± 0.16
41	-23.62 ± 0.11	-23.47 ± 0.05	-23.98 ± 0.20
42	-23.39 ± 0.02	-23.57 ± 0.07	-23.86 ± 0.13
43	-23.36 ± 0.10	-23.43 ± 0.02	-23.83 ± 0.02
44	-23.69 ± 0.14	-23.58 ± 0.02	-24.00 ± 0.04
45	-23.54 ± 0.02	-23.55 ± 0.00	-24.09 ± 0.06
46	-23.41 ± 0.08	-23.48 ± 0.08	-24.05 ± 0.09

Table 14.5: Mean $\delta^{13}\text{C}_{\text{VPDB}}$ and 1 x standard deviation results for 46 toner printed documents

Sample Number	Manufacturer	Make	Model	Paper Brand	Paper Origin
1	HP	Business Inkjet	2800	Staples	Australia
2	HP	Deskjet	5652	Reflex	Australia
3	Canon	MX	870	Australian	Australia
4	Canon	Pixma	iP4600	Unknown	Unknown
5	Epson	Stylus Photo	R2400	Officemax Multipurpose	China
6	Epson	Stylus Photo	R390	Unknown	N/A
7	Epson	Stylus Pro	4800	Canon High Quality 106gsm	Japan
8	Brother	MFC	215C	Ditto	Australia
9	Epson	Stylus Photo	R2400	EXP800	Thailand
10	Epson	Stylus Photo	R2400	EXPGreen 50% Recycled	Australia
11	Epson	Stylus Photo	R2400	Excellent Copy Paper	Indonesia
12	Epson	Stylus Photo	R2400	Value Plus	N/A
13	Epson	Stylus Photo	R2400	Reflex	Australia

Table 14.6: Printer manufacturer and paper information for inkjet printers tested for their effect on the $\delta^{13}\text{C}$ abundance values of paper.

Sample Number	Blank Paper Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1$ Std Dev (‰)	Printed Paper Non-image Area Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1$ Std Dev (‰)	Printed Paper on Sampled on Image Area Mean $\delta^{13}\text{C}_{\text{VPDB}} \pm 1$ Std Dev (‰)
1	-23.70 ± 0.06	-23.64 ± 0.08	-23.81 ± 0.03
2	-23.74 ± 0.07	-27.73 ± 0.03	-27.75 ± 0.04
3	-23.40 ± 0.10	-23.63 ± 0.24	-23.61 ± 0.00
4	-23.40 ± 0.07	-23.63 ± 0.03	-23.86 ± 0.09
5	-23.55 ± 0.02	-23.48 ± 0.03	-23.72 ± 0.04
6	-23.34 ± 0.07	-23.33 ± 0.12	-23.53 ± 0.12
7	-23.37 ± 0.07	-23.33 ± 0.02	-23.49 ± 0.04
8	-25.08 ± 0.10	-25.05 ± 0.03	-25.20 ± 0.05
9	-23.72 ± 0.09	-23.69 ± 0.20	-23.97 ± 0.14
10	-26.21 ± 0.14	-26.18 ± 0.17	-26.14 ± 0.06
11	-23.76 ± 0.21	-23.84 ± 0.19	-24.24 ± 0.08
12	-26.51 ± 0.06	-26.52 ± 0.12	-26.50 ± 0.04
13	-23.62 ± 0.11	-23.44 ± 0.04	-23.55 ± 0.04

Table 14.7: Mean $\delta^{13}\text{C}_{\text{VPDB}}$ and 1 x standard deviation results for 14 inkjet printed documents

Sample Weight (μg)	Instrument Response (mV)	Raw $\delta^{18}\text{O}$ (‰)	mV/ μg Sample
Paper			
98	3099	57.062	31.62
143	4690	47.511	32.80
245	8137	25.651	33.21
357	12214	28.023	34.21
496	16909	29.322	34.09
International Cellulose			
94	3294	28.279	35.04
152	5312	33.097	34.95
248	9071	35.118	36.58
339	12483	36.271	36.82
503	18129	37.757	36.04
IAEA-601			
96	831	28.983	8.66
156	1197	26.19	7.67
251	1957	27.059	7.80
341	2479	26.486	7.27
493	3730	26.909	7.57
IAEA-602			
99	779	77.49	7.87
147	1207	79.994	8.21
241	1816	71.389	7.54
350	2950	76.298	8.43
506	3710	76.334	7.33

Table 14.8: Summarized data for linear range experiments for $\delta^{18}\text{O}$ method validation

Sample Number	Printer Manufacturer	Make	Model	Paper Brand	Paper Country of Production
1	Brother	MFC	7340	Unknown	N/A
2	Kyocer	KM	2050	Office National	N/A
4	Fuji Xerox	DocuCentre	C2200	Office Elements	Thailand
5	Canon	LBP	3100 B	Office Elements	Thailand
6	Brother	HL	2142	Office Elements	Thailand
7	Toshiba	estudio	255	Office Elements	Thailand
8	HP	Laserjet	P1505n	Office Elements	Thailand
9	Fuji Xerox	DocuPrint	240A	Australian	Australia
10	Fuji Xerox	DocuCentre	336	Australian	Australia
11	Fuji Xerox	Apeos Port IV	C5570	Xerox Recycled supreme	Japan
12	Fuji Xerox	Apeos Port II	C4300	Staples Carbon Neutral	Australia
13	Lexmark	T644		Staples	Australia
17	Ricoh	Aficio	MP C2000	Staples	Australia
18	HP	Laserjet	P4515x	Unknown	Unknown
19	Ricoh	Aficio	MP 3350	Reflex	Australia
20	Ricoh	Aficio	2075	Staples	Australia
21	Fuji Xerox	DocuColor	252	Xerox	France
23	Lexmark	E450dn		EXP800	Thailand
24	Fuji Xerox	Apeos Port	5070	Staples	Australia
27	Lexmark	HFDI	C935	Mondi Colour Copy Paper	Austria
29	Konica Minolta	Bizhub	C280	EXP800	Thailand
34	Ricoh	Aficio	SP C420DN	Xerox	France
44	Ricoh	Aficio	2060	Reflex	Australia

Table 14.9: Printer manufacturer and paper information for toner printers tested for $\delta^{18}\text{O}_{\text{VSMOW}}$

Sample Number	Blank Paper Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev}$ (‰)	Printed Paper Non-image Area Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev}$ (‰)	Printed Paper Sampled on Image Area Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ Std Dev}$ (‰)
1	28.62 ± 0.43	28.48 ± 0.03	28.09 ± 0.16
2	21.02 ± 0.34	20.39 ± 0.11	19.46 ± 0.24
4	20.33 ± 0.32	20.11 ± 0.05	19.87 ± 0.12
5	26.38 ± 0.30	25.87 ± 0.19	24.47 ± 0.32
6	20.37 ± 0.05	20.24 ± 0.09	20.02 ± 0.19
7	19.98 ± 0.13	20.17 ± 0.35	20.66 ± 0.84
8	19.92 ± 0.53	19.96 ± 0.22	18.91 ± 0.15
9	27.13 ± 0.00	27.23 ± 0.14	26.89 ± 0.05
10	27.35 ± 0.17	27.71 ± 0.07	27.59 ± 0.35
11	28.49 ± 0.47	29.09 ± 0.17	28.32 ± 0.14
12	28.13 ± 0.21	28.54 ± 0.36	28.39 ± 0.18
13	27.64 ± 0.01	27.62 ± 0.08	27.16 ± 0.15
17	27.18 ± 0.04	26.63 ± 0.14	26.32 ± 0.04
18	28.51 ± 0.19	28.20 ± 0.15	27.89 ± 0.09
19	28.66 ± 0.51	28.40 ± 0.09	26.55 ± 0.03
20	28.96 ± 0.21	28.20 ± 0.13	27.78 ± 0.06
21	28.64 ± 0.16	28.28 ± 0.15	27.77 ± 0.02
23	27.41 ± 0.57	26.80 ± 0.15	26.65 ± 0.18
24	28.93 ± 0.23	28.34 ± 0.09	27.97 ± 0.08
27	21.61 ± 0.16	21.73 ± 0.15	21.34 ± 0.03
29	21.61 ± 0.16	21.21 ± 1.09	21.79 ± 0.07
34	29.05 ± 0.05	28.27 ± 0.10	28.18 ± 0.10
44	23.63 ± 0.20	23.47 ± 0.14	23.25 ± 0.12

Table 14.10: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ and 1 x standard deviation results for 22 toner printed documents

Sample Number	Manufacturer	Make	Model	Paper Brand	Paper Origin
1	HP	Business Inkjet	2800	Staples	Australia
2	HP	Deskjet	5652	Reflex	Australia
3	Canon	MX	870	Australian	Australia
4	Canon	Pixma	iP4600	Unknown	Unknown
6	Epson	Stylus Photo	R390	Unknown	N/A
7	Epson	Stylus Pro	4800	Canon High Quality 106gsm	Japan
8	Brother	MFC	215C	Ditto	Australia
13	Epson	Stylus Photo	R2400	Reflex	Australia

Table 14.11: Printer manufacturer and paper information for inkjet printers tested for $\delta^{18}\text{O}_{\text{VSMOW}}$

Sample Number	Blank Paper Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ St Dev}$ (‰)	Printed Paper Non-image Area Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ St Dev}$ (‰)	Printed Paper on Sampled on Image Area Mean $\delta^{18}\text{O}_{\text{VSMOW}} \pm 1 \text{ St Dev}$ (‰)
1	27.49 ± 0.30	27.81 ± 0.29	27.59 ± 0.21
2	28.54 ± 0.14	28.68 ± 0.09	28.52 ± 0.13
3	29.14 ± 0.17	29.19 ± 0.18	28.96 ± 0.25
4	28.99 ± 0.27	28.75 ± 0.17	28.70 ± 0.06
6	28.80 ± 0.01	28.78 ± 0.08	28.93 ± 0.11
7	22.02 ± 0.20	22.54 ± 0.44	22.89 ± 0.19
8	29.07 ± 0.12	28.68 ± 0.10	29.17 ± 0.27
13	29.23 ± 0.28	28.77 ± 0.17	28.63 ± 0.06

Table 14.12: Mean $\delta^{18}\text{O}_{\text{VSMOW}}$ and 1 x standard deviation results for 8 inkjet printed documents

Sample Number	Brand	Country of Origin	Mean $\delta^2\text{H}_{\text{VSMOW}} \pm 1 \text{ Std Dev } (\text{‰})$
1	Unknown	Unknown	-20.91 \pm 0.49
13	Staples Carbon Neutral	Australia	-18.73 \pm 0.35
14	Staples	Australia	-22.13 \pm 0.64
15	Staples	Australia	-19.13 \pm 2.26
17	Staples Carbon Neutral	Australia	-20.68 \pm 1.45
19	Corporate Express	Australia	-16.65 \pm 2.07
20	4CC	Sweden	-21.46 \pm 0.92
22	Staples	Australia	-19.80 \pm 1.20
23	Staples	Australia	-16.89 \pm 0.23
24	Unknown	Unknown	-20.46 \pm 0.42
25	Unknown	Unknown	-20.84 \pm 0.04
26	Unknown	Unknown	-23.06 \pm 2.17
28	Unknown	Unknown	-20.06 \pm 2.08
29	Reflex	Australia	-25.73 \pm 2.06
30	Unknown	Unknown	-25.82 \pm 0.99
31	Staples	Australia	-24.10 \pm 0.38
109	Officemax	South Africa	-22.37 \pm 1.09
66	Olympic	Australia	-11.55 \pm 0.68
111	Bilt Matrix	India	-31.29 \pm 0.80
117	Spilman	Australia	-23.23 \pm 0.30
74	Reflex	Australia	-23.92 \pm 0.35
75	Reflex	Australia	-16.29 \pm 0.92
94	Double A	Thailand	-41.12 \pm 2.29
95	Double A	Thailand	-43.31 \pm 2.74
118	Spilman	Australia	-24.60 \pm 2.23

Table 14.13: Paper samples selected and measured for their $\delta^2\text{H}_{\text{VSMOW}}$ (‰) values

Brand and Ream Designator	Mean Grammage (gsm)	Mean Apparent Sheet Density (kg/m³)
Lazer IT A	80.53	776.39
Lazer IT B	80.13	777.61
Lazer IT C	80.29	780.30
Lazer IT D	80.49	777.71
Lazer IT E	80.32	780.56
Lazer IT F	79.98	776.48
Lazer IT G	80.18	772.86
Mean and Standard Deviation	80.28 ± 0.2	777.42 ± 2.6
Officemax A	81.16	796.00
Officemax B	81.45	793.92
Officemax C	81.37	801.32
Officemax D	82.55	797.26
Officemax E	81.95	796.85
Officemax F	81.40	789.54
Officemax G	80.40	796.57
Mean and Standard Deviation	81.47 ± 0.66	795.92 ± 3.58
Olympic A	80.49	796.60
Olympic B	80.64	804.35
Olympic C	80.41	814.74
Olympic D	80.13	796.44
Olympic E	81.17	808.87
Olympic F	81.14	808.60
Olympic G	81.18	808.53
Mean and Standard Deviation	80.74 ± 0.42	805.45 ± 6.8
Docucopy A	79.92	732.93
Docucopy B	82.73	776.51
Docucopy C	80.95	735.58
Docucopy D	82.77	755.50
Docucopy E	81.47	769.71
Docucopy F	83.80	756.61
Docucopy G	81.16	770.77
Mean and Standard Deviation	81.83 ± 1.32	756.80 ± 17.18
Fuji Xerox A	81.84	809.54
Fuji Xerox B	81.65	796.58
Fuji Xerox C	80.43	794.36
Fuji Xerox D	80.78	801.36
Fuji Xerox E	80.95	800.26
Fuji Xerox F	81.63	801.85
Fuji Xerox G	81.64	800.75
Mean and Standard Deviation	81.27 ± 0.54	800.67 ± 4.78

Brand and Ream Designator	Mean Grammage (gsm)	Mean Apparent Sheet Density (kg/m³)
Double A A	81.38	791.27
Double A B	80.17	780.99
Double A C	80.42	778.83
Double A D	79.97	774.12
Double A E	80.60	780.98
Double A F	80.46	767.76
Double A G	81.22	764.75
Mean and Standard Deviation	80.60 ± 0.52	776.96 ± 8.96
Reflex A	84.10	833.15
Reflex B	81.67	825.82
Reflex C	82.39	824.33
Reflex D	82.74	830.72
Reflex E	81.52	812.81
Reflex F	82.08	824.09
Reflex G	81.92	817.60
Mean and Standard Deviation	82.35 ± 0.88	824.07 ± 7.05
Mean of 7 replicates from 6 reams of the same brand		
Reflex 2A	83.35 ± 0.78	839.06 ± 21.25
Reflex 2B	82.52 ± 0.71	830.77 ± 6.42
Reflex 2C	82.33 ± 0.38	819.84 ± 7.47
Reflex 2D	82.45 ± 0.66	823.81 ± 7.49
Reflex 2E	81.94 ± 0.72	814.32 ± 5.83
Reflex 2F	82.00 ± 0.72	817.11 ± 6.71
Mean and Standard Deviation	82.43 ± 0.51	824.15 ± 9.28

Table 14.14: Mean grammage and apparent sheet density measured for papers

Sample Number	Grammage (gsm)	Difference to Sample 71	Apparent Sheet Density (kg/m ³)	Difference to Sample 71
71	83.40		819.89	
43	79.24	4.17	792.41	27.48
46	80.11	3.29	837.54	17.65
47	82.06	1.34	831.68	11.79
48	79.71	3.69	806.28	13.61
49	78.80	4.60	793.84	26.05
50	83.45	0.05	808.33	11.56
53	81.32	2.08	793.14	26.75
54	80.93	2.47	755.26	64.63
56	80.32	3.08	812.73	7.16
66	80.00	3.40	784.40	35.49
67	82.49	0.92	819.69	0.20
68	81.20	2.20	821.39	1.50
69	80.29	3.12	827.71	7.82
72	83.99	0.59	836.79	16.90
73	83.59	0.18	843.37	23.48
74	83.71	0.31	832.39	12.50
75	85.09	1.69	840.30	20.41
76	84.28	0.88	845.61	25.72
77	85.80	2.40	884.00	64.11
79	83.53	0.13	861.74	41.85
80	83.13	0.28	845.65	25.76
81	84.38	0.98	850.38	30.49
82	81.90	1.50	804.79	15.11
83	78.99	4.41	821.18	1.29
84	81.90	1.50	808.62	11.27
85	82.21	1.20	806.65	13.24
86	79.98	3.43	840.86	20.97
87	81.18	2.22	833.77	13.87
88	78.45	4.95	755.09	64.80
89	78.50	4.90	761.60	58.29
90	79.15	4.25	799.29	20.60
91	80.06	3.35	782.10	37.79
92	80.93	2.47	791.16	28.73
93	80.44	2.97	796.71	23.18
94	80.43	2.98	797.86	22.03
95	80.82	2.58	677.77	142.12
96	79.95	3.45	782.06	37.84
97	80.91	2.49	784.54	35.35

Sample Number	Grammage (gsm)	Difference to Sample 71	Apparent Sheet Density (kg/m ³)	Difference to Sample 71
71	83.40		819.89	
98	80.99	2.41	764.53	55.36
99	82.05	1.36	787.92	31.98
101	82.74	0.67	773.72	46.17
102	77.88	5.53	776.80	43.09
103	81.30	2.10	797.55	22.34
104	79.93	3.48	788.27	31.63
105	80.38	3.02	751.43	68.46
106	80.23	3.17	735.43	84.46
107	78.50	4.91	723.92	95.97
108	80.56	2.85	757.20	62.69
109	80.73	2.67	773.63	46.26
112	84.18	0.77	821.07	1.18
113	81.11	2.30	807.58	12.31
114	81.58	1.82	798.65	21.24
115	79.72	3.68	779.36	40.53
117	82.49	0.92	834.64	14.74
118	81.72	1.68	839.67	19.78
120	80.16	3.25	781.02	38.87
121	79.82	3.59	822.86	2.97
122	84.85	1.44	828.80	8.90
124	78.95	4.45	788.27	31.62
125	78.93	4.48	782.92	36.97

Table 14.15: Grammage and Apparent Sheet Density results for background papers, with green cells denoting difference values larger than the defined benchmark values.

Plate Number	Sample Number	Measurement (mm)				Calculated Rf of Peaks			Differentiate from Sample 71?
		Peak 1	Peak 2	Peak 3	Mobile Phase End Point	RF Comp 1	RF Comp 2	RF Comp 3	
2	71	47.6	52.7	54.6	67.5	0.71	0.78	0.81	
	43	53.1	54.7			0.79	0.81		
	67	52.8	54.6			0.78	0.81		
	68	53.0	54.8			0.79	0.81		x
	74	53.5	55.0			0.79	0.81		
3	71	52.2	N/A	56.2	64.7	0.82		0.87	
	75	52.6	55.4	56.4		0.81	0.86	0.87	
	76	52.8	55.9	56.4		0.82	0.86	0.87	
	77	52.6	55.5	56.2		0.81	0.86	0.87	
	78	52.2	55.4	56.3		0.81	0.86	0.87	
5	71	52.9	55.2	57.7	67.5	0.78	0.82	0.85	
	84	56.5	58.3			0.84	0.86		x
	86	51.1	56.6	58		0.76	0.84	0.86	
	87	51.0	56.1	58.1		0.76	0.83	0.86	
	88	54.8	56.5			0.81	0.84		x
6	71	50.1	54.7	56.8	66.5	0.75	0.82	0.85	
	88	55.2	57			0.83	0.86		x
	89	55.2	57			0.83	0.86		x
	92	54.2	55.9			0.82	0.84		x
	93	56.5	58.6			0.85	0.88		
7	71	51.5	56.1	57.9	68.2	0.76	0.82	0.85	
	94	51.5	56.0	58		0.76	0.82	0.85	
	95	51.5	55.9	57.9		0.76	0.82	0.85	
	96	56.2	58.2			0.82	0.85		
	97	56.2	58.2			0.82	0.85		
9	71	50.5	54.7	56.6	65.1	0.78	0.84	0.87	
	103	52.9	54.9			0.81	0.84		x
	104	52.9	54.9			0.81	0.84		x
10	71	49.7	53.6	55.3	62.2	0.80	0.86	0.89	

	105	48.9				0.79			x
	106	48.9				0.79			x
	107	48.9	52.9	55.3		0.79	0.85	0.89	
	109	48.8	53.8	55.3		0.78	0.86	0.89	
11	71	50.3	53.2	54.9	63.5	0.79	0.84	0.86	
	102	47.1	48.2			0.74	0.76		x
	112	48.5	52.1	54.1		0.76	0.82	0.85	
12	71	59.1	60.3	61.5	67.1	0.88	0.90	0.92	
	113	59.2	60.6			0.88	0.90		x
	114	54.9	58.2	59.5		0.82	0.87	0.89	x
	P3	54.9	58.2	59.5		0.82	0.87	0.89	x
13	71	59.2	60.3	61.5	66.9	0.88	0.90	0.92	
	115	59.7	61.1			0.89	0.91		
	116	55.3	59	60.4		0.83	0.88	0.90	x
14	71	46	52.2	54.8	65.8	0.70	0.79	0.83	
	117	44.9	50.6	53.7		0.68	0.77	0.82	
	118	44.9	50.6	53.7		0.68	0.77	0.82	
	120	48.8	52.1			0.74	0.79		
	121	N/A							x
15	71	46.4	52.4	55.5	66.9	0.69	0.78	0.83	
	122	N/A							x
	124	51.1	53.9			0.76	0.81		
	125	51.1	53.9			0.76	0.81		
	48	51.1	53.9			0.76	0.81		
16	71	55.8	57.8		65.3	0.85	0.89		
	49	54.6	57.3			0.84	0.88		
	53	52.7	55.1			0.81	0.84		x
17	71	52.8	57.0	58.6	68.3	0.77	0.83	0.86	
	46	51.8	56.4	58.1		0.76	0.83	0.85	
	72	51.8	56.4	58.1		0.76	0.83	0.85	
	57	51.8	56.4	58.1		0.76	0.83	0.85	
	73	51.8	56.4	58.1		0.76	0.83	0.85	
18	71	52.1	56.3	58.0	67.7	0.77	0.83	0.86	
	69	51.9	56.3	58.3		0.77	0.83	0.86	

	119	51.9	56.3	58.3		0.77	0.83	0.86	
	70	51.9	56.3	58.3		0.77	0.83	0.86	
	47	51.9	56.3	58.3		0.77	0.83	0.86	
19	71		57.2	59.3	66.9		0.86	0.89	
	57	53.5				0.80			x
	99	53.2	57.1	59.2		0.80	0.85	0.88	
	54		57.1	59.2			0.85	0.88	
	51	53.2	57.1	59.2		0.80	0.85	0.88	
Number of Comparisons			55						
Number Differentiated			19						
Not differentiated			36						
Discrimination Power			0.35						

Table 14.16: Results of analyses undertaken using Thin Layer Chromatography

15. Appendix 2 - ANOVA tables for Carbon Homogeneity Experiments

15.1 Single Ream Homogeneity Study

The following tables in this and Appendix 3 were generated using GraphPad Prism. For paper samples, non-parametric test was selected due to the relative inhomogeneity of the material. Further information regarding the interpretation of these results can be found at: <http://www.graphpad.com/guides/prism/6/statistics/>

DOCUCOPY

Kruskal-Wallis test			
P value	0.0003		
Exact or approximate P value?	Gaussian Approximation		
P value summary	***		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	6		
Kruskal-Wallis statistic	23.3		
Dunn's Multiple Comparison Test	Difference in rank sum	Significant? P < 0.05?	Summary
IP_IL_3 vs IP_RL_3	4.67	No	ns
IP_IL_3 vs IP_IL_7	27.3	No	ns
IP_IL_3 vs IP_RL_7	14.9	No	ns
IP_IL_3 vs 7P_I	-1.52	No	ns
IP_IL_3 vs 25P_I	22.3	No	ns
IP_RL_3 vs IP_IL_7	22.7	No	ns
IP_RL_3 vs IP_RL_7	10.2	No	ns
IP_RL_3 vs 7P_I	-6.19	No	ns
IP_RL_3 vs 25P_I	17.6	No	ns
IP_IL_7 vs IP_RL_7	-12.4	No	ns
IP_IL_7 vs 7P_I	-28.9	Yes	**
IP_IL_7 vs 25P_I	-5.08	No	ns
IP_RL_7 vs 7P_I	-16.4	No	ns
IP_RL_7 vs 25P_I	7.35	No	ns
7P_I vs 25P_I	23.8	Yes	**

DOUBLE A

Kruskal-Wallis test			
P value	< 0.0001		
Exact or approximate P value?	Gaussian Approximation		
P value summary	****		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	6		
Kruskal-Wallis statistic	38.43		
Dunn's Multiple Comparison Test	Difference in rank sum	Significant? P < 0.05?	Summary
IP_IL_3 vs IP_RL_3	-21.67	No	ns
IP_IL_3 vs IP_IL_7	-20.52	No	ns
IP_IL_3 vs IP_RL_7	-12.38	No	ns
IP_IL_3 vs 7P_I	-0.6667	No	ns
IP_IL_3 vs 25P_I	11.65	No	ns
IP_RL_3 vs IP_IL_7	1.143	No	ns
IP_RL_3 vs IP_RL_7	9.286	No	ns
IP_RL_3 vs 7P_I	21	No	ns
IP_RL_3 vs 25P_I	33.32	Yes	**
IP_IL_7 vs IP_RL_7	8.143	No	ns
IP_IL_7 vs 7P_I	19.86	No	ns
IP_IL_7 vs 25P_I	32.18	Yes	***
IP_RL_7 vs 7P_I	11.71	No	ns
IP_RL_7 vs 25P_I	24.03	Yes	**
7P_I vs 25P_I	12.32	No	ns

FUJI XEROX

Kruskal-Wallis test			
P value	< 0.0001		
Exact or approximate P value?	Gaussian Approximation		
P value summary	****		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	6		
Kruskal-Wallis statistic	31.35		
Dunn's Multiple Comparison Test	Difference in rank sum	Significant? P < 0.05?	Summary
IP_IL_3 vs IP_RL_3	21	No	ns
IP_IL_3 vs IP_IL_7	18.86	No	ns
IP_IL_3 vs IP_RL_7	-10.5	No	ns
IP_IL_3 vs 7P_I	-12.29	No	ns
IP_IL_3 vs 25P_I	11.63	No	ns
IP_RL_3 vs IP_IL_7	-2.143	No	ns
IP_RL_3 vs IP_RL_7	-31.5	Yes	*
IP_RL_3 vs 7P_I	-33.29	Yes	*
IP_RL_3 vs 25P_I	-9.375	No	ns
IP_IL_7 vs IP_RL_7	-29.36	Yes	**
IP_IL_7 vs 7P_I	-31.14	Yes	***
IP_IL_7 vs 25P_I	-7.232	No	ns
IP_RL_7 vs 7P_I	-1.786	No	ns
IP_RL_7 vs 25P_I	22.13	Yes	*
7P_I vs 25P_I	23.91	Yes	**

HP EVERYDAY

Kruskal-Wallis test			
P value	0.0861		
Exact or approximate P value?	Gaussian Approximation		
P value summary	ns		
Do the medians vary signif. (P < 0.05)	No		
Number of groups	6		
Kruskal-Wallis statistic	9.64		
Dunn's Multiple Comparison Test			
	Difference in rank sum	Significant? P < 0.05?	Summary
IP_IL_3 vs IP_RL_3	30.33	No	ns
IP_IL_3 vs IP_IL_7	8.69	No	ns
IP_IL_3 vs IP_RL_7	1.833	No	ns
IP_IL_3 vs 7P_I	7.119	No	ns
IP_IL_3 vs 25P_I	3.553	No	ns
IP_RL_3 vs IP_IL_7	-21.64	No	ns
IP_RL_3 vs IP_RL_7	-28.5	No	ns
IP_RL_3 vs 7P_I	-23.21	No	ns
IP_RL_3 vs 25P_I	-26.78	No	ns
IP_IL_7 vs IP_RL_7	-6.857	No	ns
IP_IL_7 vs 7P_I	-1.571	No	ns
IP_IL_7 vs 25P_I	-5.137	No	ns
IP_RL_7 vs 7P_I	5.286	No	ns
IP_RL_7 vs 25P_I	1.72	No	ns
7P_I vs 25P_I	-3.566	No	ns

LAZER IT

Kruskal-Wallis test			
P value	0.0006		
Exact or approximate P value?	Gaussian Approximation		
P value summary	***		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	6		
Kruskal-Wallis statistic	21.52		
Dunn's Multiple Comparison Test	Difference in rank sum	Significant? P < 0.05?	Summary
IP_IL_3 vs IP_RL_3	-8.167	No	ns
IP_IL_3 vs IP_IL_7	-2.857	No	ns
IP_IL_3 vs IP_RL_7	2.5	No	ns
IP_IL_3 vs 7P_I	5.429	No	ns
IP_IL_3 vs 25P_I	18.4	No	ns
IP_RL_3 vs IP_IL_7	5.31	No	ns
IP_RL_3 vs IP_RL_7	10.67	No	ns
IP_RL_3 vs 7P_I	13.6	No	ns
IP_RL_3 vs 25P_I	26.56	Yes	*
IP_IL_7 vs IP_RL_7	5.357	No	ns
IP_IL_7 vs 7P_I	8.286	No	ns
IP_IL_7 vs 25P_I	21.25	Yes	*
IP_RL_7 vs 7P_I	2.929	No	ns
IP_RL_7 vs 25P_I	15.9	No	ns
7P_I vs 25P_I	12.97	No	ns

OFFICEMAX

Kruskal-Wallis test			
P value	< 0.0001		
Exact or approximate P value?	Gaussian Approximation		
P value summary	****		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	6		
Kruskal-Wallis statistic	30.69		
Dunn's Multiple Comparison Test	Difference in rank sum	Significant? P < 0.05?	Summary
IP_IL_3 vs IP_RL_3	-1	No	ns
IP_IL_3 vs IP_IL_7	9.333	No	ns
IP_IL_3 vs IP_RL_7	11.9	No	ns
IP_IL_3 vs 7P_I	-5.095	No	ns
IP_IL_3 vs 25P_I	-16.63	No	ns
IP_RL_3 vs IP_IL_7	10.33	No	ns
IP_RL_3 vs IP_RL_7	12.9	No	ns
IP_RL_3 vs 7P_I	-4.095	No	ns
IP_RL_3 vs 25P_I	-15.63	No	ns
IP_IL_7 vs IP_RL_7	2.571	No	ns
IP_IL_7 vs 7P_I	-14.43	No	ns
IP_IL_7 vs 25P_I	-25.96	Yes	***
IP_RL_7 vs 7P_I	-17	No	ns
IP_RL_7 vs 25P_I	-28.53	Yes	***
7P_I vs 25P_I	-11.53	No	ns

REFLEX

Kruskal-Wallis test			
P value	< 0.0001		
Exact or approximate P value?	Gaussian Approximation		
P value summary	****		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	6		
Kruskal-Wallis statistic	36.1		
Dunn's Multiple Comparison Test	Difference in rank sum	Significant? P < 0.05?	Summary
IP_IL_3 vs IP_RL_3	3.333	No	ns
IP_IL_3 vs IP_IL_7	-0.7619	No	ns
IP_IL_3 vs IP_RL_7	11.95	No	ns
IP_IL_3 vs 7P_I	19.38	No	ns
IP_IL_3 vs 25P_I	30.33	Yes	*
IP_RL_3 vs IP_IL_7	-4.095	No	ns
IP_RL_3 vs IP_RL_7	8.619	No	ns
IP_RL_3 vs 7P_I	16.05	No	ns
IP_RL_3 vs 25P_I	27	Yes	*
IP_IL_7 vs IP_RL_7	12.71	No	ns
IP_IL_7 vs 7P_I	20.14	No	ns
IP_IL_7 vs 25P_I	31.1	Yes	***
IP_RL_7 vs 7P_I	7.429	No	ns
IP_RL_7 vs 25P_I	18.38	No	ns
7P_I vs 25P_I	10.95	No	ns

15.2 Between Ream Homogeneity Study

DOCUCOPY

Kruskal-Wallis test			
P value	0.1497		
Exact or approximate P value?	Approximate		
P value summary	ns		
Do the medians vary signif. (P < 0.05)	No		
Number of groups	7		
Kruskal-Wallis statistic	9.453		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
A vs. B	14.71	No	ns
A vs. C	1.286	No	ns
A vs. D	9.429	No	ns
A vs. E	3.143	No	ns
A vs. F	12.18	No	ns
A vs. G	-2.988	No	ns
B vs. C	-13.43	No	ns
B vs. D	-5.286	No	ns
B vs. E	-11.57	No	ns
B vs. F	-2.536	No	ns
B vs. G	-17.7	No	ns
C vs. D	8.143	No	ns
C vs. E	1.857	No	ns
C vs. F	10.89	No	ns
C vs. G	-4.274	No	ns
D vs. E	-6.286	No	ns
D vs. F	2.75	No	ns
D vs. G	-12.42	No	ns
E vs. F	9.036	No	ns
E vs. G	-6.131	No	ns
F vs. G	-15.17	No	ns

DOUBLE A

Kruskal-Wallis test			
P value	< 0.0001		
Exact or approximate P value?	Approximate		
P value summary	****		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	7		
Kruskal-Wallis statistic	29.99		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
A vs. B	2.19	No	ns
A vs. C	1.524	No	ns
A vs. D	5.857	No	ns
A vs. E	7.143	No	ns
A vs. F	27	Yes	**
A vs. G	25.71	Yes	**
B vs. C	-0.6667	No	ns
B vs. D	3.667	No	ns
B vs. E	4.952	No	ns
B vs. F	24.81	Yes	*
B vs. G	23.52	Yes	*
C vs. D	4.333	No	ns
C vs. E	5.619	No	ns
C vs. F	25.48	Yes	*
C vs. G	24.19	Yes	*
D vs. E	1.286	No	ns
D vs. F	21.14	No	ns
D vs. G	19.86	No	ns
E vs. F	19.86	No	ns
E vs. G	18.57	No	ns
F vs. G	-1.286	No	ns

FUJI XEROX

Kruskal-Wallis test				
P value	0.3116			
Exact or approximate P value?	Approximate			
P value summary	ns			
Do the medians vary signif. (P < 0.05)	No			
Number of groups	7			
Kruskal-Wallis statistic	7.101			
Dunn's multiple comparisons test				
	Mean rank diff.	Significant?	Summary	Adjusted P Value
A vs. B	13.36	No	ns	> 0.9999
A vs. C	11.5	No	ns	> 0.9999
A vs. D	9.357	No	ns	> 0.9999
A vs. E	4.643	No	ns	> 0.9999
A vs. F	6.857	No	ns	> 0.9999
A vs. G	17.79	No	ns	0.418
B vs. C	-1.857	No	ns	> 0.9999
B vs. D	-4	No	ns	> 0.9999
B vs. E	-8.714	No	ns	> 0.9999
B vs. F	-6.5	No	ns	> 0.9999
B vs. G	4.429	No	ns	> 0.9999
C vs. D	-2.143	No	ns	> 0.9999
C vs. E	-6.857	No	ns	> 0.9999
C vs. F	-4.643	No	ns	> 0.9999
C vs. G	6.286	No	ns	> 0.9999
D vs. E	-4.714	No	ns	> 0.9999
D vs. F	-2.5	No	ns	> 0.9999
D vs. G	8.429	No	ns	> 0.9999
E vs. F	2.214	No	ns	> 0.9999
E vs. G	13.14	No	ns	> 0.9999
F vs. G	10.93	No	ns	> 0.9999

LAZER IT

Kruskal-Wallis test			
P value	0.1312		
Exact or approximate P value?	Approximate		
P value summary	ns		
Do the medians vary signif. (P < 0.05)	No		
Number of groups	7		
Kruskal-Wallis statistic	9.848		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
A vs. B	12.19	No	ns
A vs. C	20.69	No	ns
A vs. D	11.65	No	ns
A vs. E	16.26	No	ns
A vs. F	20.54	No	ns
A vs. G	18.26	No	ns
B vs. C	8.5	No	ns
B vs. D	-0.5357	No	ns
B vs. E	4.071	No	ns
B vs. F	8.357	No	ns
B vs. G	6.071	No	ns
C vs. D	-9.036	No	ns
C vs. E	-4.429	No	ns
C vs. F	-0.1429	No	ns
C vs. G	-2.429	No	ns
D vs. E	4.607	No	ns
D vs. F	8.893	No	ns
D vs. G	6.607	No	ns
E vs. F	4.286	No	ns
E vs. G	2	No	ns
F vs. G	-2.286	No	ns

OFFICEMAX

Kruskal-Wallis test			
P value	< 0.0001		
Exact or approximate P value?	Approximate		
P value summary	****		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	7		
Kruskal-Wallis statistic	31.3		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
A vs. B	6.929	No	ns
A vs. C	7	No	ns
A vs. D	0.03571	No	ns
A vs. E	14.14	No	ns
A vs. F	-17.88	No	ns
A vs. G	-16.55	No	ns
B vs. C	0.07143	No	ns
B vs. D	-6.893	No	ns
B vs. E	7.214	No	ns
B vs. F	-24.81	Yes	*
B vs. G	-23.48	Yes	*
C vs. D	-6.964	No	ns
C vs. E	7.143	No	ns
C vs. F	-24.88	Yes	*
C vs. G	-23.55	Yes	*
D vs. E	14.11	No	ns
D vs. F	-17.92	No	ns
D vs. G	-16.58	No	ns
E vs. F	-32.02	Yes	***
E vs. G	-30.69	Yes	***
F vs. G	1.333	No	ns

OLYMPIC

Kruskal-Wallis test			
P value	0.6551		
Exact or approximate P value?	Approximate		
P value summary	ns		
Do the medians vary signif. (P < 0.05)	No		
Number of groups	7		
Kruskal-Wallis statistic	4.159		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
A vs. B	7.5	No	ns
A vs. C	11.36	No	ns
A vs. D	2.5	No	ns
A vs. E	3.5	No	ns
A vs. F	11	No	ns
A vs. G	9.071	No	ns
B vs. C	3.857	No	ns
B vs. D	-5	No	ns
B vs. E	-4	No	ns
B vs. F	3.5	No	ns
B vs. G	1.571	No	ns
C vs. D	-8.857	No	ns
C vs. E	-7.857	No	ns
C vs. F	-0.3571	No	ns
C vs. G	-2.286	No	ns
D vs. E	1	No	ns
D vs. F	8.5	No	ns
D vs. G	6.571	No	ns
E vs. F	7.5	No	ns
E vs. G	5.571	No	ns
F vs. G	-1.929	No	ns

REFLEX

Kruskal-Wallis test			
P value	0.007		
Exact or approximate P value?	Approximate		
P value summary	**		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	7		
Kruskal-Wallis statistic	17.7		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
A vs. B	-0.2143	No	ns
A vs. C	4.286	No	ns
A vs. D	-1.714	No	ns
A vs. E	-15.14	No	ns
A vs. F	11.93	No	ns
A vs. G	11.86	No	ns
B vs. C	4.5	No	ns
B vs. D	-1.5	No	ns
B vs. E	-14.93	No	ns
B vs. F	12.14	No	ns
B vs. G	12.07	No	ns
C vs. D	-6	No	ns
C vs. E	-19.43	No	ns
C vs. F	7.643	No	ns
C vs. G	7.571	No	ns
D vs. E	-13.43	No	ns
D vs. F	13.64	No	ns
D vs. G	13.57	No	ns
E vs. F	27.07	Yes	**
E vs. G	27	Yes	**
F vs. G	-0.07143	No	ns

16. Appendix 3 - ANOVA tables for Oxygen Homogeneity Experiments

16.1 Single Ream Homogeneity Study

DOCUCOPY

Kruskal-Wallis test			
P value	< 0.0001		
Exact or approximate P value?	Approximate		
P value summary	****		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	6		
Kruskal-Wallis statistic	33.38		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
3S_1P_1L vs. 7S_1P_1L	5.643	No	ns
3S_1P_1L vs. 3S_1P_RL	10	No	ns
3S_1P_1L vs. 7S_1P_RL	13.83	No	ns
3S_1P_1L vs. 7S_7P	11.07	No	ns
3S_1P_1L vs. 25S_25P	-13.7	No	ns
7S_1P_1L vs. 3S_1P_RL	4.357	No	ns
7S_1P_1L vs. 7S_1P_RL	8.19	No	ns
7S_1P_1L vs. 7S_7P	5.429	No	ns
7S_1P_1L vs. 25S_25P	-19.34	Yes	*
3S_1P_RL vs. 7S_1P_RL	3.833	No	ns
3S_1P_RL vs. 7S_7P	1.071	No	ns
3S_1P_RL vs. 25S_25P	-23.7	No	ns
7S_1P_RL vs. 7S_7P	-2.762	No	ns
7S_1P_RL vs. 25S_25P	-27.53	Yes	***
7S_7P vs. 25S_25P	-24.77	Yes	***

FUJI XEROX

Kruskal-Wallis test			
P value	< 0.0001		
Exact or approximate P value?	Approximate		
P value summary	****		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	6		
Kruskal-Wallis statistic	27.03		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
3S_1P_1L vs. 7S_1P_1L	-8.667	No	ns
3S_1P_1L vs. 3S_1P_RL	-8.833	No	ns
3S_1P_1L vs. 7S_1P_RL	-18.08	No	ns
3S_1P_1L vs. 7S_7P	-18.17	No	ns
3S_1P_1L vs. 25S_25P	6.625	No	ns
7S_1P_1L vs. 3S_1P_RL	-0.1667	No	ns
7S_1P_1L vs. 7S_1P_RL	-9.417	No	ns
7S_1P_1L vs. 7S_7P	-9.5	No	ns
7S_1P_1L vs. 25S_25P	15.29	No	ns
3S_1P_RL vs. 7S_1P_RL	-9.25	No	ns
3S_1P_RL vs. 7S_7P	-9.333	No	ns
3S_1P_RL vs. 25S_25P	15.46	No	ns
7S_1P_RL vs. 7S_7P	-0.08333	No	ns
7S_1P_RL vs. 25S_25P	24.71	Yes	**
7S_7P vs. 25S_25P	24.79	Yes	***

HP EVERYDAY

Kruskal-Wallis test			
P value	< 0.0001		
Exact or approximate P value?	Approximate		
P value summary	****		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	5		
Kruskal-Wallis statistic	34.14		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
3S_1P_1L vs. 7S_1P_1L	-2.762	No	ns
3S_1P_1L vs. 7S_1P_RL	4.381	No	ns
3S_1P_1L vs. 1S_7P	2.583	No	ns
3S_1P_1L vs. 1S_25P	23.67	Yes	*
7S_1P_1L vs. 7S_1P_RL	7.143	No	ns
7S_1P_1L vs. 1S_7P	5.345	No	ns
7S_1P_1L vs. 1S_25P	26.43	Yes	****
7S_1P_RL vs. 1S_7P	-1.798	No	ns
7S_1P_RL vs. 1S_25P	19.29	Yes	**
1S_7P vs. 1S_25P	21.08	Yes	**

LAZER IT

Kruskal-Wallis test			
P value	0.0008		
Exact or approximate P value?	Approximate		
P value summary	***		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	6		
Kruskal-Wallis statistic	21.02		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
3S_1P_1L vs. 7S_1P_1L	-5.667	No	ns
3S_1P_1L vs. 3S_1P_RL	-7.333	No	ns
3S_1P_1L vs. 7S_1P_RL	-11.17	No	ns
3S_1P_1L vs. 7S_7P	-2.738	No	ns
3S_1P_1L vs. 25S_25P	-23.79	No	ns
7S_1P_1L vs. 3S_1P_RL	-1.667	No	ns
7S_1P_1L vs. 7S_1P_RL	-5.5	No	ns
7S_1P_1L vs. 7S_7P	2.929	No	ns
7S_1P_1L vs. 25S_25P	-18.13	No	ns
3S_1P_RL vs. 7S_1P_RL	-3.833	No	ns
3S_1P_RL vs. 7S_7P	4.595	No	ns
3S_1P_RL vs. 25S_25P	-16.46	No	ns
7S_1P_RL vs. 7S_7P	8.429	No	ns
7S_1P_RL vs. 25S_25P	-12.63	No	ns
7S_7P vs. 25S_25P	-21.05	Yes	**

OFFICEMAX

Kruskal-Wallis test			
P value	< 0.0001		
Exact or approximate P value?	Approximate		
P value summary	****		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	6		
Kruskal-Wallis statistic	27.73		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
3S_1P_1L vs. 7S_1P_1L	-5.75	No	ns
3S_1P_1L vs. 3S_1P_RL	10.5	No	ns
3S_1P_1L vs. 7S_1P_RL	-2.381	No	ns
3S_1P_1L vs. 7S_7P	0.3333	No	ns
3S_1P_1L vs. 25S_25P	-21.17	No	ns
7S_1P_1L vs. 3S_1P_RL	16.25	No	ns
7S_1P_1L vs. 7S_1P_RL	3.369	No	ns
7S_1P_1L vs. 7S_7P	6.083	No	ns
7S_1P_1L vs. 25S_25P	-15.42	No	ns
3S_1P_RL vs. 7S_1P_RL	-12.88	No	ns
3S_1P_RL vs. 7S_7P	-10.17	No	ns
3S_1P_RL vs. 25S_25P	-31.67	Yes	**
7S_1P_RL vs. 7S_7P	2.714	No	ns
7S_1P_RL vs. 25S_25P	-18.79	Yes	*
7S_7P vs. 25S_25P	-21.5	Yes	*

PAPER ONE

Kruskal-Wallis test			
P value	0.9416		
Exact or approximate P value?	Approximate		
P value summary	ns		
Do the medians vary signif. (P < 0.05)	No		
Number of groups	6		
Kruskal-Wallis statistic	1.234		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
3S_1P_1L vs. 7S_1P_1L	6	No	ns
3S_1P_1L vs. 3S_1P_RL	-3.667	No	ns
3S_1P_1L vs. 7S_1P_RL	3	No	ns
3S_1P_1L vs. 7S_7P	1.667	No	ns
3S_1P_1L vs. 25S_25P	0.1467	No	ns
7S_1P_1L vs. 3S_1P_RL	-9.667	No	ns
7S_1P_1L vs. 7S_1P_RL	-3	No	ns
7S_1P_1L vs. 7S_7P	-4.333	No	ns
7S_1P_1L vs. 25S_25P	-5.853	No	ns
3S_1P_RL vs. 7S_1P_RL	6.667	No	ns
3S_1P_RL vs. 7S_7P	5.333	No	ns
3S_1P_RL vs. 25S_25P	3.813	No	ns
7S_1P_RL vs. 7S_7P	-1.333	No	ns
7S_1P_RL vs. 25S_25P	-2.853	No	ns
7S_7P vs. 25S_25P	-1.52	No	ns

REFLEX

Kruskal-Wallis test			
P value	0.0023		
Exact or approximate P value?	Approximate		
P value summary	**		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	6		
Kruskal-Wallis statistic	18.54		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
3S_1P_1L vs. 7S_1P_1L	-19.12	No	ns
3S_1P_1L vs. 3S_1P_RL	-3	No	ns
3S_1P_1L vs. 7S_1P_RL	-28.26	No	ns
3S_1P_1L vs. 7S_7P	-29.62	No	ns
3S_1P_1L vs. 25S_25P	-11.88	No	ns
7S_1P_1L vs. 3S_1P_RL	16.12	No	ns
7S_1P_1L vs. 7S_1P_RL	-9.143	No	ns
7S_1P_1L vs. 7S_7P	-10.5	No	ns
7S_1P_1L vs. 25S_25P	7.244	No	ns
3S_1P_RL vs. 7S_1P_RL	-25.26	No	ns
3S_1P_RL vs. 7S_7P	-26.62	No	ns
3S_1P_RL vs. 25S_25P	-8.875	No	ns
7S_1P_RL vs. 7S_7P	-1.357	No	ns
7S_1P_RL vs. 25S_25P	16.39	No	ns
7S_7P vs. 25S_25P	17.74	No	ns

16.2 Between Ream Homogeneity Study

DOCUCOPY

Kruskal-Wallis test			
P value	0.2703		
Exact or approximate P value?	Approximate		
P value summary	ns		
Do the medians vary signif. (P < 0.05)	No		
Number of groups	7		
Kruskal-Wallis statistic	7.583		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
A vs. B	8.405	No	ns
A vs. C	11.98	No	ns
A vs. D	14.67	No	ns
A vs. E	17.63	No	ns
A vs. F	14.69	No	ns
A vs. G	15.83	No	ns
B vs. C	3.571	No	ns
B vs. D	6.262	No	ns
B vs. E	9.229	No	ns
B vs. F	6.286	No	ns
B vs. G	7.429	No	ns
C vs. D	2.69	No	ns
C vs. E	5.657	No	ns
C vs. F	2.714	No	ns
C vs. G	3.857	No	ns
D vs. E	2.967	No	ns
D vs. F	0.02381	No	ns
D vs. G	1.167	No	ns
E vs. F	-2.943	No	ns
E vs. G	-1.8	No	ns
F vs. G	1.143	No	ns

DOUBLE A

Kruskal-Wallis test			
P value	0.0002		
Exact or approximate P value?	Approximate		
P value summary	***		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	7		
Kruskal-Wallis statistic	26.89		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
A vs. B	2.333	No	ns
A vs. C	5.583	No	ns
A vs. D	8.333	No	ns
A vs. E	10.83	No	ns
A vs. F	-9.667	No	ns
A vs. G	19.83	Yes	*
B vs. C	3.25	No	ns
B vs. D	6	No	ns
B vs. E	8.5	No	ns
B vs. F	-12	No	ns
B vs. G	17.5	No	ns
C vs. D	2.75	No	ns
C vs. E	5.25	No	ns
C vs. F	-15.25	No	ns
C vs. G	14.25	No	ns
D vs. E	2.5	No	ns
D vs. F	-18	No	ns
D vs. G	11.5	No	ns
E vs. F	-20.5	No	ns
E vs. G	9	No	ns
F vs. G	29.5	Yes	****

FUJI XEROX

Kruskal-Wallis test			
P value	0.0027		
Exact or approximate P value?	Approximate		
P value summary	**		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	7		
Kruskal-Wallis statistic	20.05		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
A vs. B	15.45	No	ns
A vs. C	14.5	No	ns
A vs. D	21.77	No	ns
A vs. E	29.47	Yes	**
A vs. F	21.58	Yes	*
A vs. G	20.5	No	ns
B vs. C	-0.9524	No	ns
B vs. D	6.314	No	ns
B vs. E	14.01	No	ns
B vs. F	6.131	No	ns
B vs. G	5.048	No	ns
C vs. D	7.267	No	ns
C vs. E	14.97	No	ns
C vs. F	7.083	No	ns
C vs. G	6	No	ns
D vs. E	7.7	No	ns
D vs. F	-0.1833	No	ns
D vs. G	-1.267	No	ns
E vs. F	-7.883	No	ns
E vs. G	-8.967	No	ns
F vs. G	-1.083	No	ns

LAZER IT

Kruskal-Wallis test			
P value	0.2523		
Exact or approximate P value?	Approximate		
P value summary	ns		
Do the medians vary signif. (P < 0.05)	No		
Number of groups	7		
Kruskal-Wallis statistic	7.811		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
A vs. B	-0.1905	No	ns
A vs. C	-7.429	No	ns
A vs. D	1.286	No	ns
A vs. E	7.429	No	ns
A vs. F	-1	No	ns
A vs. G	11.14	No	ns
B vs. C	-7.238	No	ns
B vs. D	1.476	No	ns
B vs. E	7.619	No	ns
B vs. F	-0.8095	No	ns
B vs. G	11.33	No	ns
C vs. D	8.714	No	ns
C vs. E	14.86	No	ns
C vs. F	6.429	No	ns
C vs. G	18.57	No	ns
D vs. E	6.143	No	ns
D vs. F	-2.286	No	ns
D vs. G	9.857	No	ns
E vs. F	-8.429	No	ns
E vs. G	3.714	No	ns
F vs. G	12.14	No	ns

OFFICEMAX

Kruskal-Wallis test			
P value	< 0.0001		
Exact or approximate P value?	Approximate		
P value summary	****		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	7		
Kruskal-Wallis statistic	30.91		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
A vs. B	6.5	No	ns
A vs. C	4.929	No	ns
A vs. D	8.5	No	ns
A vs. E	9.214	No	ns
A vs. F	29.43	Yes	**
A vs. G	30.57	Yes	**
B vs. C	-1.571	No	ns
B vs. D	2	No	ns
B vs. E	2.714	No	ns
B vs. F	22.93	Yes	*
B vs. G	24.07	Yes	*
C vs. D	3.571	No	ns
C vs. E	4.286	No	ns
C vs. F	24.5	Yes	*
C vs. G	25.64	Yes	*
D vs. E	0.7143	No	ns
D vs. F	20.93	No	ns
D vs. G	22.07	No	ns
E vs. F	20.21	No	ns
E vs. G	21.36	No	ns
F vs. G	1.143	No	ns

OLYMPIC

Kruskal-Wallis test			
P value	0.0045		
Exact or approximate P value?	Approximate		
P value summary	**		
Do the medians vary signif. (P < 0.05)	Yes		
Number of groups	7		
Kruskal-Wallis statistic	18.8		
Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary
A vs. B	16	No	ns
A vs. C	14.36	No	ns
A vs. D	15.42	No	ns
A vs. E	25.67	Yes	**
A vs. F	24.17	Yes	*
A vs. G	23.67	Yes	*
B vs. C	-1.643	No	ns
B vs. D	-0.5833	No	ns
B vs. E	9.667	No	ns
B vs. F	8.167	No	ns
B vs. G	7.667	No	ns
C vs. D	1.06	No	ns
C vs. E	11.31	No	ns
C vs. F	9.81	No	ns
C vs. G	9.31	No	ns
D vs. E	10.25	No	ns
D vs. F	8.75	No	ns
D vs. G	8.25	No	ns
E vs. F	-1.5	No	ns
E vs. G	-2	No	ns
F vs. G	-0.5	No	ns

17. Appendix 4 – Proposed Technical Report Appendix

The following section provides an example of what should be included either within the body of the report or referenced in the report and included as an appendix. The appendix should be tailored as appropriate to the tests conducted in each case.

Overview of Paper Examination Procedures

A range of examinations were conducted with the aim of determining whether two or more papers did or did not originate from the same source. The tests utilise range in complexity from simple physical and light examinations to more complex chemical testing. In this case, the tests conducted include:

- Physical Examinations – include measuring the weight, dimensions and thickness of the paper sample and subsequently calculating the grammage (grams per metre²) and sheet density (thickness per kg³) of the sheets. Once calculated, these attributes are compared for similarities or differences.
- Light Examinations – include comparing the reaction of the paper samples to ultraviolet light, transmitted light and observing the light in normal (white) light conditions.
- Chemical - Thin Layer Chromatography (TLC) – the dyes and brighteners included in the papers are extracted and spotted on to a silica backed aluminium plate, which is dried and a mobile (liquid) phase is allowed to migrate up the plate, separating the dyes and brighteners into their core components. These components are then compared to determine whether the same or different dyes and brighteners were used in paper production.
- Chemical - Isotope Ratio Mass Spectrometry (IRMS) – a highly discriminating technique where samples are punched from the paper sheets and measured to determine the stable isotope ratios of oxygen and carbon within the sample. The isotopes, defined as naturally occurring forms of the elements oxygen and carbon that differ in their atomic weight, are naturally occurring and vary depending on the source of the material used to produce the paper. This technique has been shown to be highly discriminating for naturally occurring materials, and for paper in particular.

At the conclusion of testing, the examiner weighs the results to determine whether the evidence supports one proposition over the other. The two propositions being tested are:

Proposition 1: The specimen and questioned papers originated from the same source

Proposition 2: The specimen and questioned papers originated from a different source

As papers are a manufactured material, even using highly discriminating techniques such as the ones here, the possibility that more than one source has produced a paper of the same physical, optical, chemical and isotopic properties cannot be excluded. Additionally, the size of a source or batch produced in the same facility cannot be defined. Due to this, when no difference is observed between questioned and specimen papers, in general, it is not possible to weigh the proposition in terms of the paper having solely originated from the same source.

The weight of the evidence should be scaled to formulate one of the following opinions:

- There is support for the proposition that the specimen and questioned papers originated from the same source, or from a source with the same physical, optical, chemical and isotopic properties.
- The evidence favours neither proposition and the results are inconclusive.
- There is support for the proposition that the specimen and questioned papers originated from different sources.
- There is strong support for the proposition that the specimen and questioned papers originated from different sources.

18. Appendix 5 - Electronic Files on Included Disc

18.1 Comparison Tables

18.1.1 Carbon Isotopes

18.1.2 Oxygen Isotopes

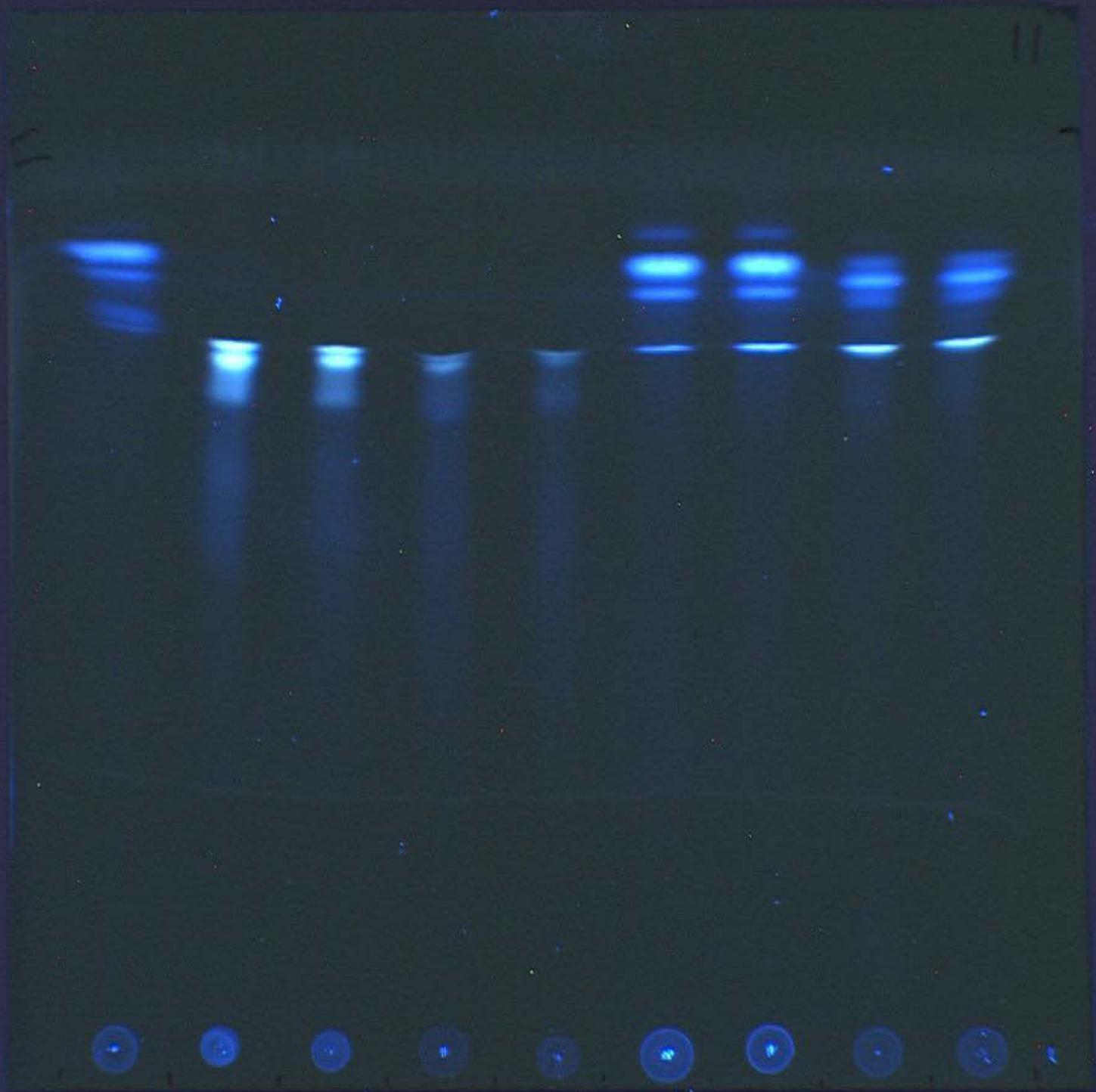
18.1.3 Hydrogen Isotopes

18.1.4 Apparent Sheet Density and Grammage

18.2 Thin Layer Chromatography Plate Images

18.3 Ambient, UV and Transmitted Light Images

18.4 Copy of Published Work

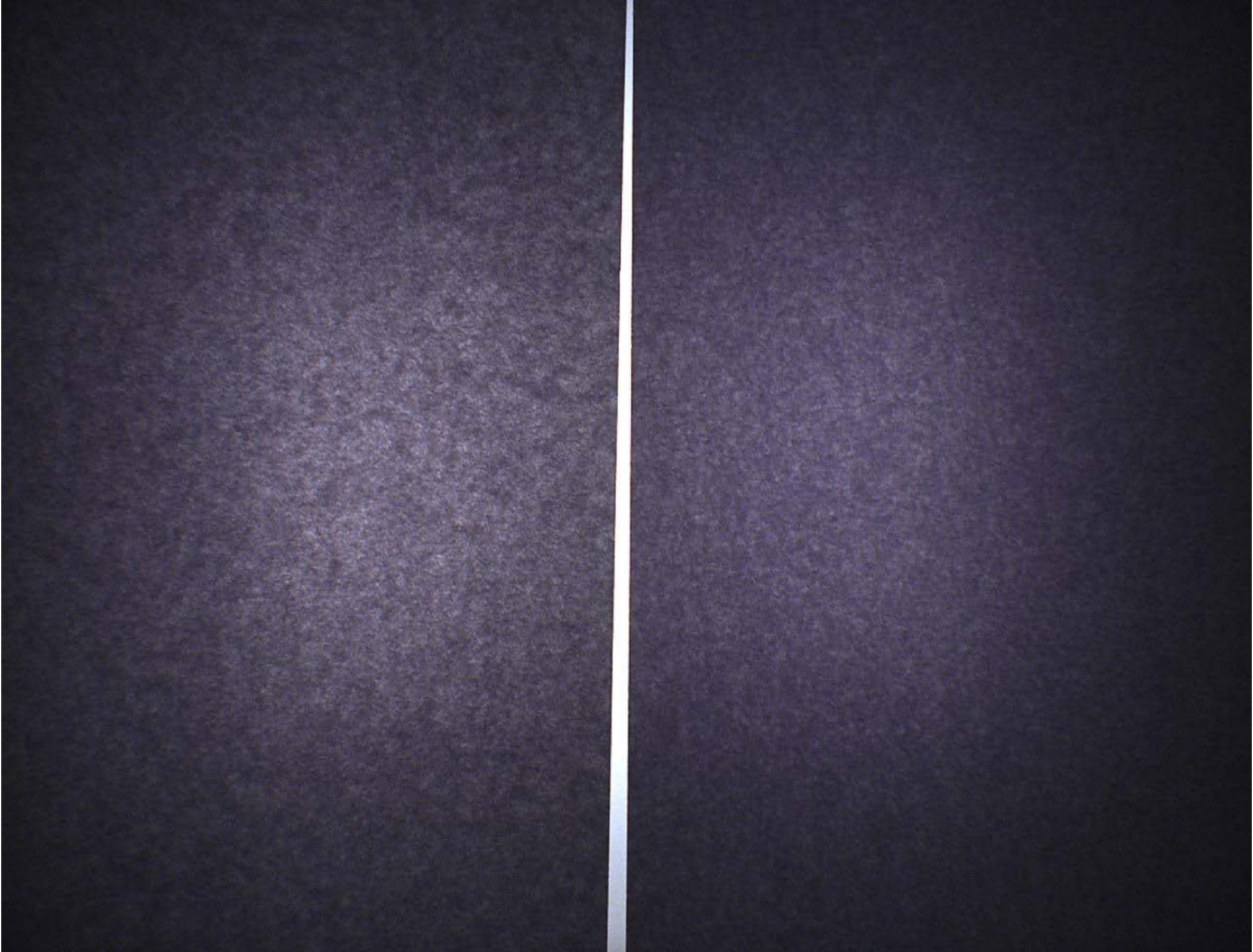


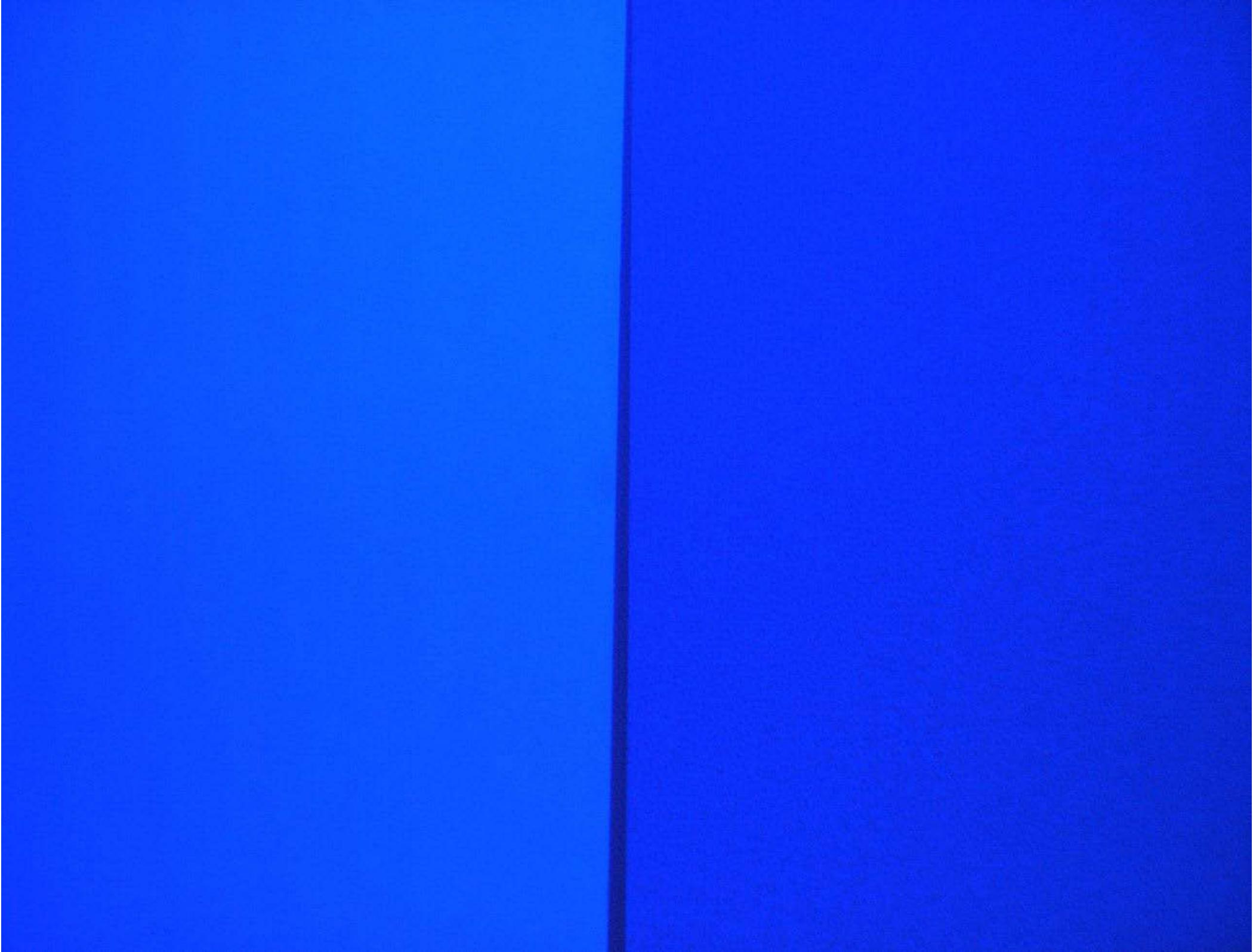
TLC Results, with plate references

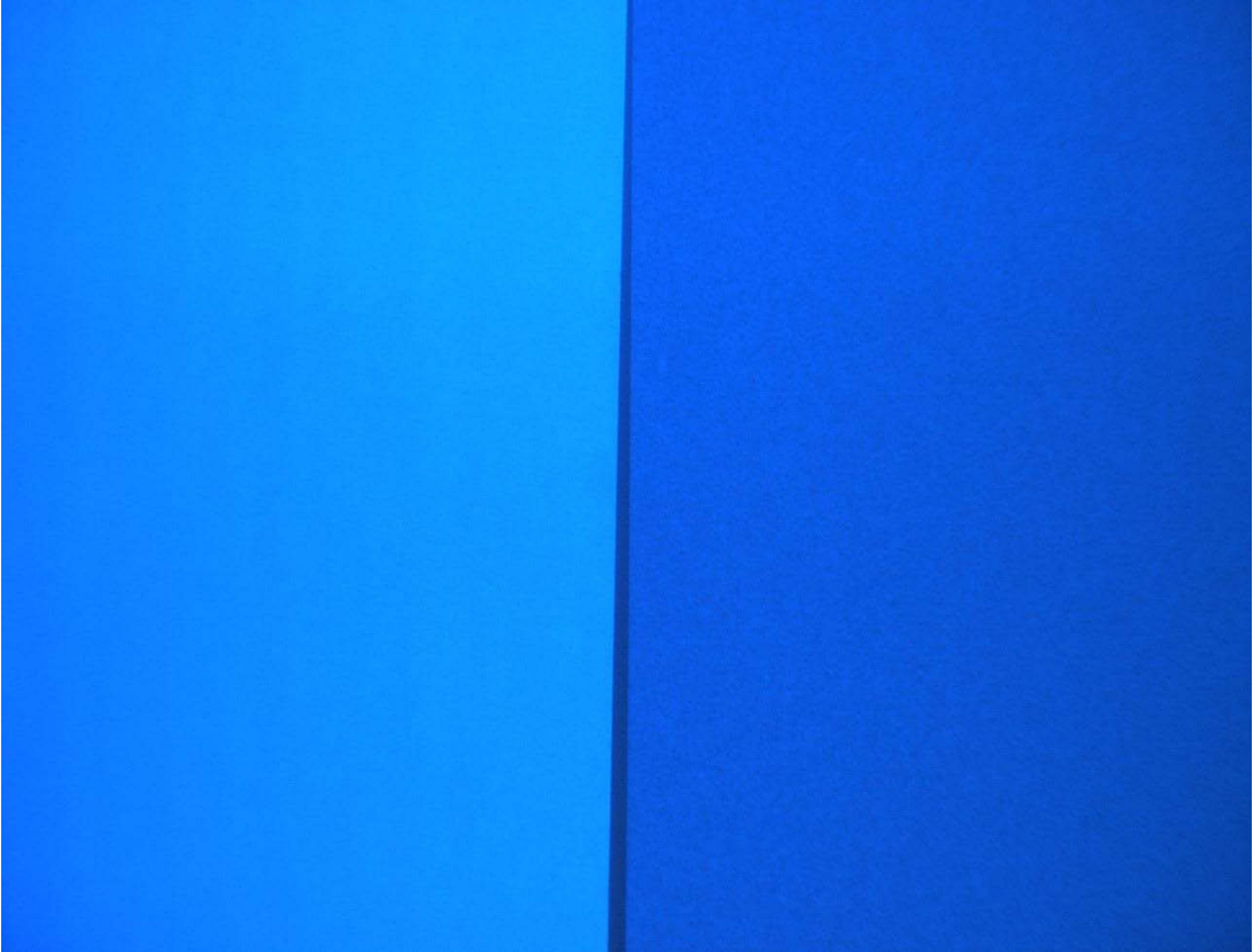
Plate Number	Sample Number	Measurement			Solvent Front	Calc Rf of Peaks			Differentiate?
		Peak 1	Peak 2	Peak 3		RF Comp 1	RF Comp 2	RF Comp 3	
2	71	47.6	52.7	54.6	67.5	0.71	0.78	0.81	
	43	53.1	54.7			0.79	0.81		n
	67	52.8	54.6			0.78	0.81		n
	68	53.0	54.8			0.79	0.81		y
	74	53.5	55.0			0.79	0.81		n
3	71	53.2	56.2		64.7	0.82	0.87		
	75	52.6	55.4	56.4		0.81	0.86	0.87	n
	76	52.8	55.9	56.4		0.82	0.86	0.87	n
	77	52.6	55.5	56.2		0.81	0.86	0.87	n
	78	52.2	55.4	56.3		0.81	0.86	0.87	n
5	71	52.9	55.2	57.7	67.5	0.78	0.82	0.85	
	84	56.5	58.3			0.84	0.86		y
	86	51.1	56.6	58		0.76	0.84	0.86	n
	87	51.0	56.1	58.1		0.76	0.83	0.86	n
	88	54.8	56.5			0.81	0.84		y
6	71	50.1	54.7	56.8	66.5	0.75	0.82	0.85	
	88	55.2	57			0.83	0.86		y
	89	55.2	57			0.83	0.86		y
	92	54.2	55.9			0.82	0.84		y
	93	56.5	58.6			0.85	0.88		n
7	71	51.5	56.1	57.9	68.2	0.76	0.82	0.85	
	94	51.5	56.0	58		0.76	0.82	0.85	n
	95	51.5	55.9	57.9		0.76	0.82	0.85	n
	96	56.2	58.2			0.82	0.85		n
	97	56.2	58.2			0.82	0.85		n
9	71	50.5	54.7	56.6	65.1	0.78	0.84	0.87	
	103	52.9	54.9			0.81	0.84		y
	104	52.9	54.9			0.81	0.84		y
10	71	49.7	53.6	55.3	62.2	0.80	0.86	0.89	
	105	48.9				0.79			y
	106	48.9				0.79			y
	107	48.9	52.9	55.3		0.79	0.85	0.89	n
	109	48.8	53.8	55.3		0.78	0.86	0.89	n
11	71	50.3	53.2	54.9	63.5	0.79	0.84	0.86	
	102	47.1	48.2			0.74	0.76		y
	112	48.5	52.1	54.1		0.76	0.82	0.85	n
12	71	59.1	60.3	61.5	67.1	0.88	0.90	0.92	
	113	59.2	60.6			0.88	0.90		y
	114	54.9	58.2	59.5		0.82	0.87	0.89	y
	P3	54.9	58.2	59.5		0.82	0.87	0.89	y
13	71	59.2	60.3	61.5	66.9	0.88	0.90	0.92	
	115	59.7	61.1			0.89	0.91		n
	116	55.3	59	60.4		0.83	0.88	0.90	y
14	71	46	52.2	54.8	65.8	0.70	0.79	0.83	
	117	44.9	50.6	53.7		0.68	0.77	0.82	n
	118	44.9	50.6	53.7		0.68	0.77	0.82	n
	120	48.8	52.1			0.74	0.79		n
	121	NONE							y
15	71	46.4	52.4	55.5	66.9	0.69	0.78	0.83	
	122	NONE							y
	124	51.1	53.9			0.76	0.81		n
	125	51.1	53.9			0.76	0.81		n
	48	51.1	53.9		0.76	0.81		n	
16	71	55.8	57.8		65.3	0.85	0.89		
	49	54.6	57.3			0.84	0.88		n
	53	52.7	55.1			0.81	0.84		y
17	71	52.8	57.0	58.6	68.3	0.77	0.83	0.86	
	46	51.8	56.4	58.1		0.76	0.83	0.85	n
	72	51.8	56.4	58.1		0.76	0.83	0.85	n
	57	51.8	56.4	58.1		0.76	0.83	0.85	n
	73	51.8	56.4	58.1		0.76	0.83	0.85	n
18	71	52.1	56.3	58.0	67.7	0.77	0.83	0.86	
	69	51.9	56.3	58.3		0.77	0.83	0.86	n
	119	51.9	56.3	58.3		0.77	0.83	0.86	n
	70	51.9	56.3	58.3		0.77	0.83	0.86	n
	47	51.9	56.3	58.3		0.77	0.83	0.86	n
19	71		57.2	59.3	66.9		0.86	0.89	
	57	53.5				0.80			y
	99	53.2	57.1	59.2		0.80	0.85	0.88	n
	54		57.1	59.2			0.85	0.88	n
	51	53.2	57.1	59.2		0.80	0.85	0.88	n

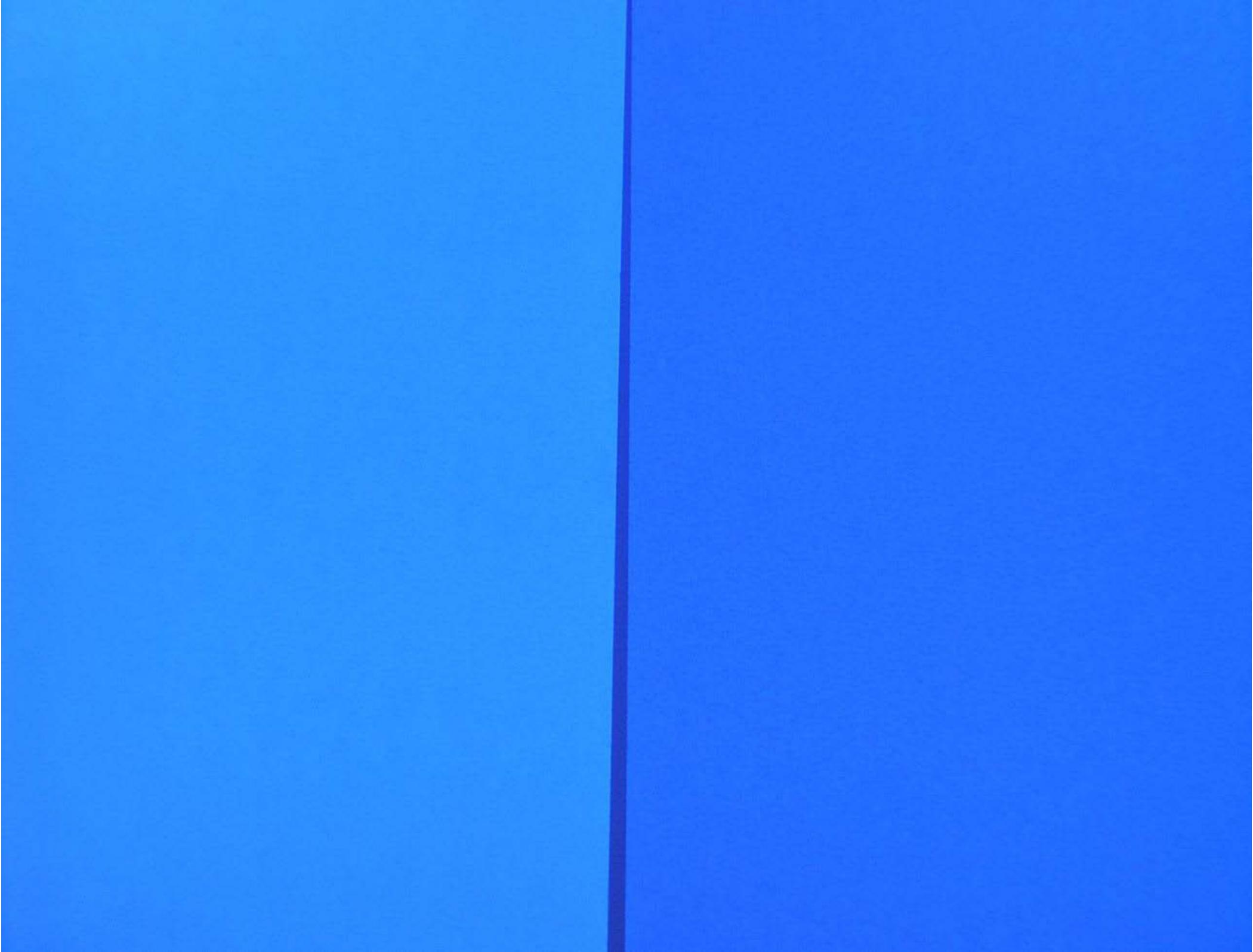
Differentiate?
yes 19
no 36
Comparisons 55

DF 0.35

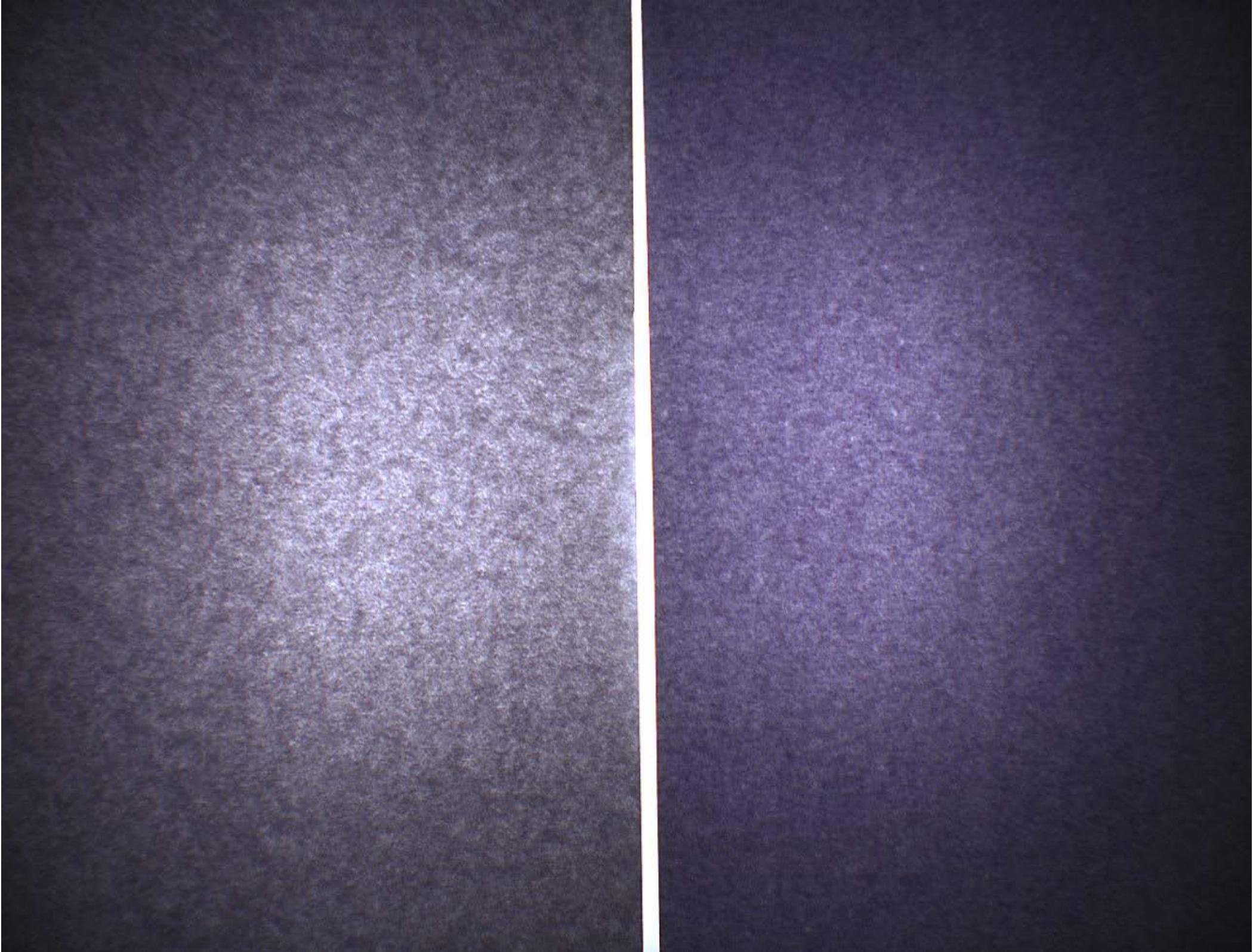


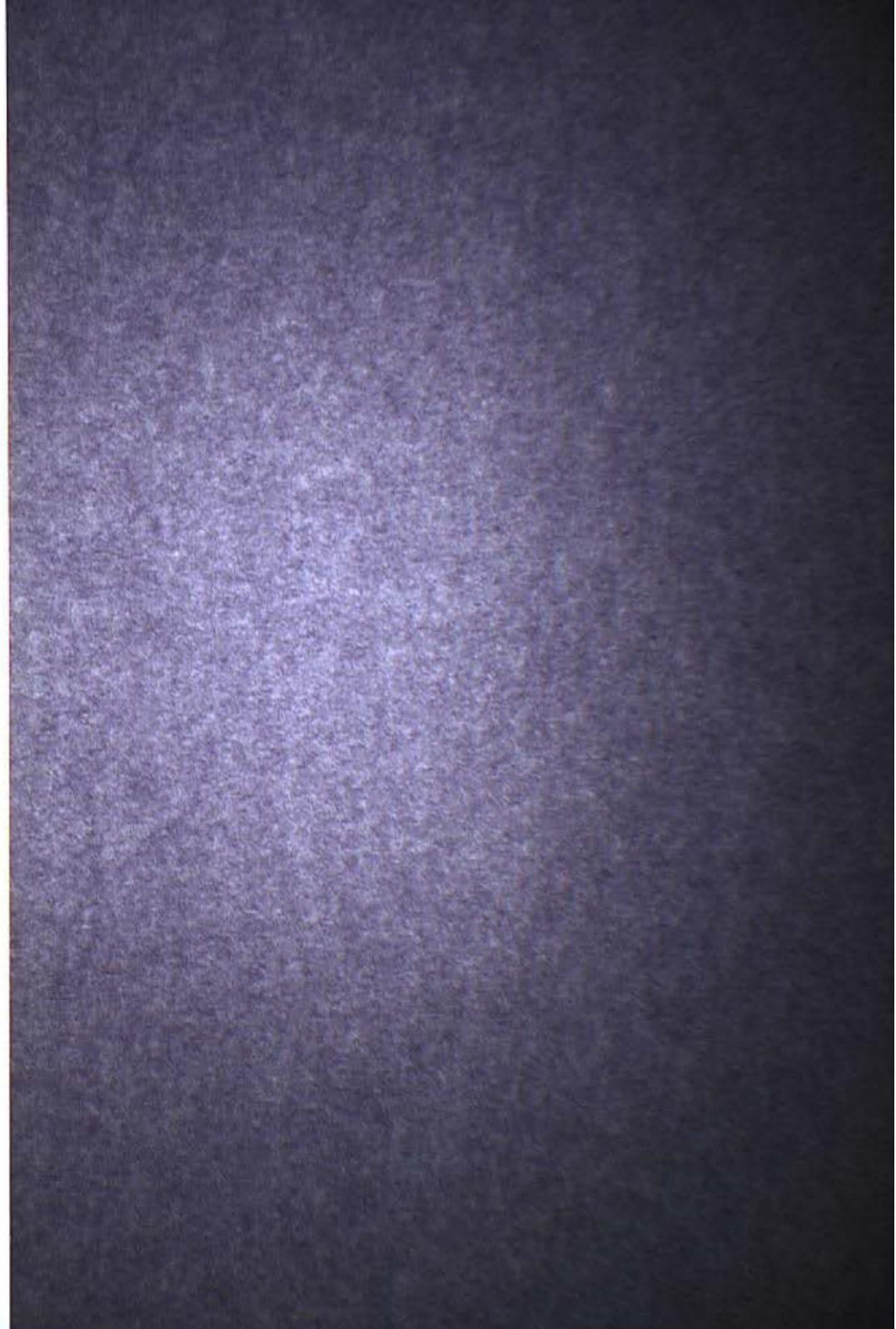
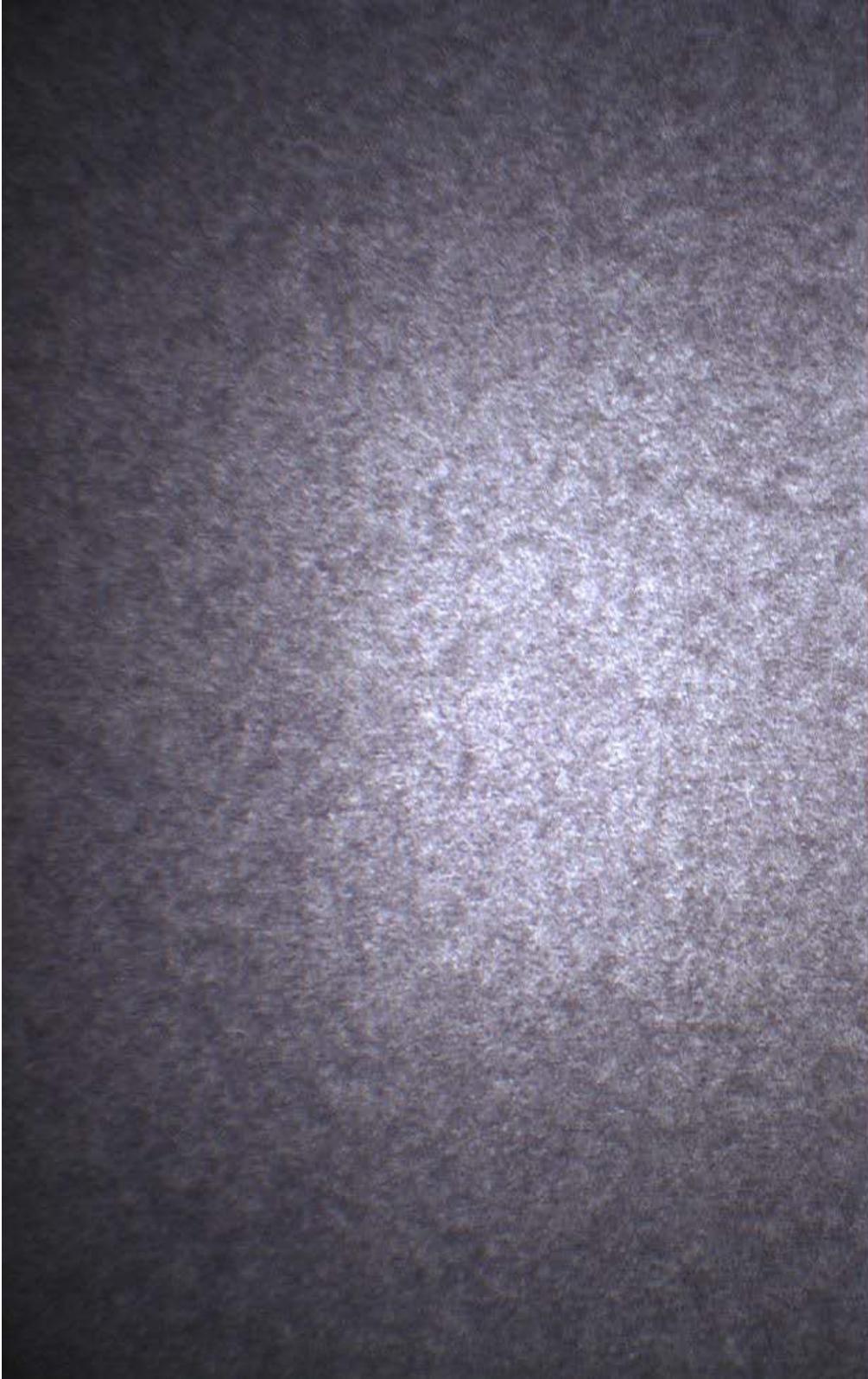


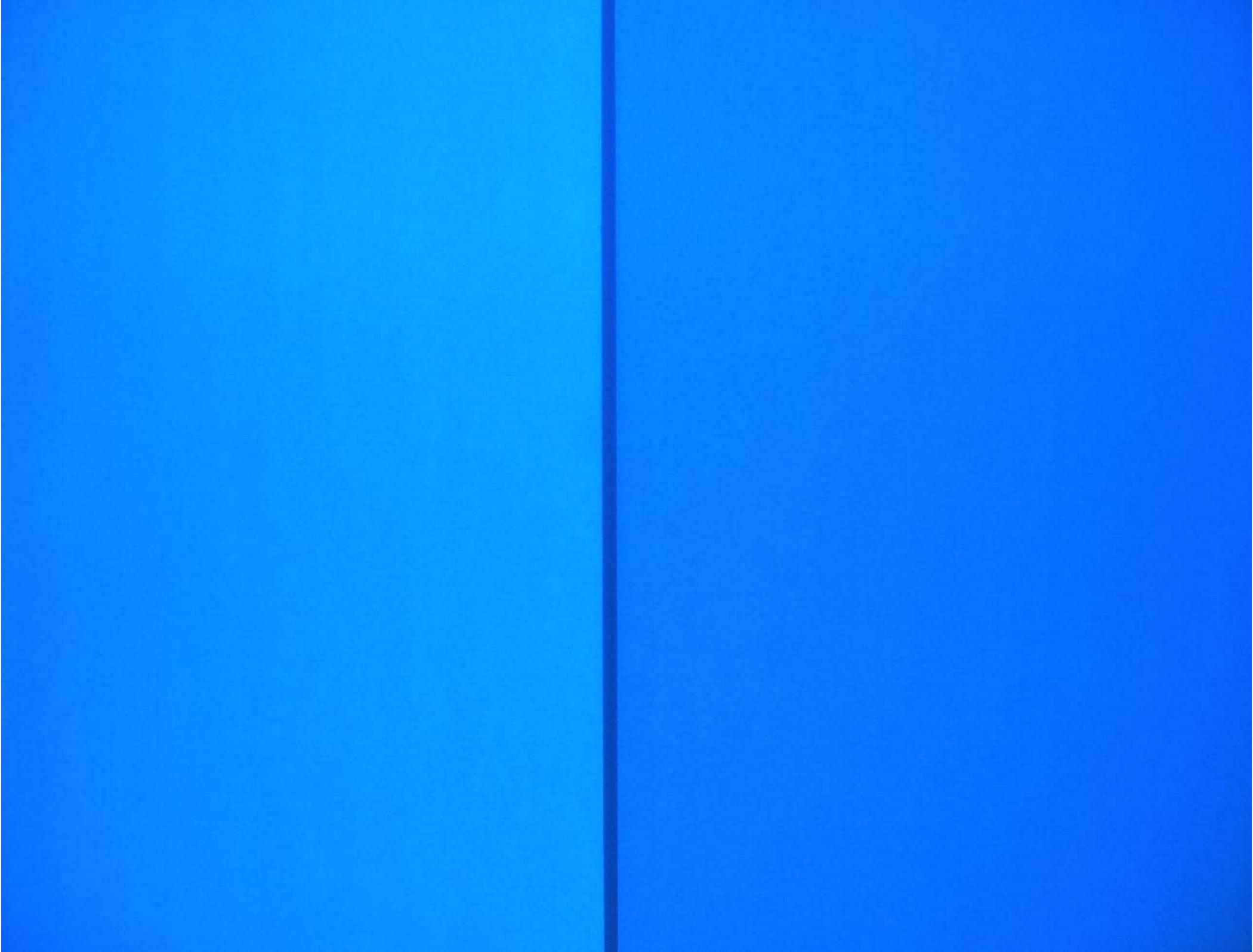


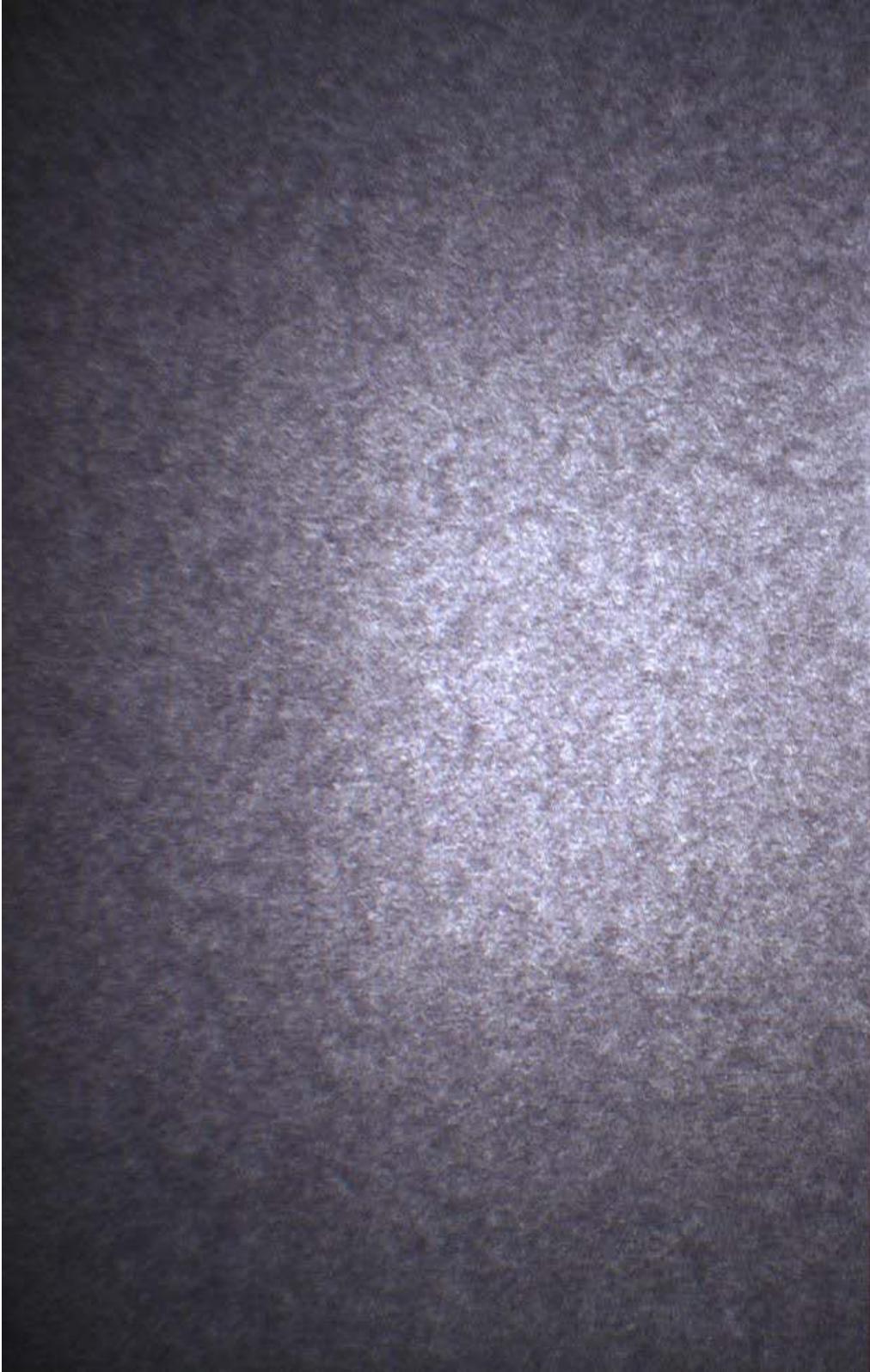


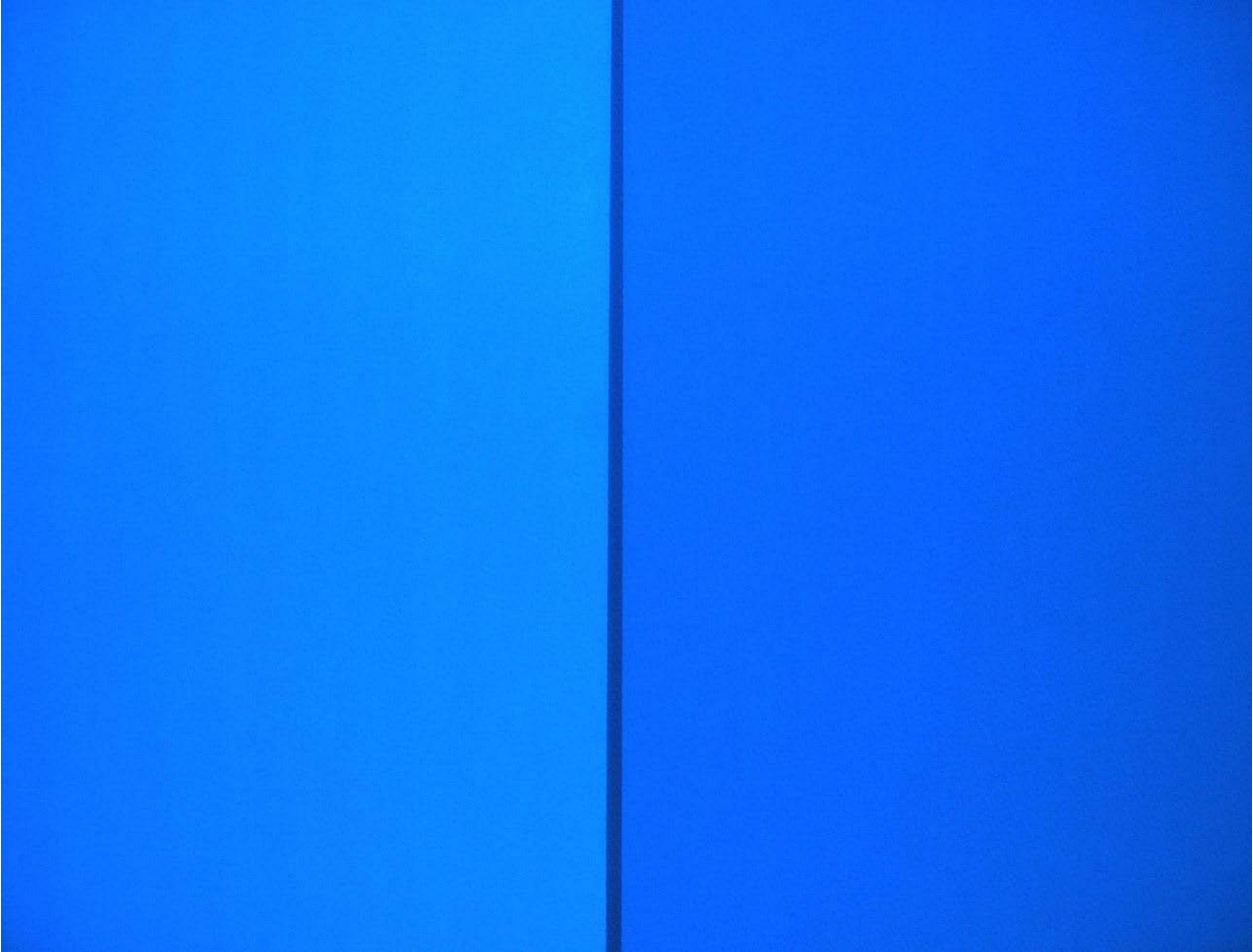


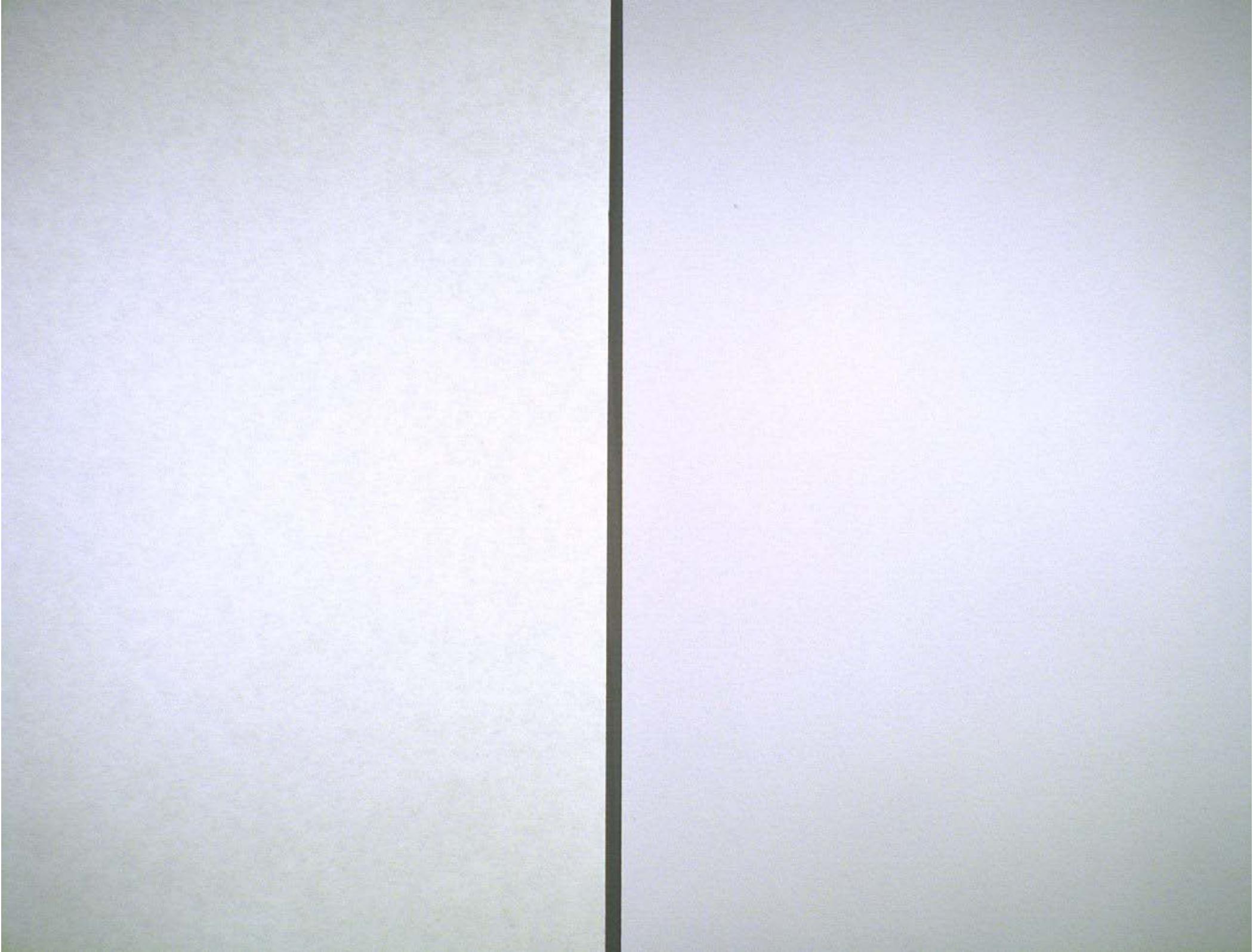


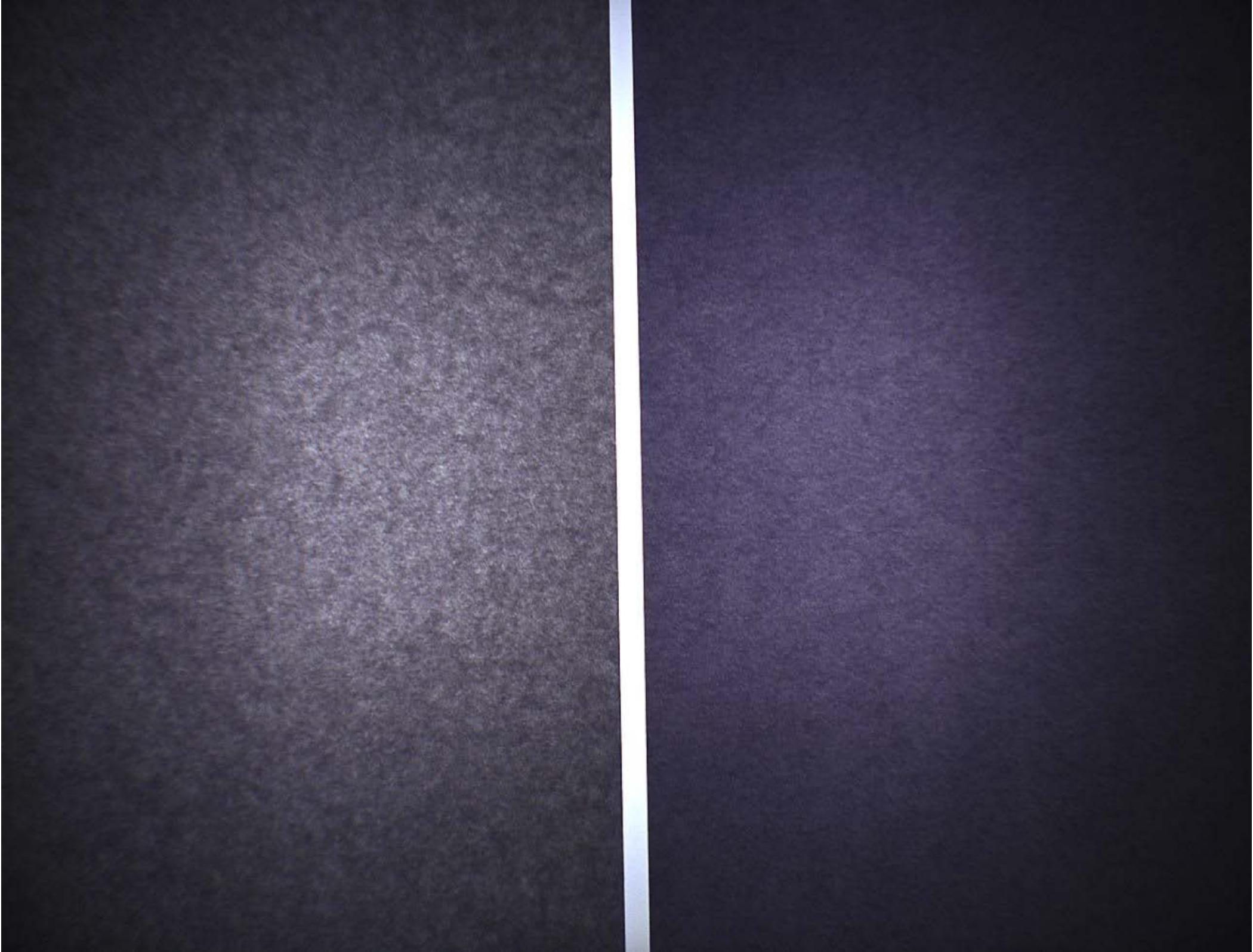


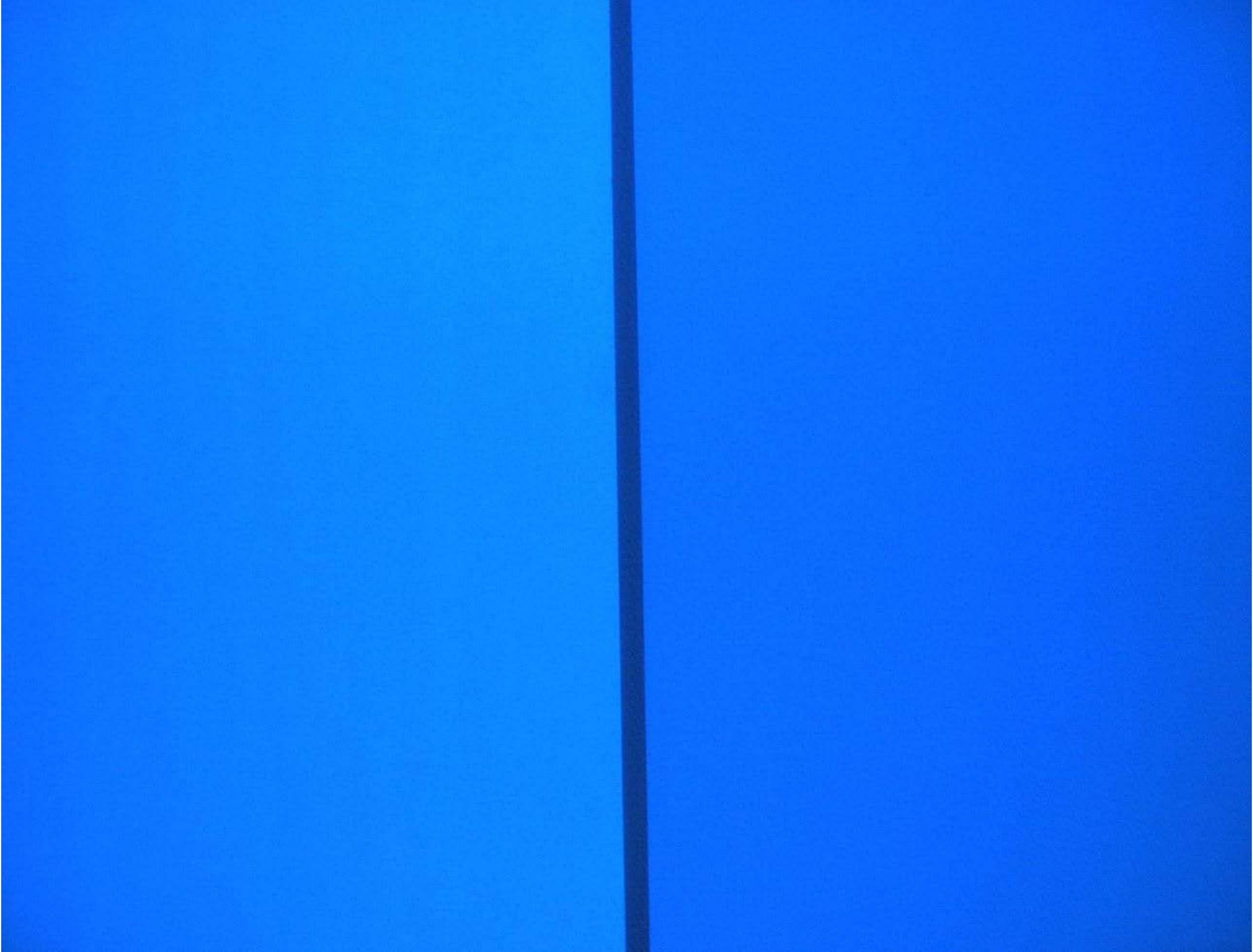


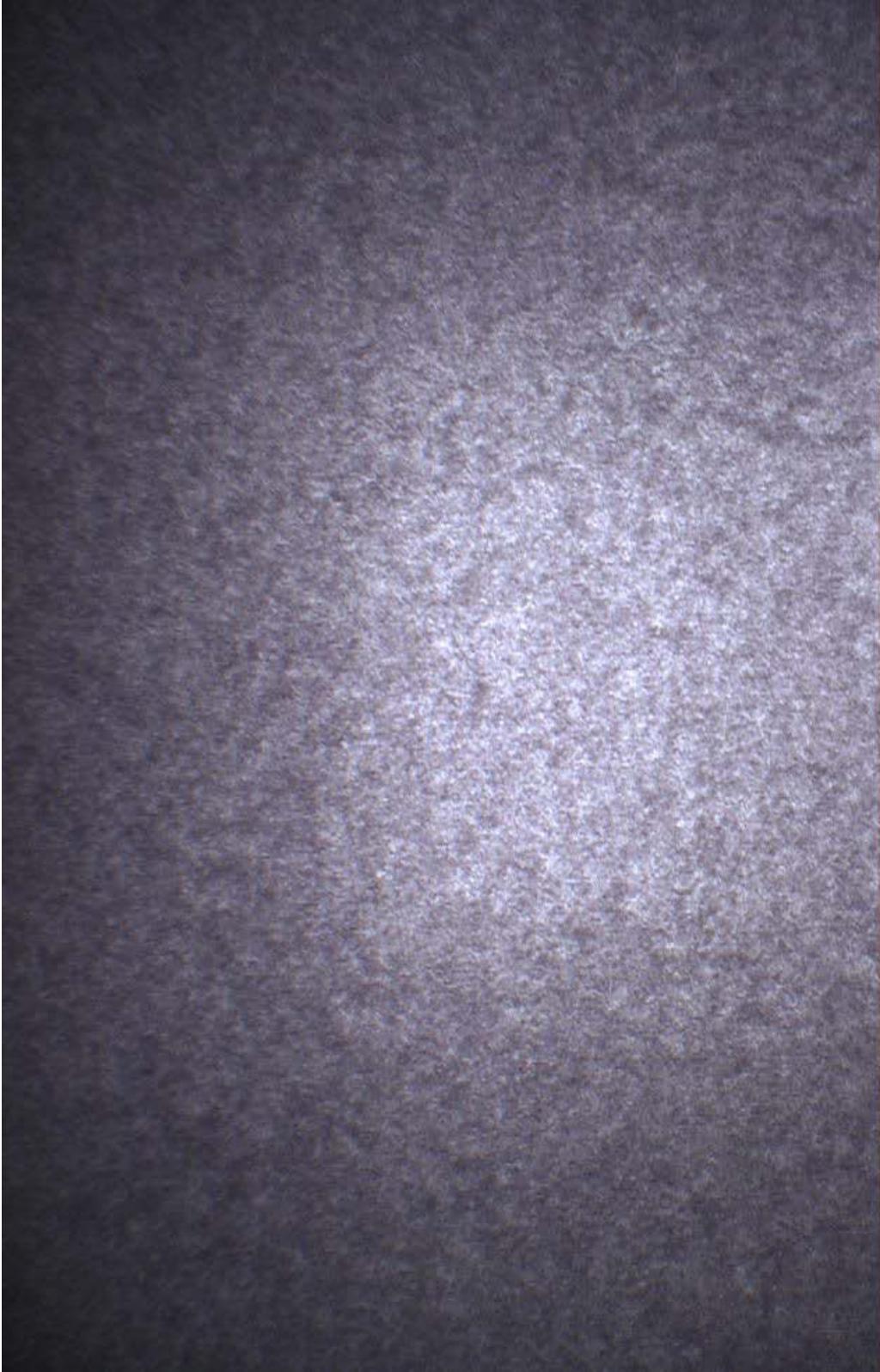


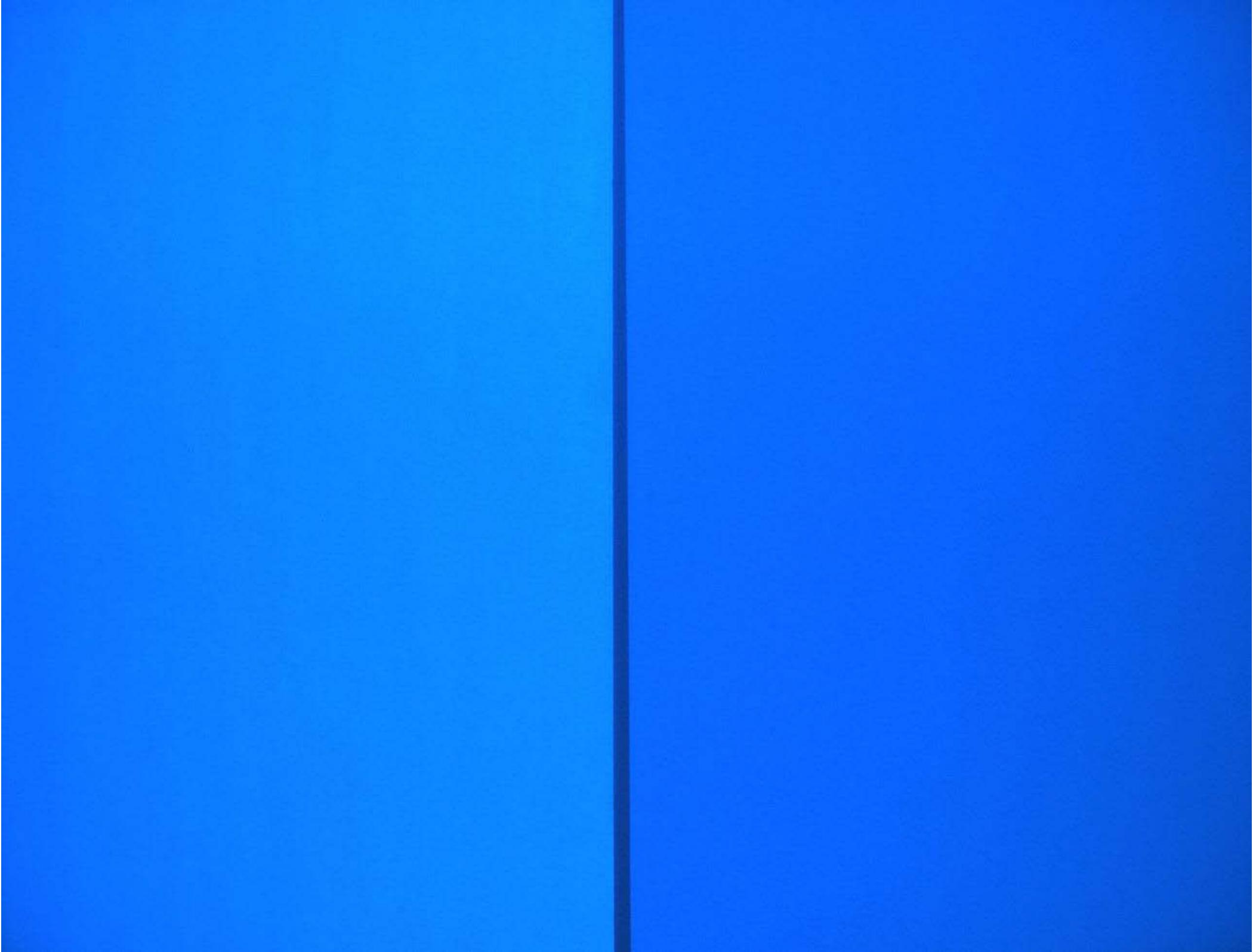


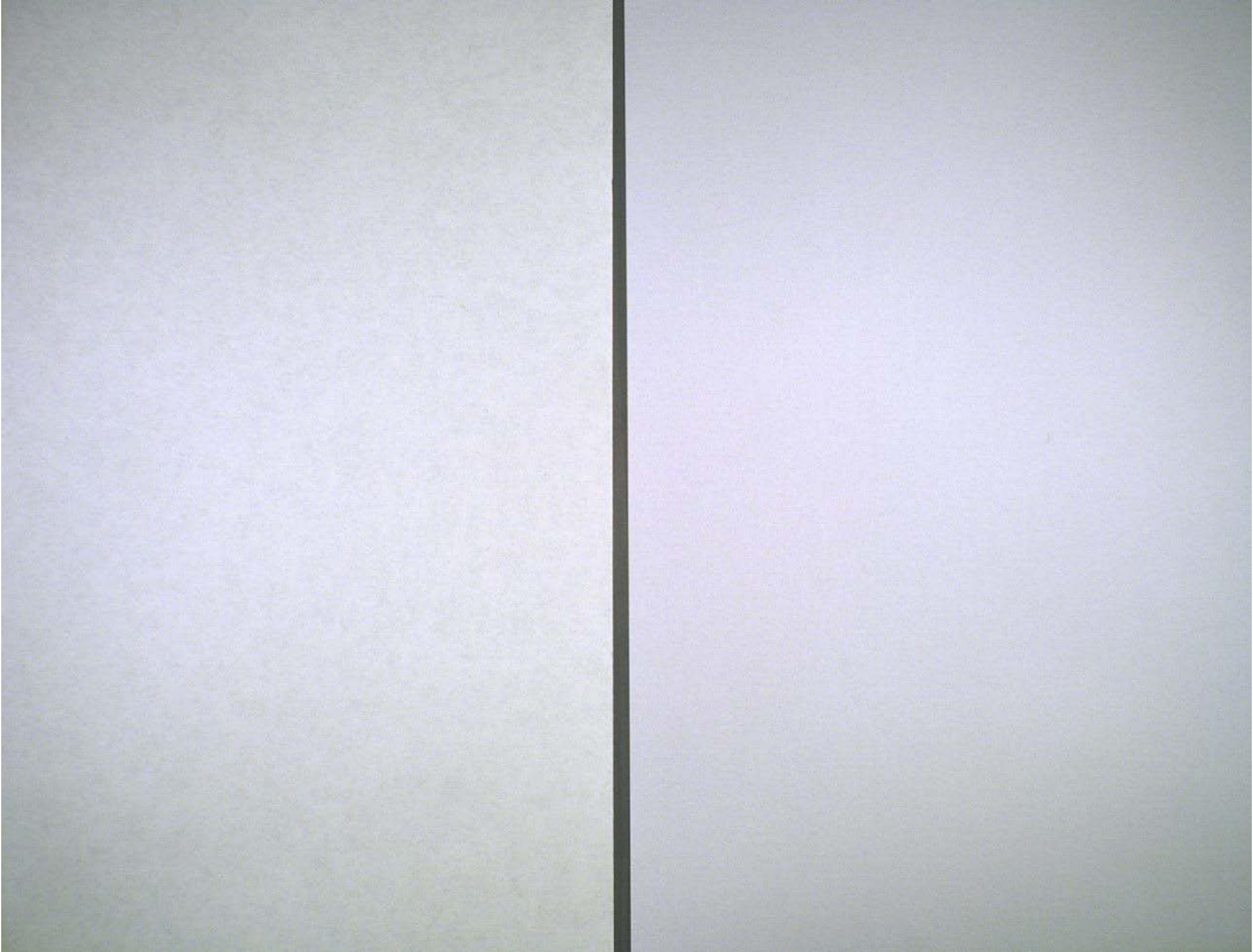


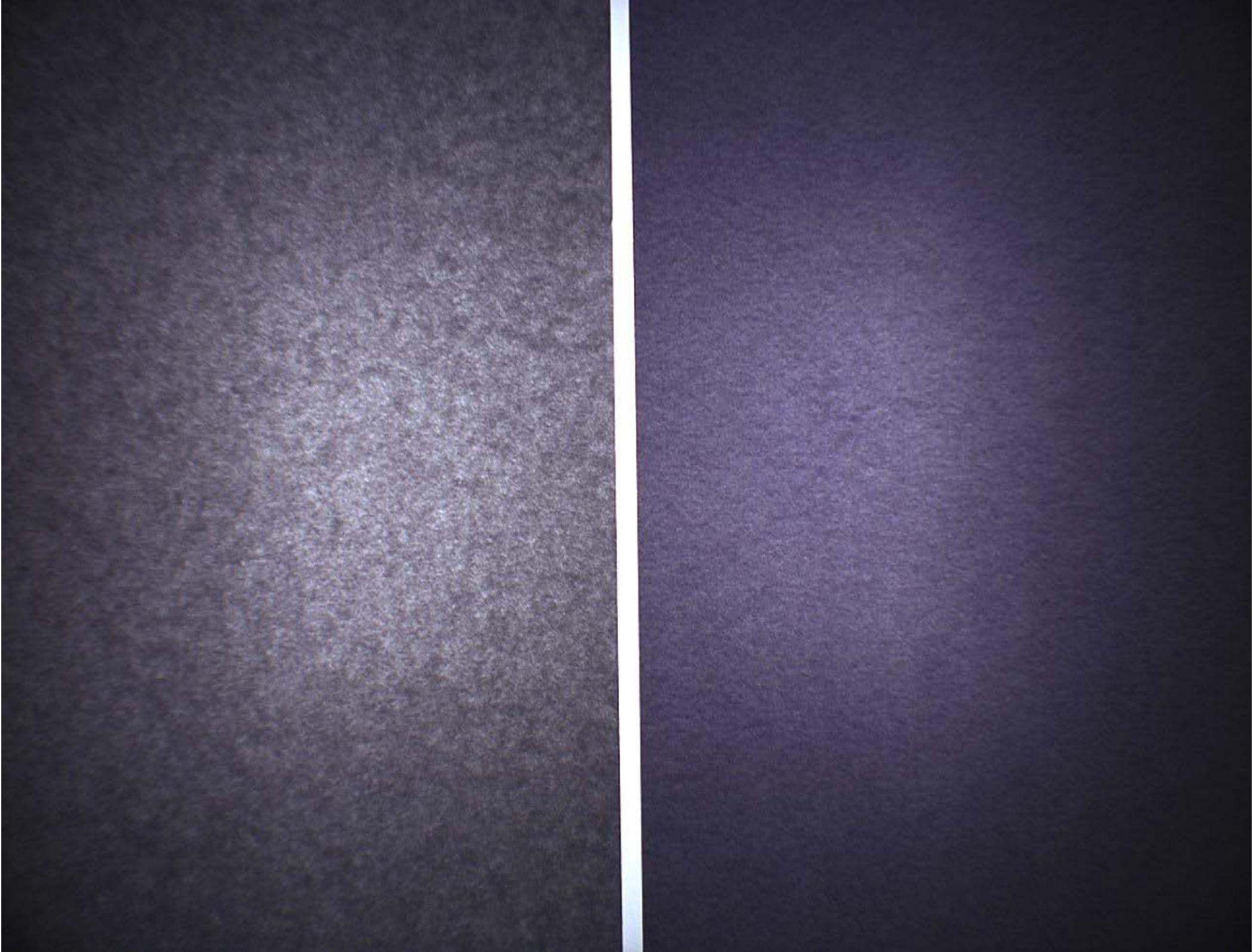


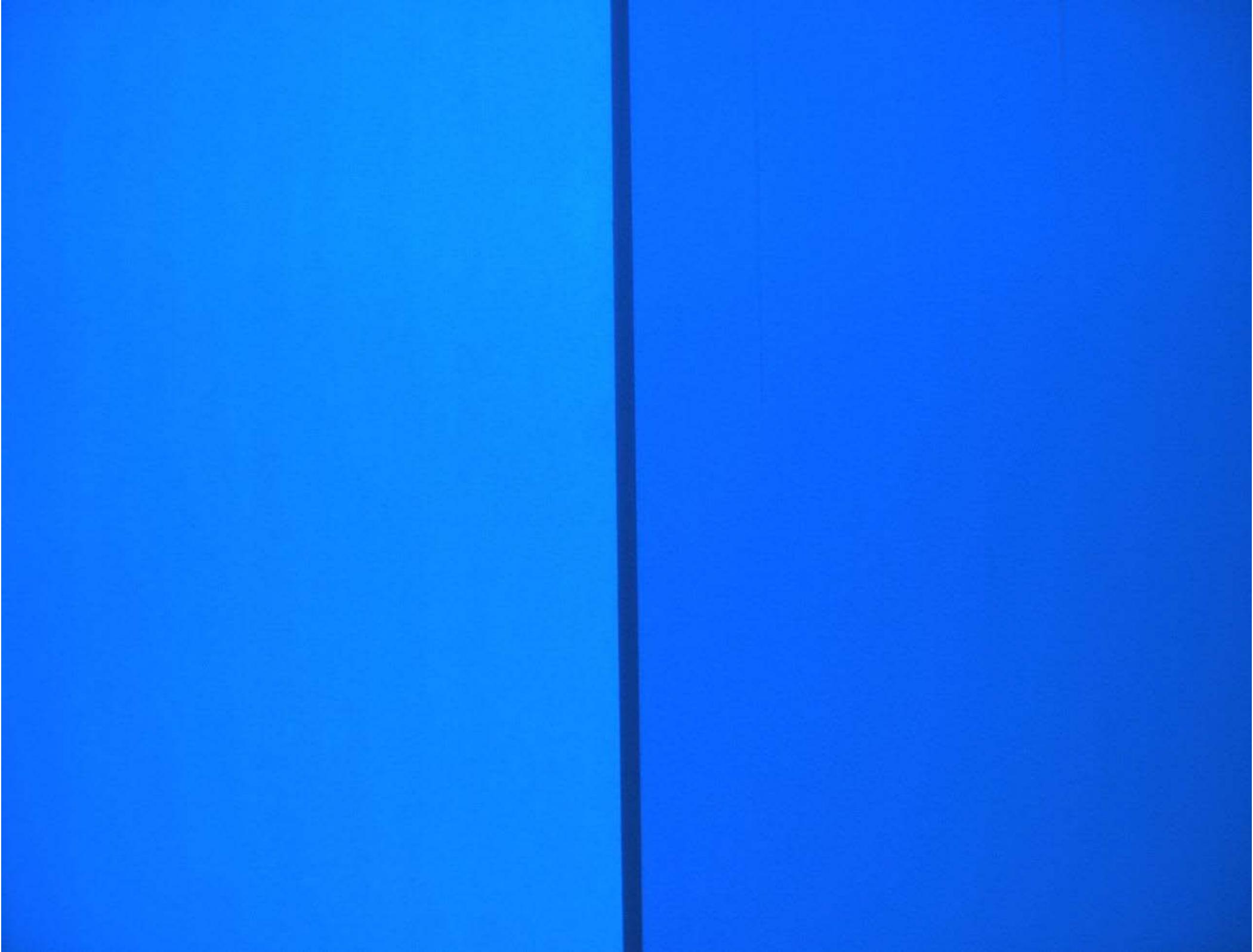


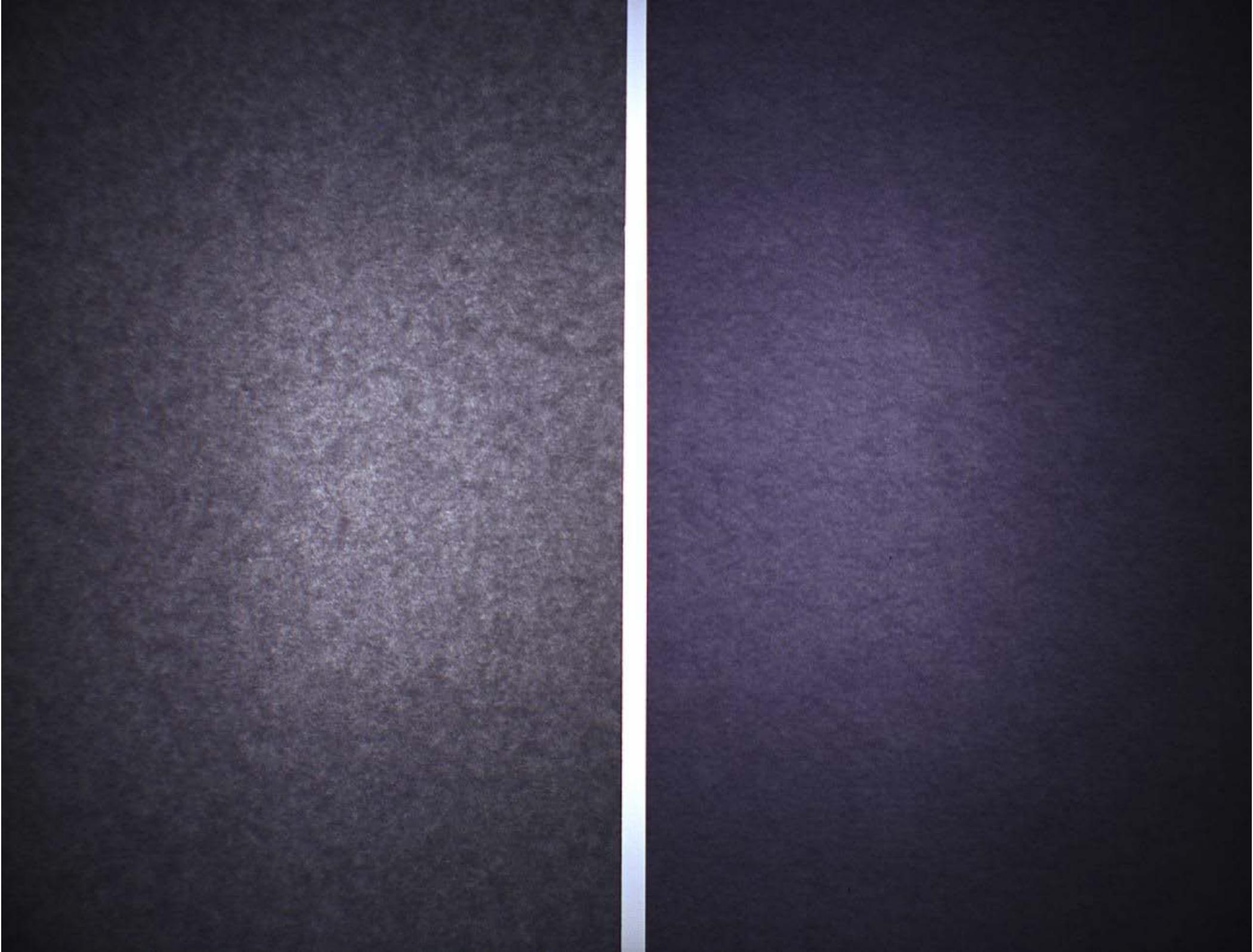


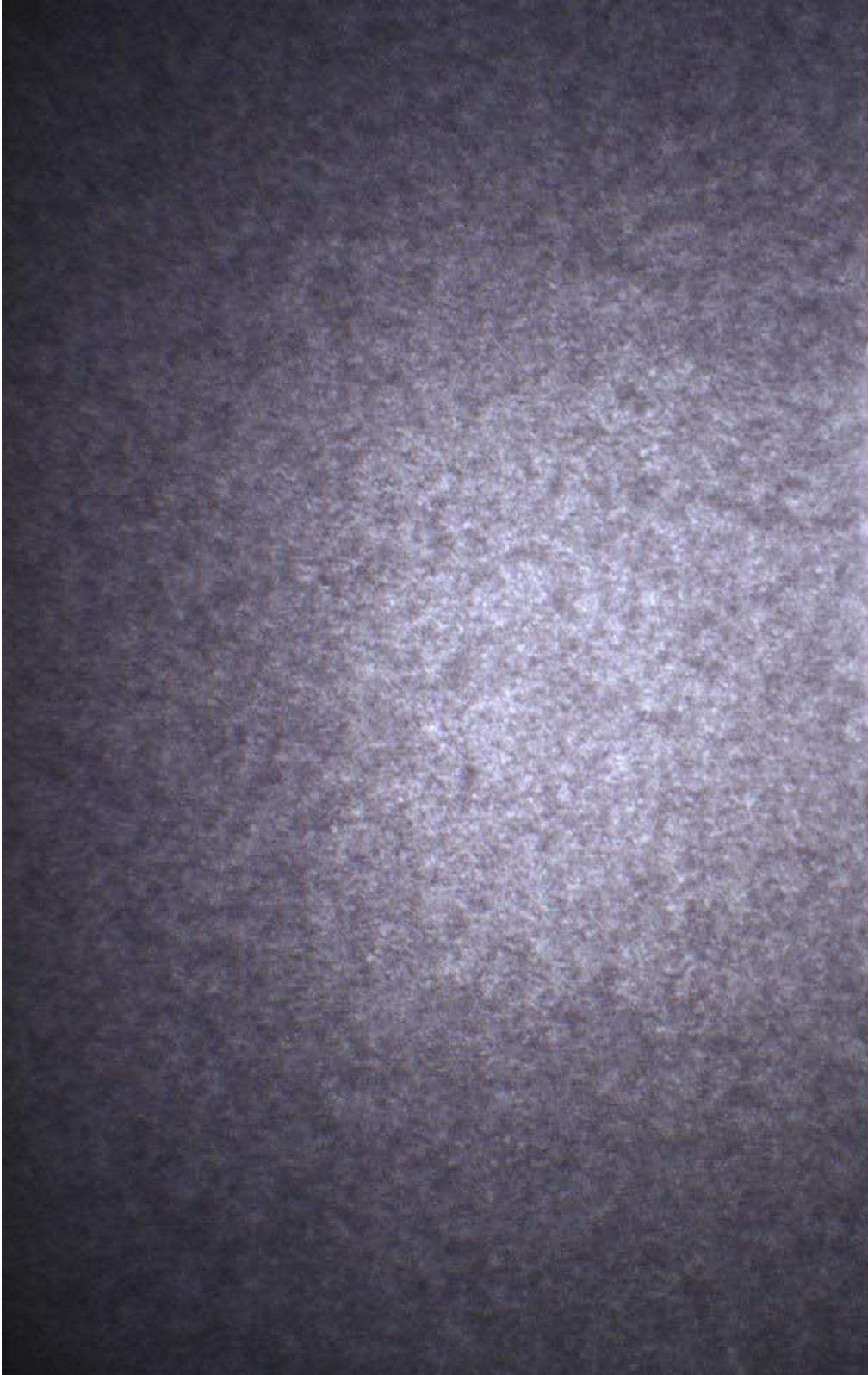


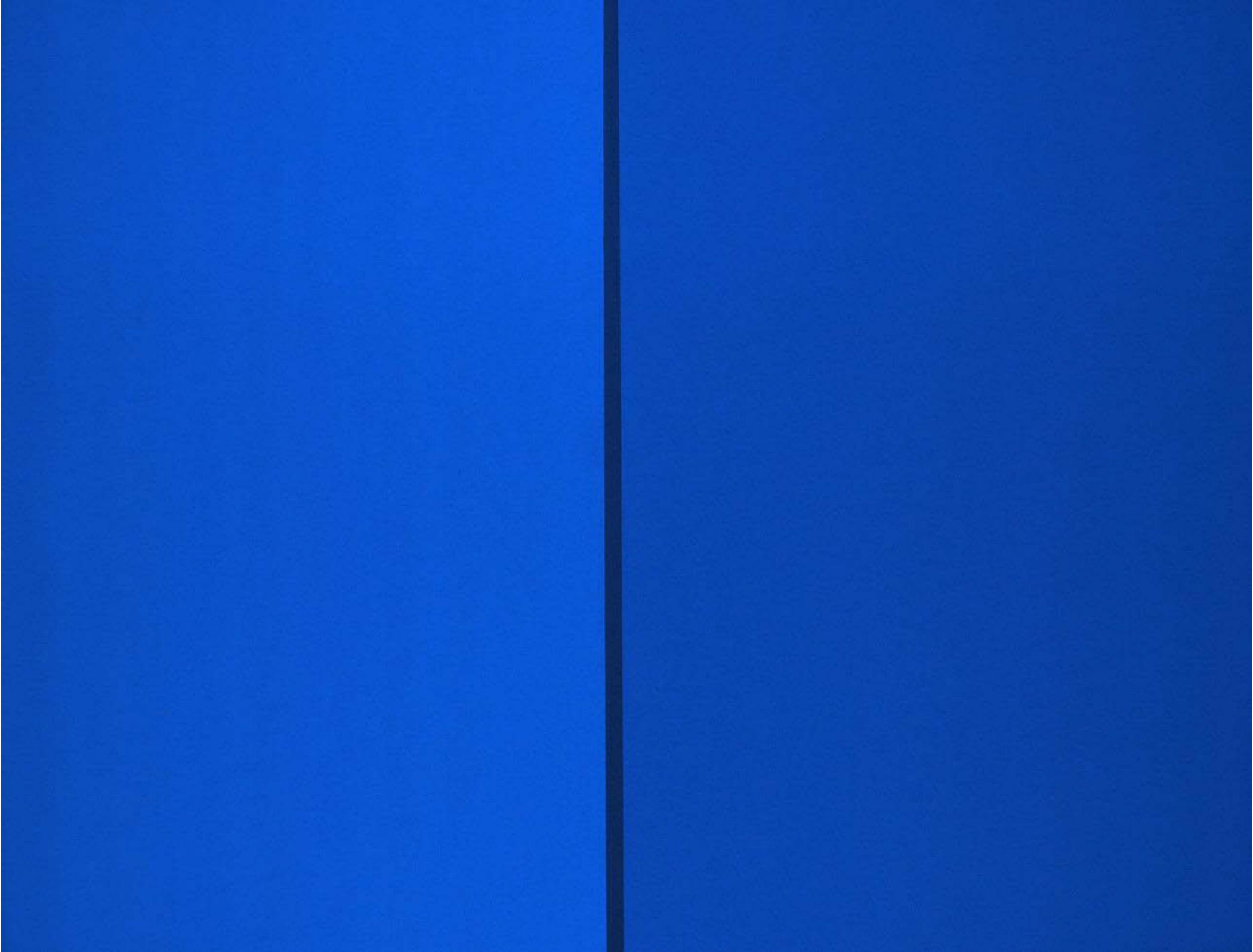


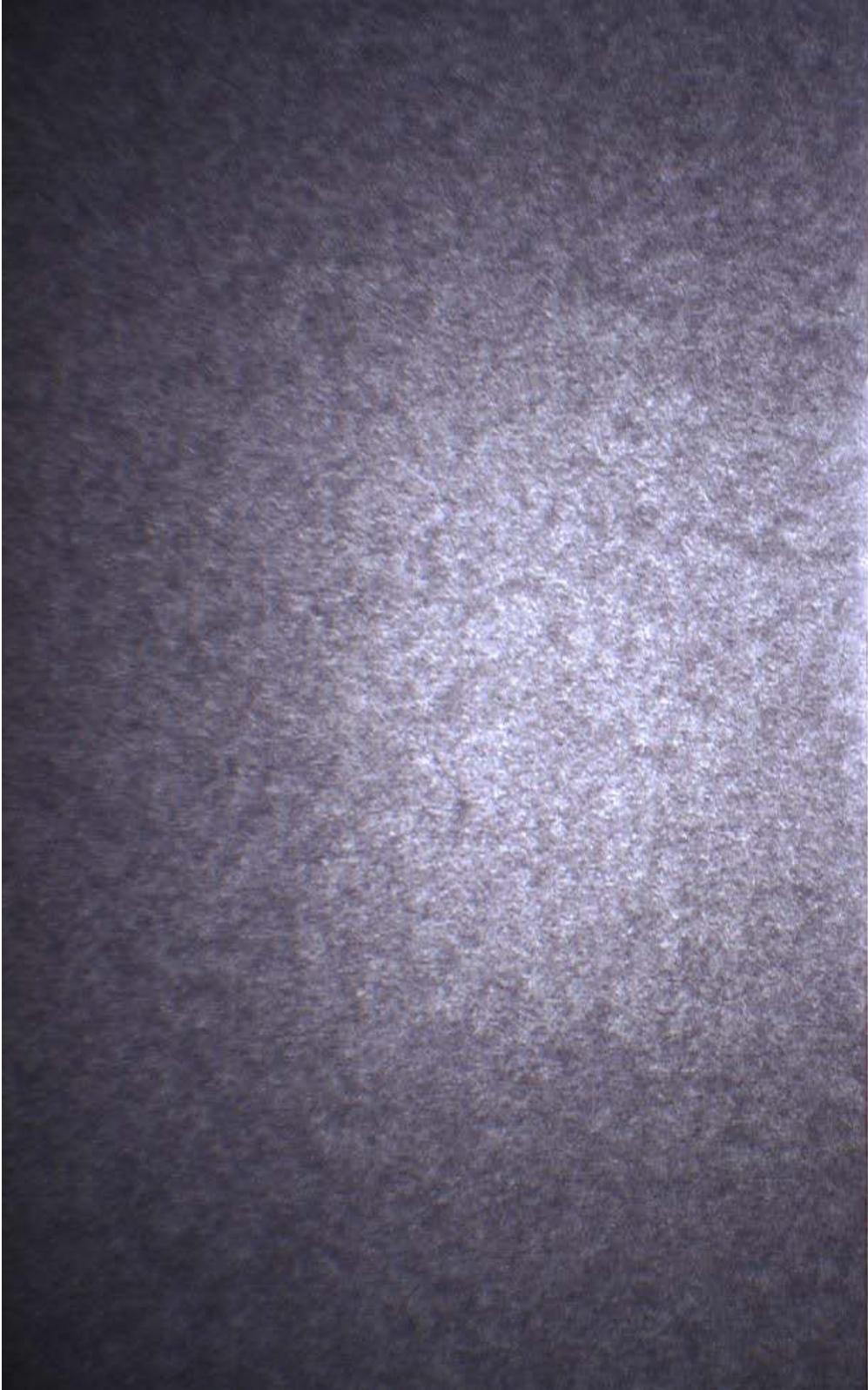


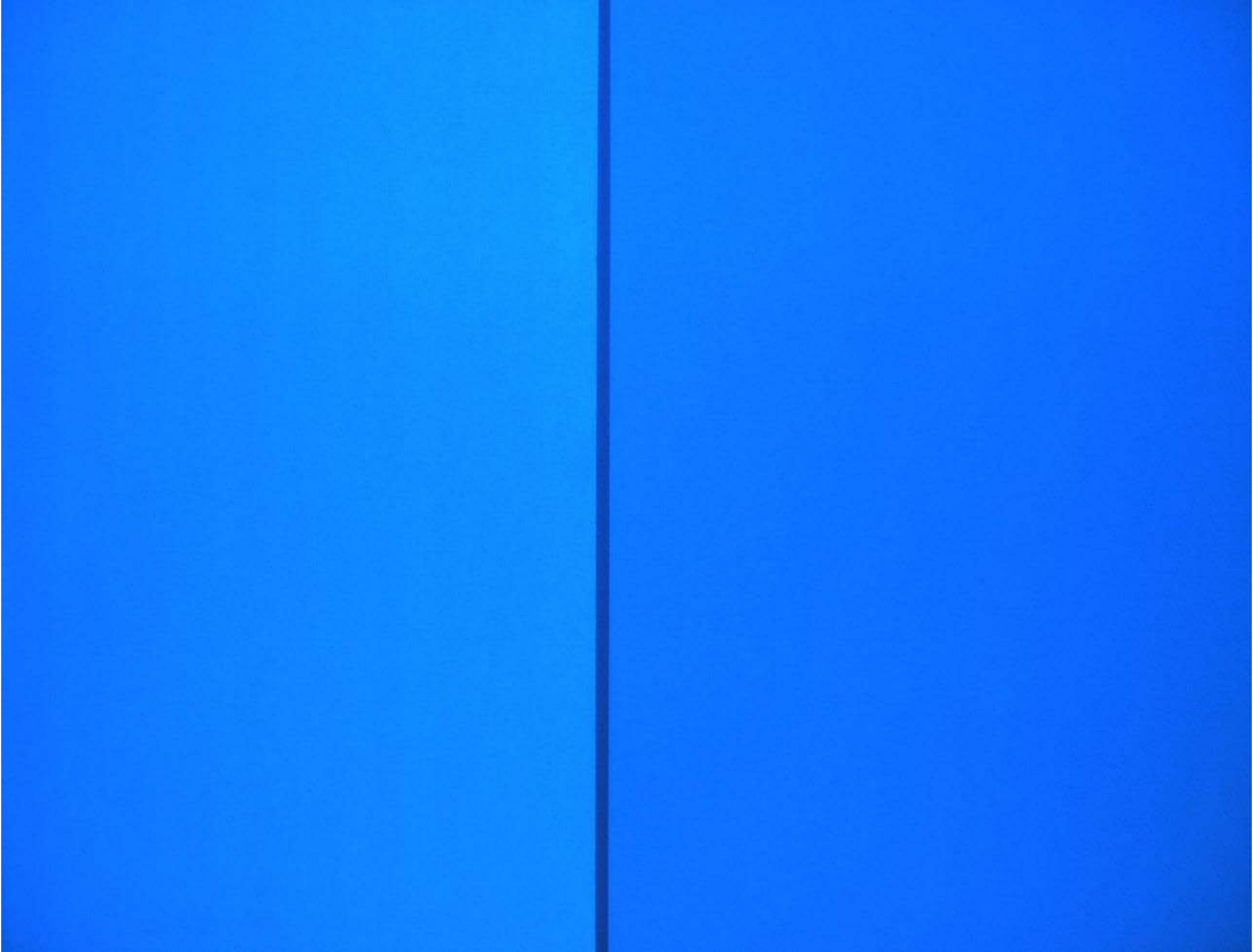


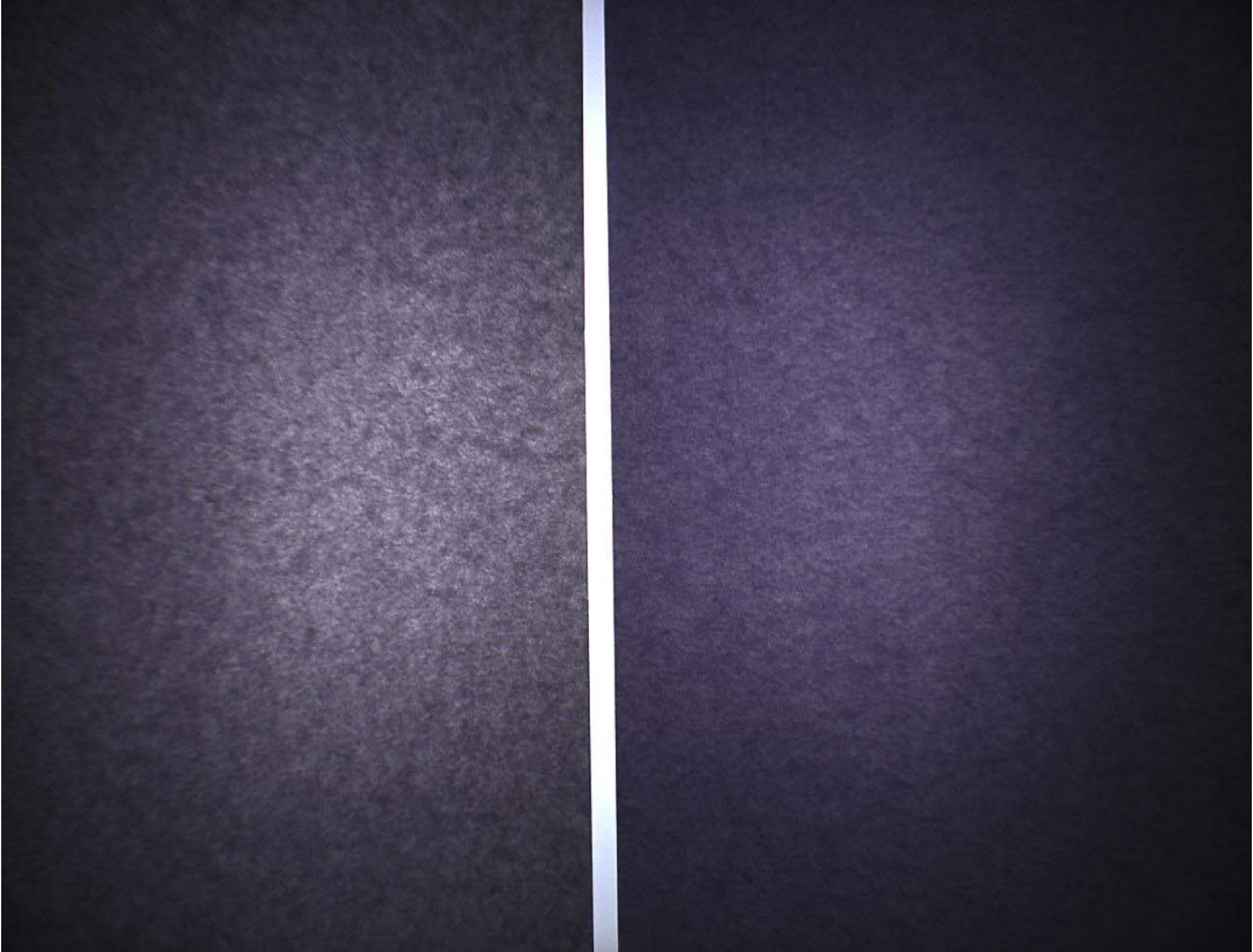


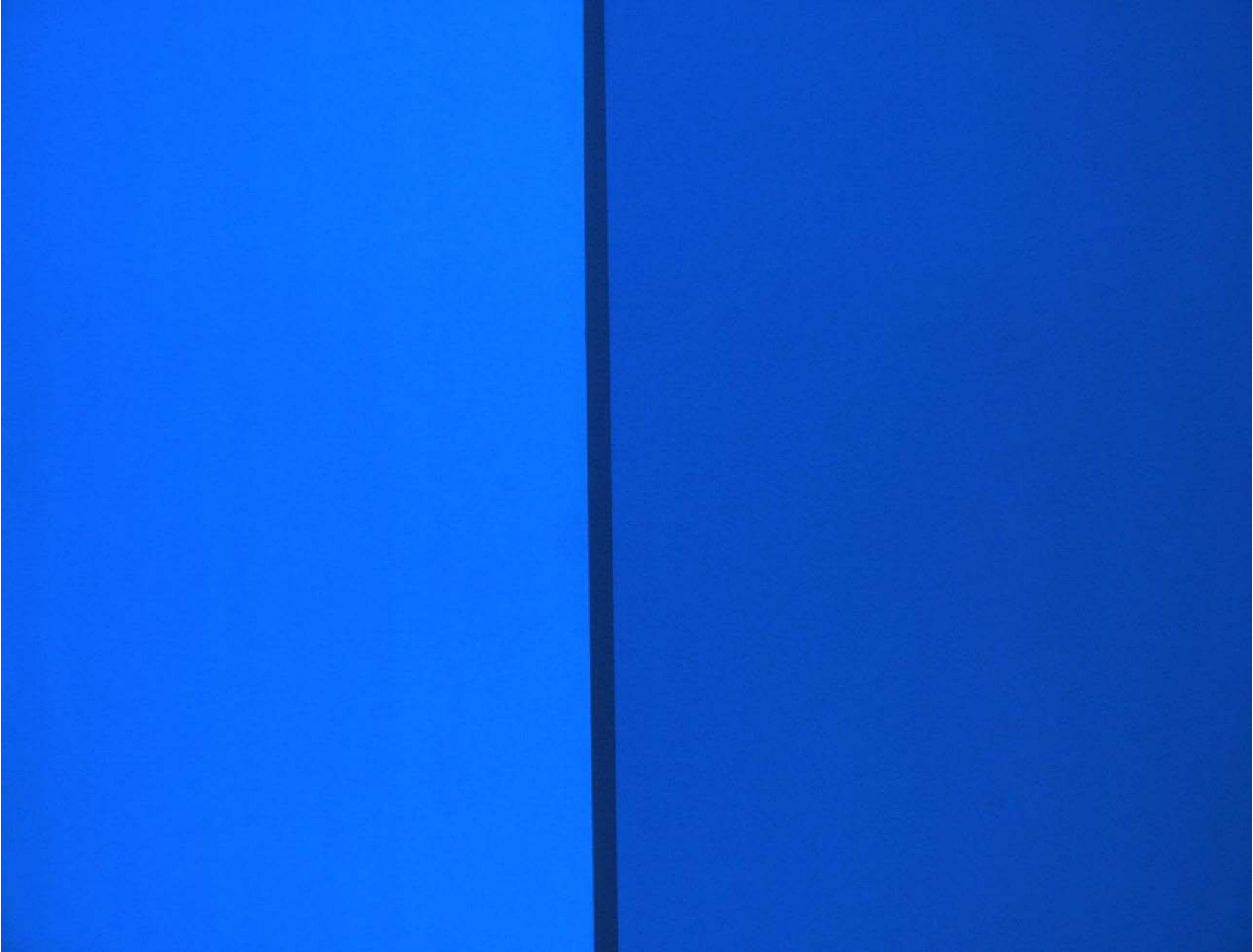


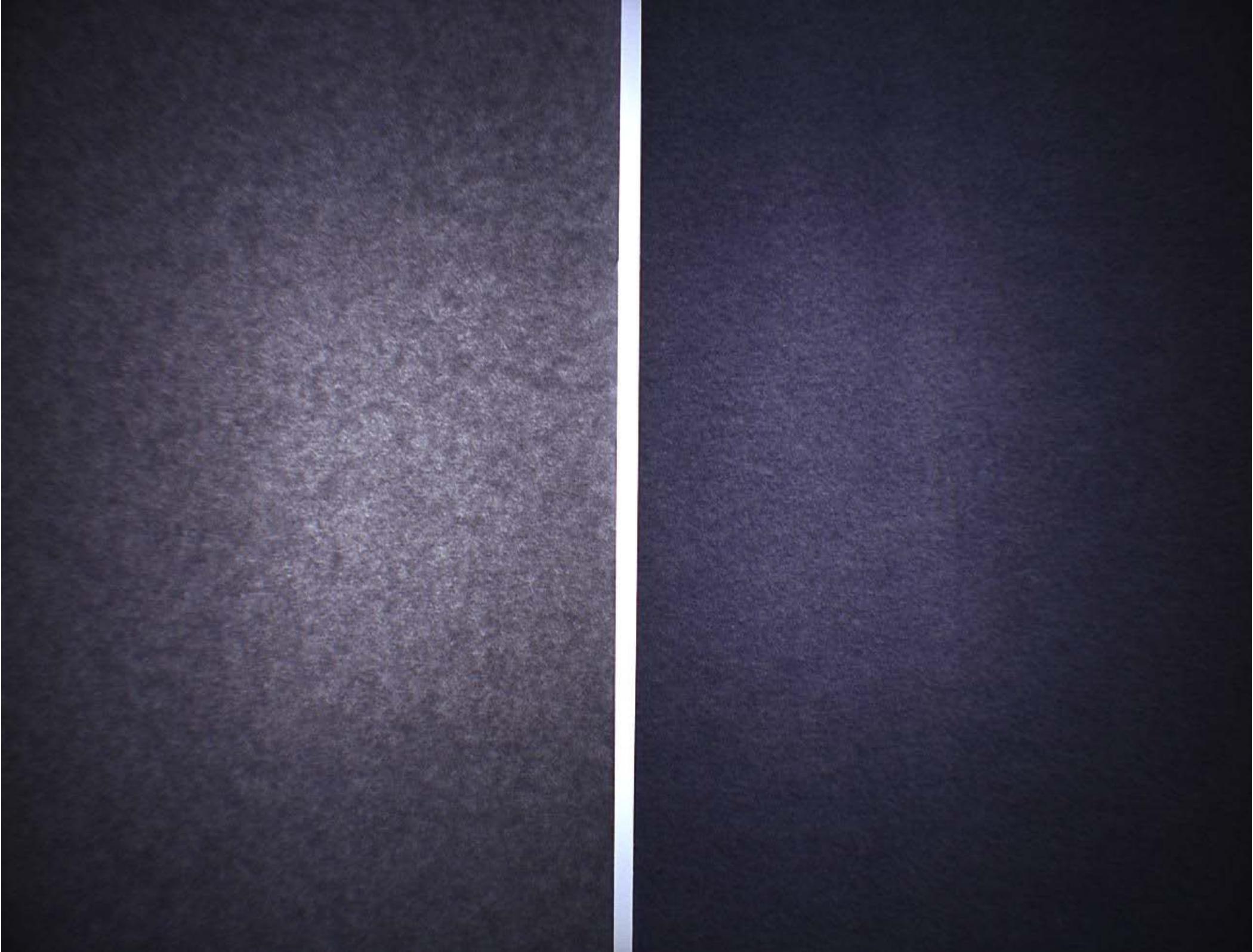


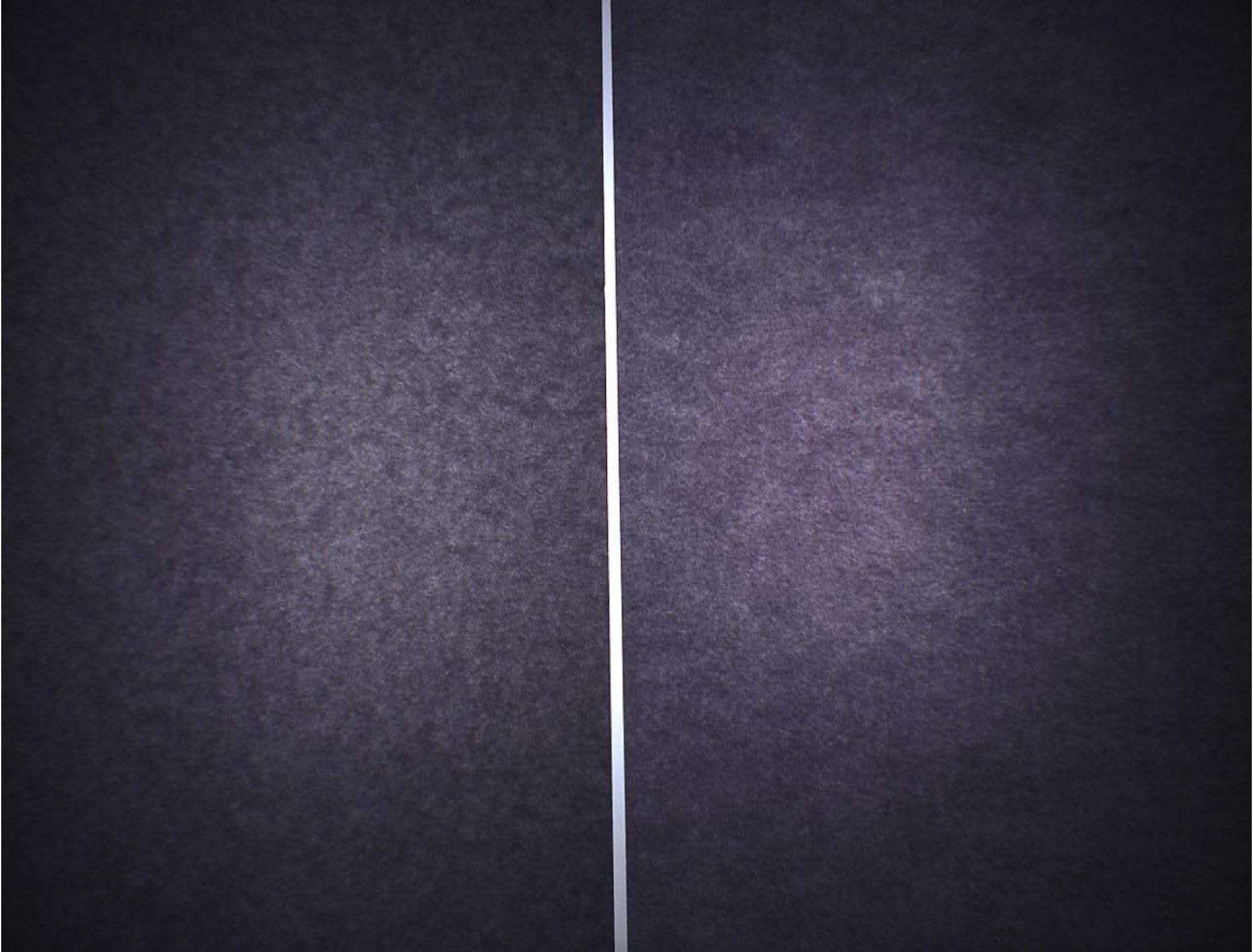


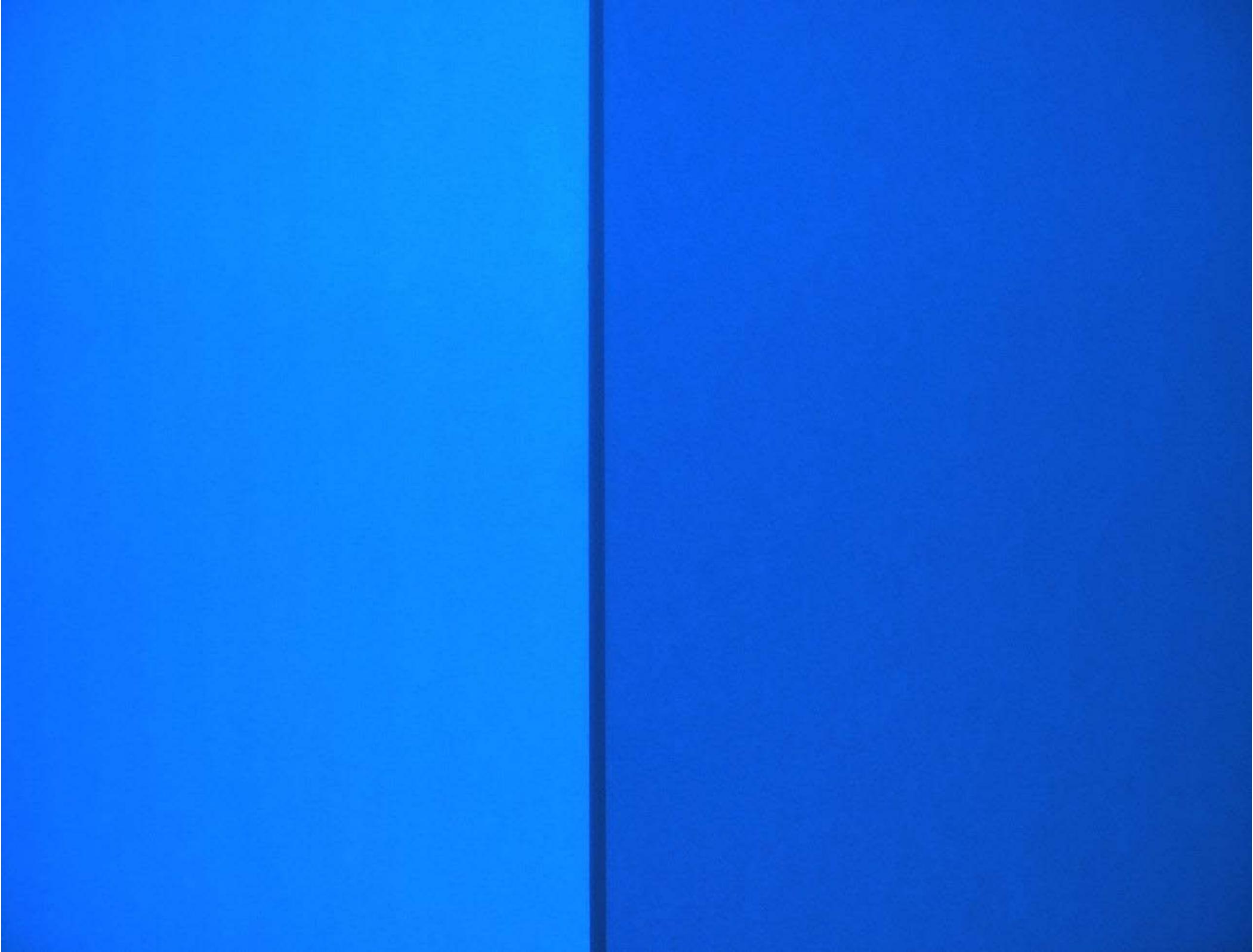


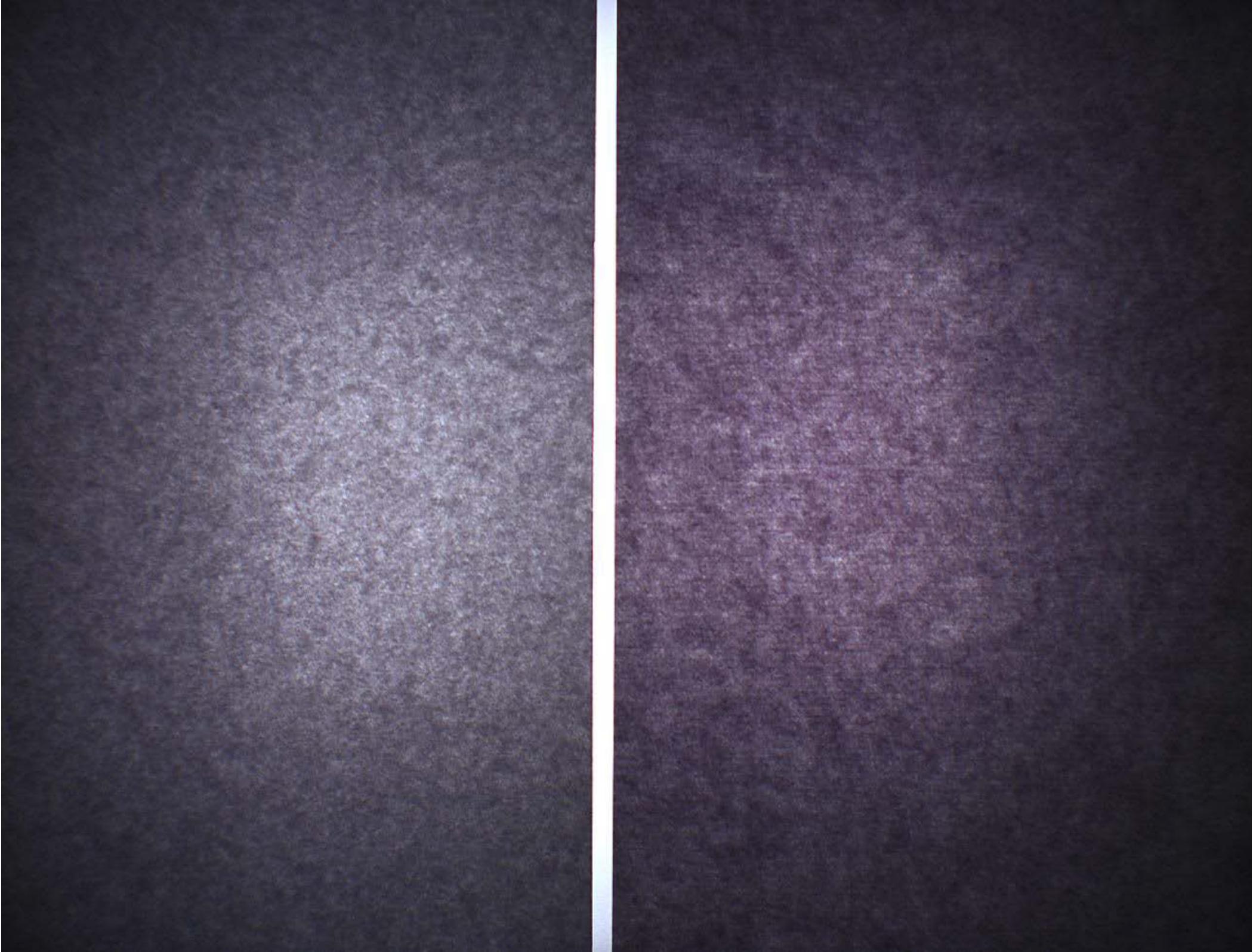


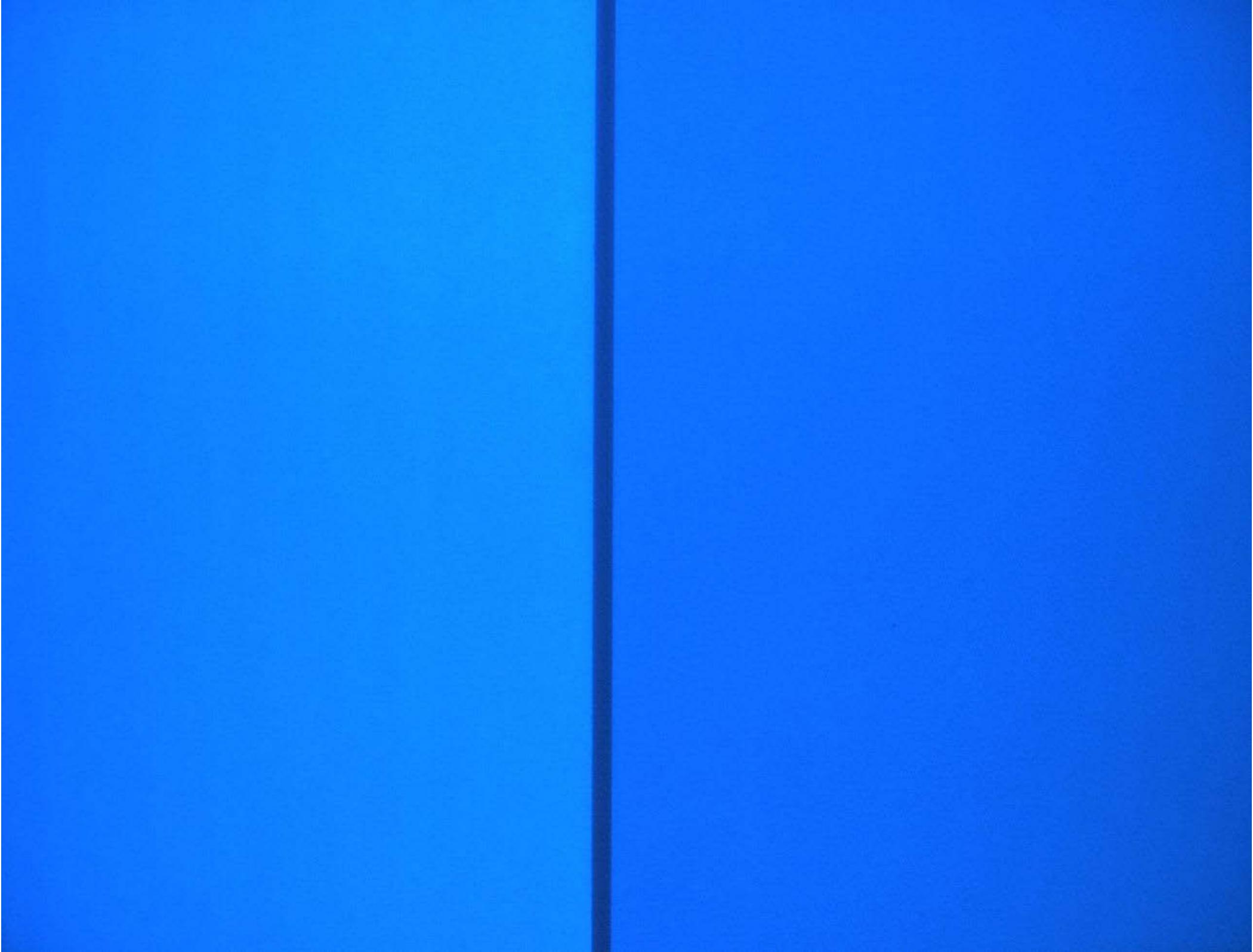


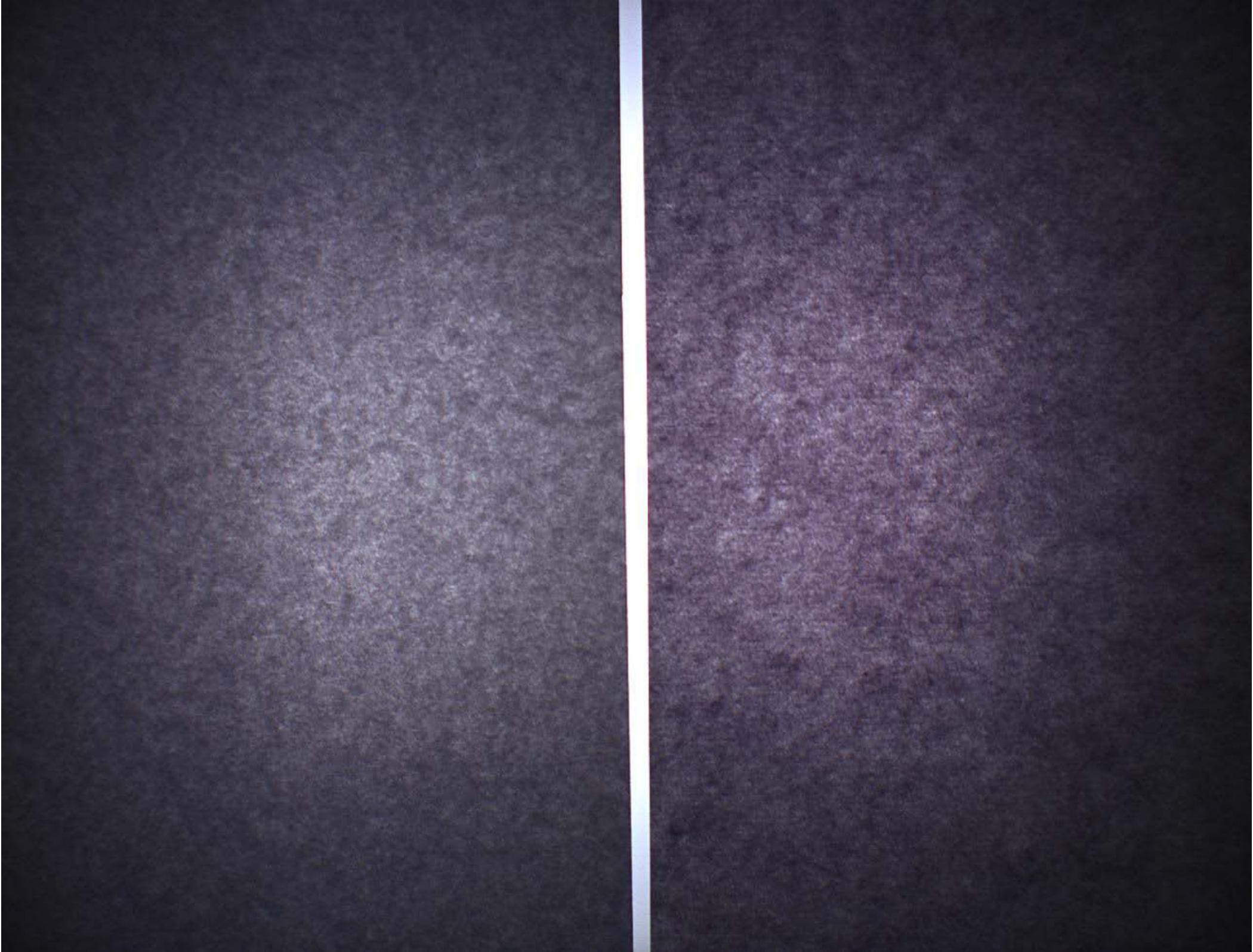


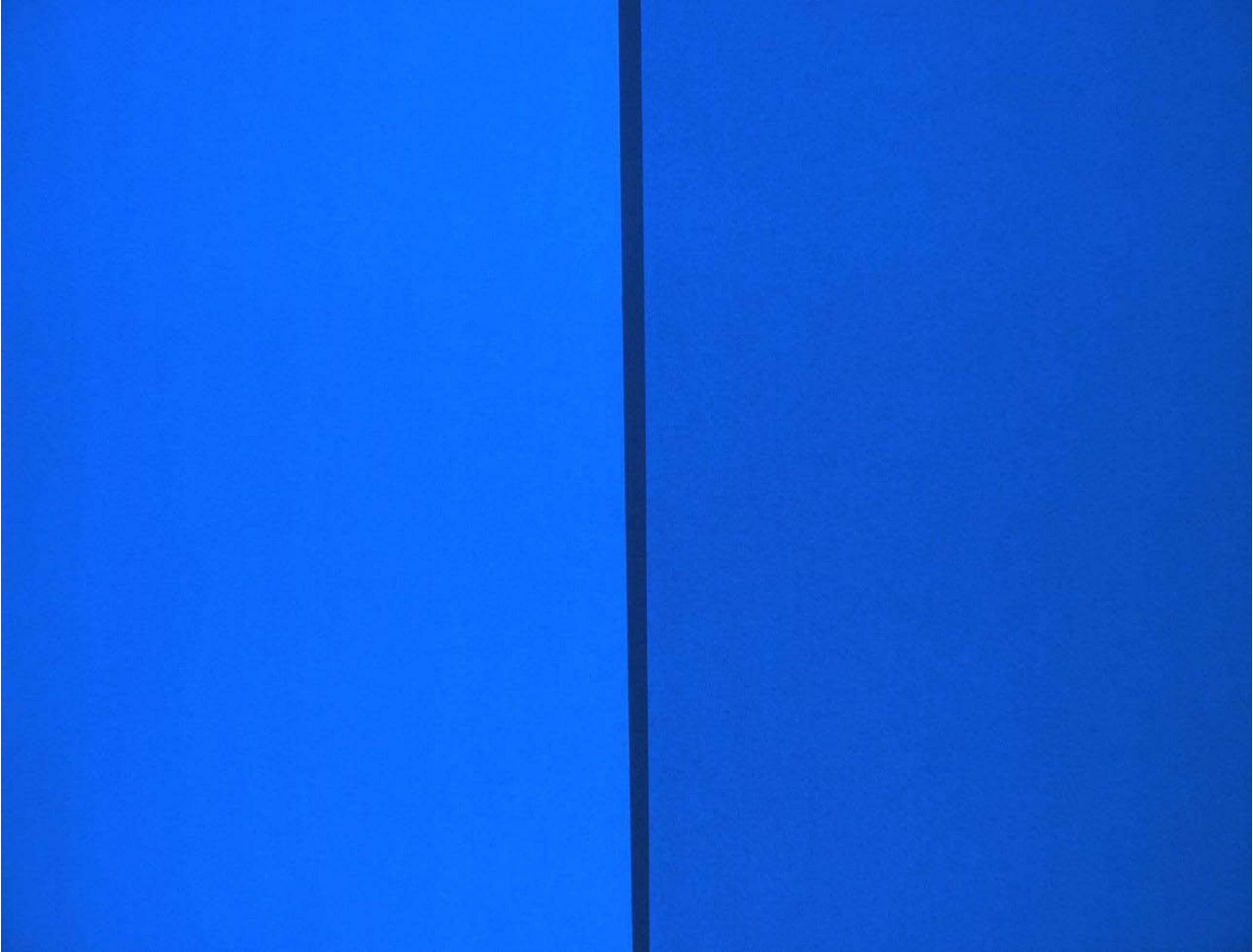


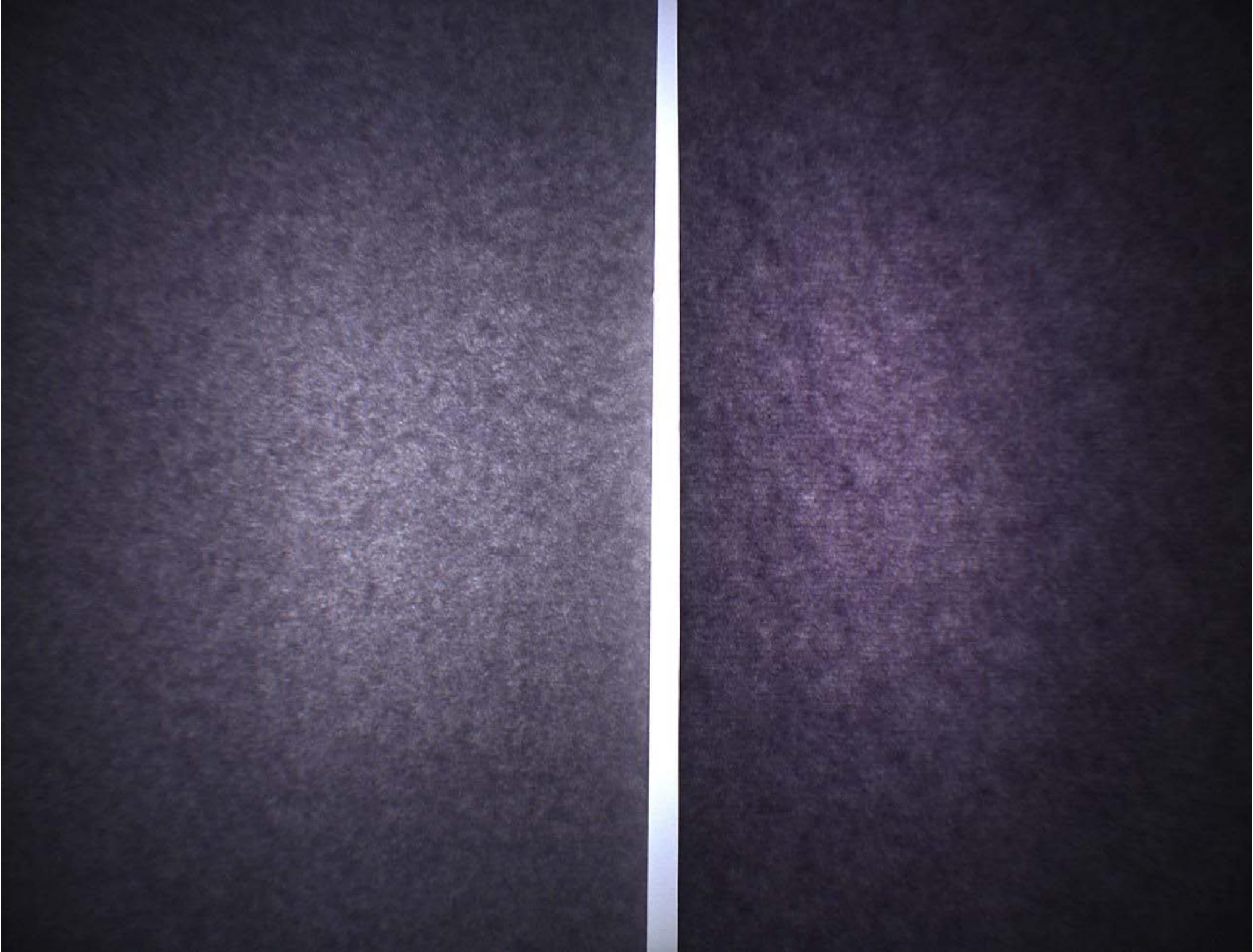


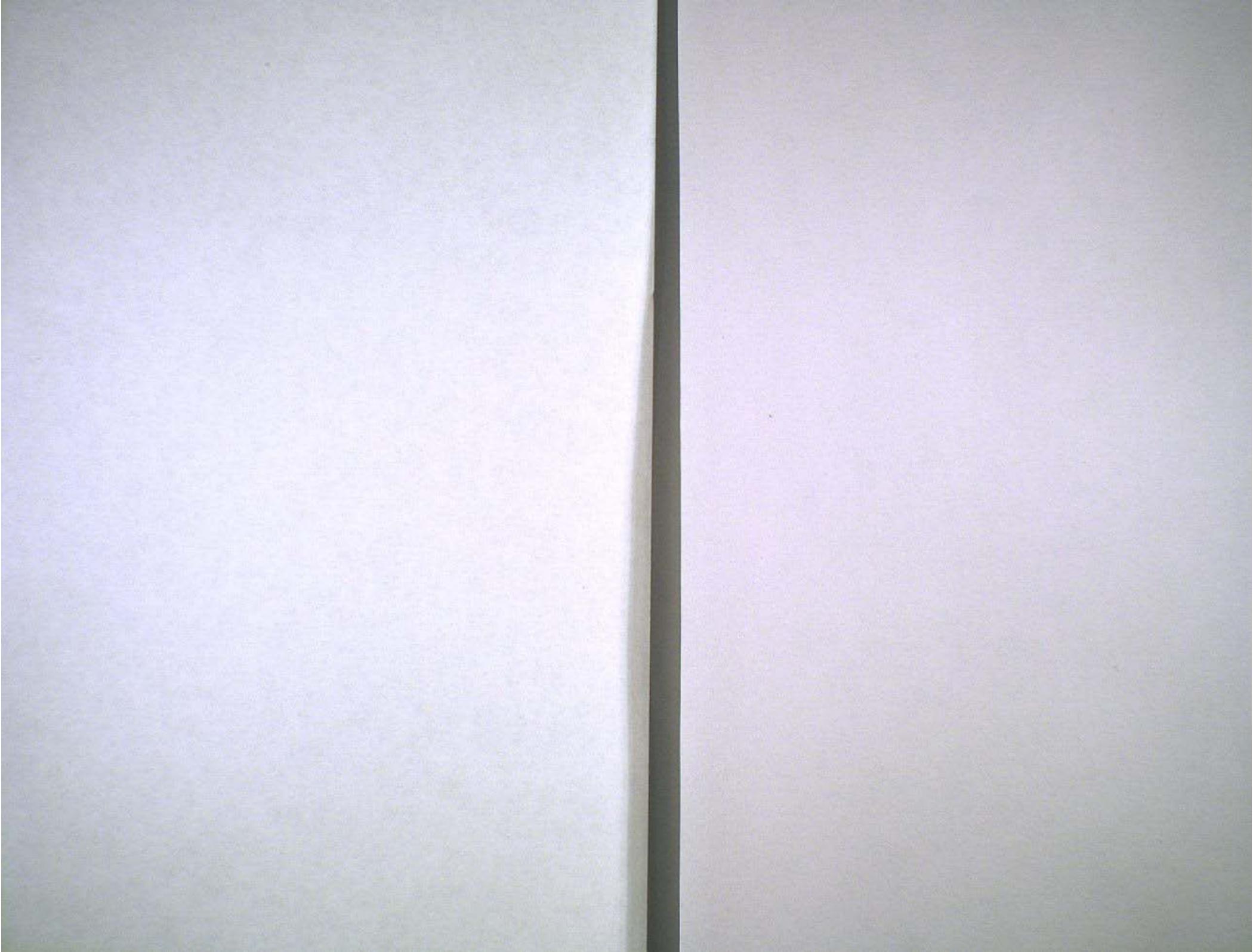


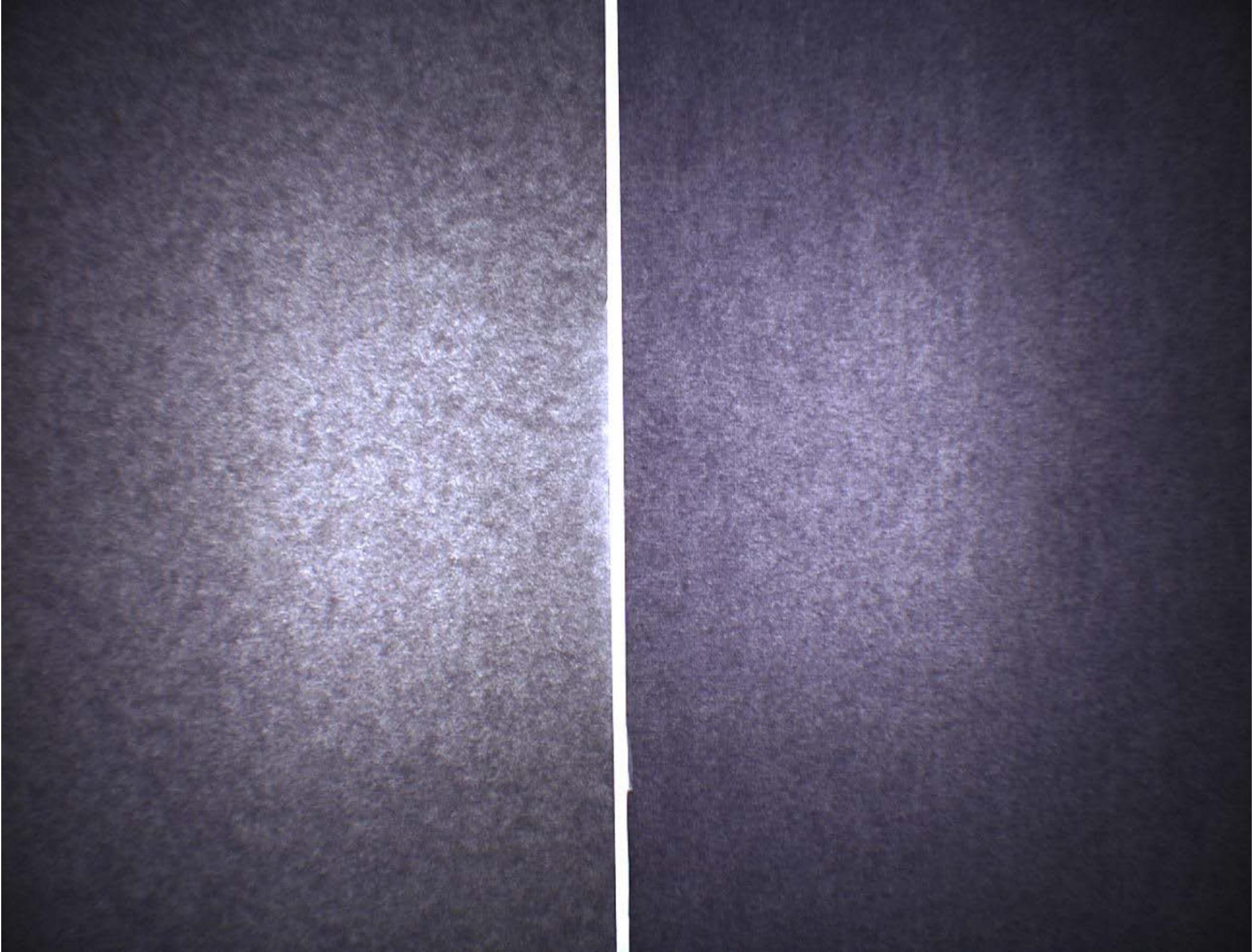


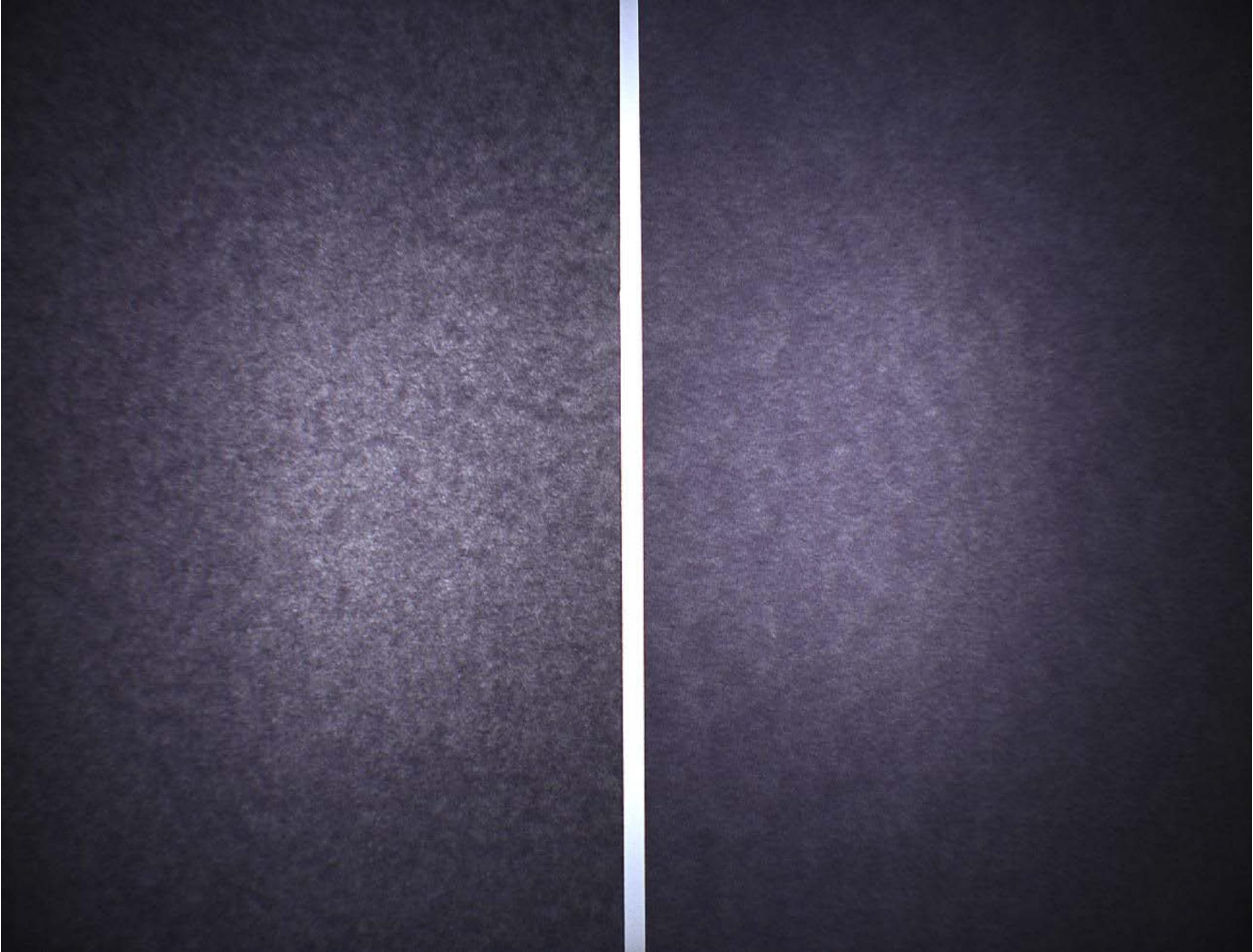


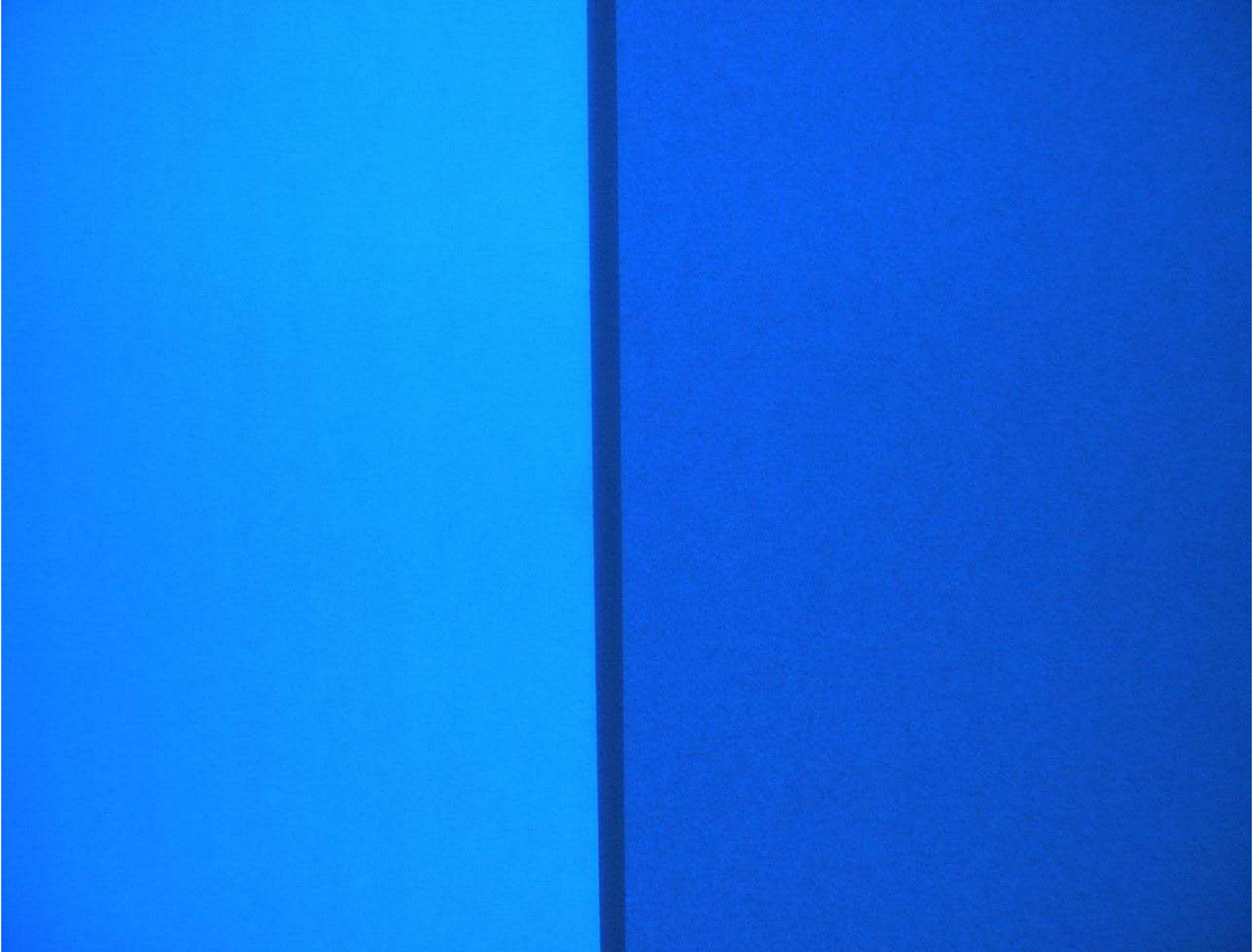


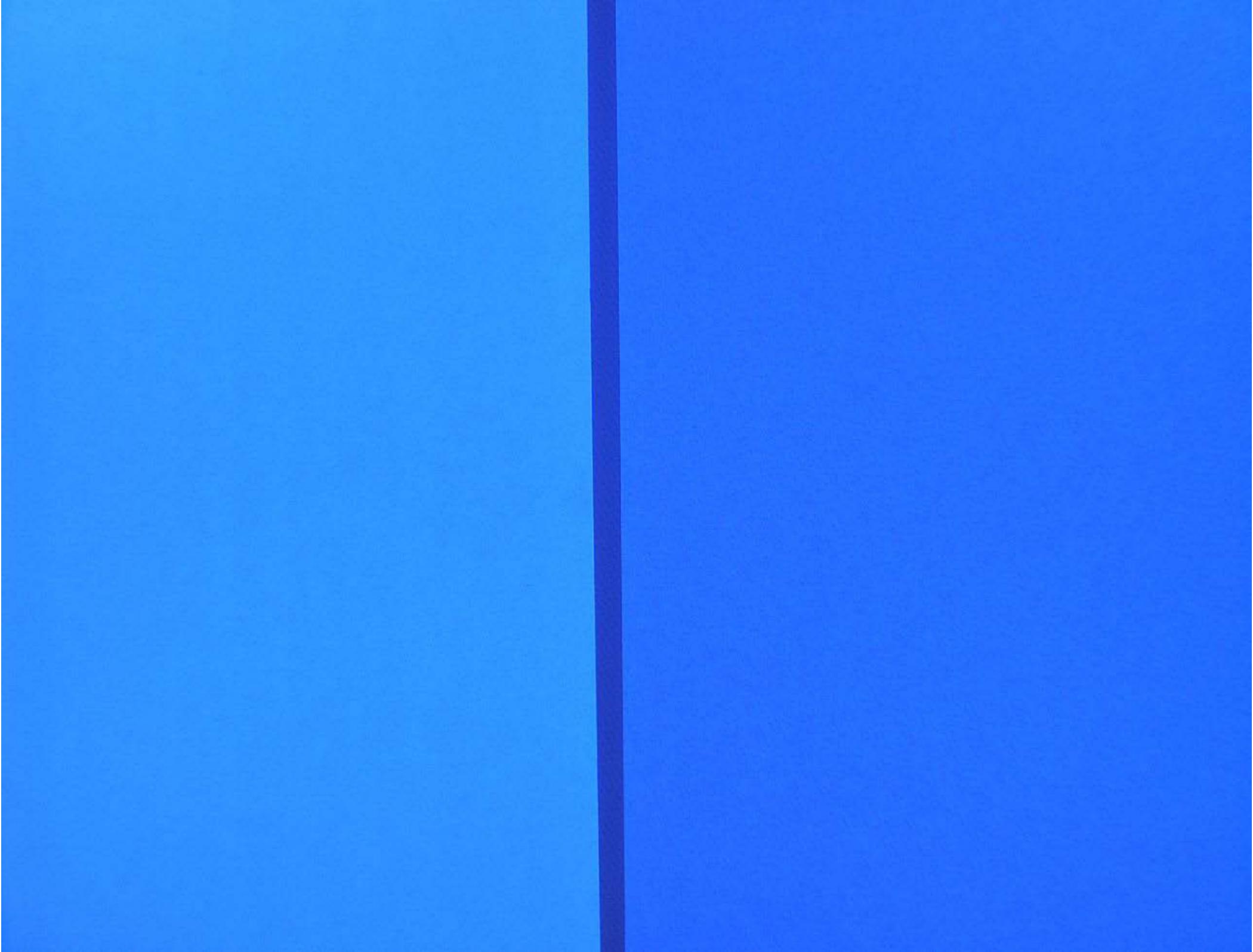


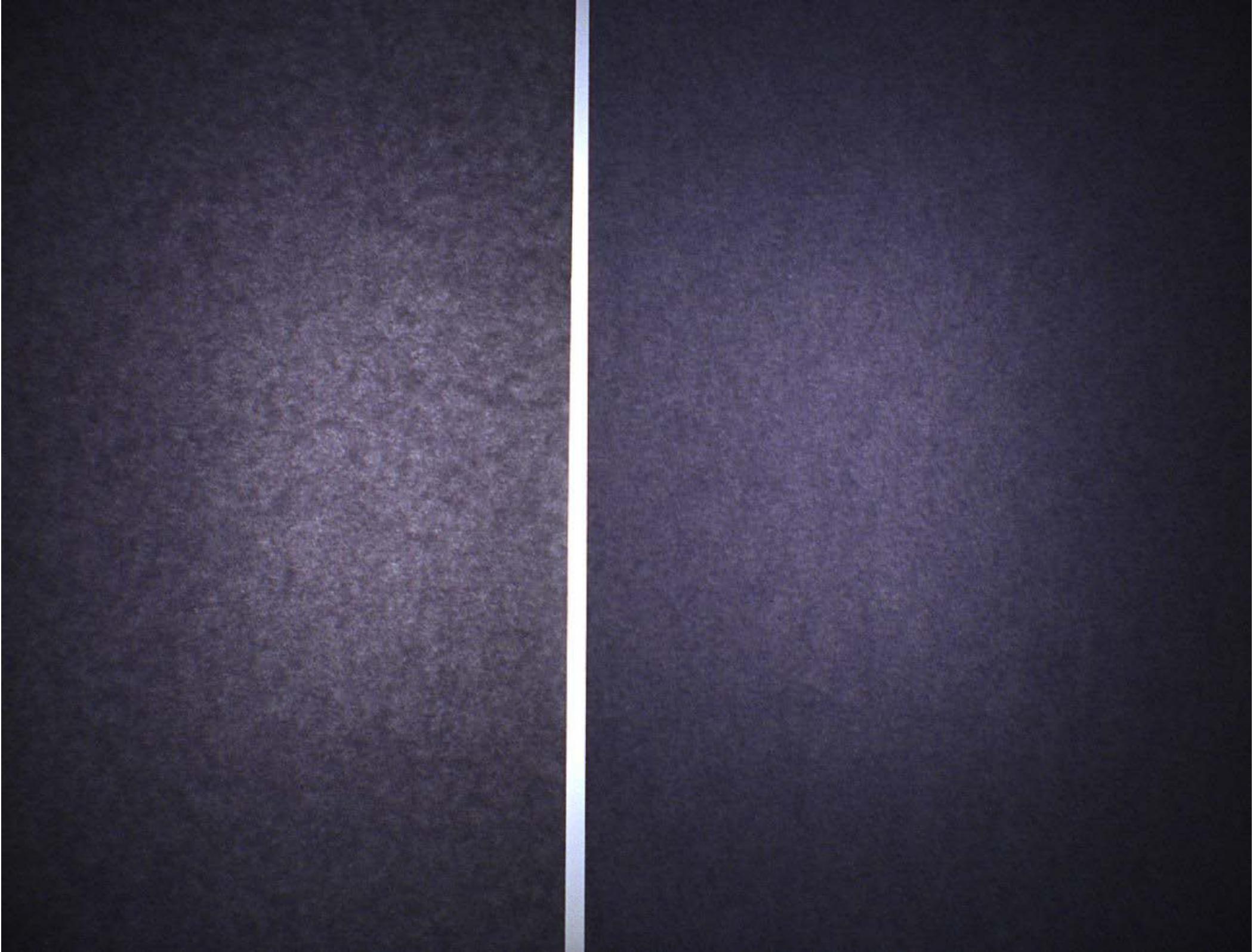














The forensic analysis of office paper using oxygen isotope ratio mass spectrometry. Part 1: Understanding the background population and homogeneity of paper for the comparison and discrimination of samples



Kylie Jones ^{a,b,*}, Sarah Benson ^a, Claude Roux ^b

^a Forensics, Australian Federal Police, P.O. Box 401, Canberra, ACT 2601, Australia

^b Centre for Forensic Science, University of Technology, Sydney, P.O. Box 123, Broadway, NSW 2007, Australia

ARTICLE INFO

Article history:

Received 1 December 2015
Received in revised form 27 January 2016
Accepted 22 February 2016
Available online 2 March 2016

Keywords:

Document paper
IRMS
Isotope
Document examination
Oxygen isotope

ABSTRACT

Isotope ratio mass spectrometry (IRMS) using carbon isotopes has previously been shown to be a robust and discriminating technique for the comparison of document papers. This study aims to examine the inter and intra sample variability for oxygen isotopes measured in standard 80gsm white document papers, to inform the comparison of document papers in forensic casework.

123 paper samples collected from Australia and New Zealand over a 24-month period were measured for their bulk oxygen isotopic abundance and were found to sit within a range of 15 ‰. A homogeneity study was undertaken which included examining the variability of samples at the sheet, ream and brand source levels. The results of this study were used to construct guidelines for sample comparison and as such, 95% confidence intervals were observed to be inappropriate for use given the high intra sample variability. Instead, a 1.4 ‰ discrimination range (0.7 ‰ either side of the measured value) was defined for use as a benchmark for discrimination when samples were measured in the same sequence. Utilising this value, 82% of the samples could be discriminated using a paired comparison, demonstrating a strong potential for use within forensic casework.

Crown Copyright © 2016 Published by Elsevier Ireland Ltd. All rights reserved.

[Production note: this paper is not included in this digital copy due to copyright restrictions.]

View/Download from: [Publisher's site](#)

* Corresponding author at: P.O. Box 401, Canberra, ACT 2601, Australia.

Tel.: +61 2 62036078.

E-mail address: kylie.jones@afp.gov.au (K. Jones).



The forensic analysis of office paper using carbon isotope ratio mass spectrometry – Part 1: Understanding the background population and homogeneity of paper for the comparison and discrimination of samples



Kylie Jones^{a,b,*}, Sarah Benson^a, Claude Roux^b

^a Forensic and Data Centres, Australian Federal Police, P.O. Box 401, Canberra, ACT 2601, Australia

^b Centre for Forensic Science, University of Technology Sydney, P.O. Box 123, Broadway, Sydney, NSW 2007, Australia

ARTICLE INFO

Article history:

Received 24 October 2012

Received in revised form 23 March 2013

Accepted 26 March 2013

Available online 24 April 2013

Keywords:

Isotope

Forensic

Paper

Document

IRMS

ABSTRACT

Isotope Ratio Mass Spectrometry (IRMS) has been shown to be a useful tool in the comparison of materials that are chemically identical either through man-made production processes or for materials that have been naturally produced. Paper therefore, is an ideal material for this type of measurement given that it is manufactured from a naturally produced product that can be difficult to discriminate based on physical feature comparison alone. To determine whether carbon isotopes are useful for discriminating document papers, 125 samples from Australia and New Zealand were collected over a 24-month period. When measured, a range of 8‰ was observed. A homogeneity study was undertaken to examine the range of values expected from paper sources including single sheets, single reams and multiple reams from the same brand. These results can also be used to suggest how best to sample from these different sources. After characterizing the natural variation of the material, a range of 1‰ was defined for use as a benchmark for discrimination. Utilizing this threshold, 68% of the 125 collected samples (when paired against each other) could be discriminated using the carbon isotope abundances alone. Additionally, correlation was observed when measured values were plotted against their production region of origin.

Crown Copyright © 2013 Published by Elsevier Ireland Ltd. All rights reserved.

[Production note: this paper is not included in this digital copy due to copyright restrictions.]

View/Download from: [Publisher's site](#)

* Corresponding author at: P.O. Box 401, Canberra, ACT 2601, Australia.
Tel.: +61 2 62036078.

E-mail address: kylie.jones@afp.gov.au (K. Jones).



The forensic analysis of office paper using carbon isotope ratio mass spectrometry—Part 2: Method development, validation and sample handling



Kylie Jones^{a,b,*}, Sarah Benson^a, Claude Roux^b

^a Forensic and Data Centres, Australian Federal Police, P.O. Box 401, Canberra, ACT 2601, Australia

^b Centre for Forensic Science, University of Technology Sydney, P.O. Box 123, Broadway, NSW 2007, Australia

ARTICLE INFO

Article history:

Received 5 November 2012

Received in revised form 20 April 2013

Accepted 7 May 2013

Available online 27 June 2013

Keywords:

Isotope

Forensic

Paper

Document

IRMS

Method validation

ABSTRACT

This paper describes the development and validation of a method for the analysis of office papers by measuring carbon isotopes using isotope ratio mass spectrometry (IRMS). The method development phase included testing protocols for storage, sample materials, set-up of the analytical run; and examining the effects of other paper examination procedures on IRMS results. A method validation was performed so that the Delta^{plus} XP IRMS instrument (Thermo Finnigan, Bremen, Germany) with Flash EATM 1112 could be used to measure document paper samples for forensic casework. A validation protocol that would meet international standards for laboratory accreditation (international standard ISO 17025) was structured so that the instruments performance characteristics could be observed. All performance characteristics measured were found to be within an acceptable range and an expanded measurement uncertainty for the measurement of carbon isotopes in paper was calculated at 0.26‰, with a coverage factor of 2. This method was utilized in a large-scale study, published as part one of this series, that showed that IRMS of document papers is useful as a chemical comparison technique for 80 gsm white office papers.

Crown Copyright © 2013 Published by Elsevier Ireland Ltd. All rights reserved.

[Production note: this paper is not included in this digital copy due to copyright restrictions.]

View/Download from: [Publisher's site](#)

* Corresponding author at: P.O. Box 401, Canberra, ACT 2601, Australia.
Tel.: +61 2 62036078.

E-mail address: kylie.jones@afp.gov.au (K. Jones).



The forensic analysis of office paper using carbon isotope ratio mass spectrometry. Part 3: Characterizing the source materials and the effect of production and usage on the $\delta^{13}\text{C}$ values of paper



Kylie Jones^{a,b,*}, Sarah Benson^a, Claude Roux^b

^a Forensic and Data Centres, Australian Federal Police, P.O. Box 401, Canberra, ACT 2601, Australia

^b Centre for Forensic Science, University of Technology, Sydney, P.O. Box 123, Broadway, NSW 2007, Australia

ARTICLE INFO

Article history:

Received 6 June 2013

Received in revised form 9 October 2013

Accepted 12 October 2013

Available online 19 October 2013

Keywords:

Isotope

Forensic

Paper

Document

IRMS

ABSTRACT

When undertaking any study of the isotope abundance values of a bulk material, consideration should be given to the source materials and how they are combined to reach the final product being measured. While it is demonstrative to measure and record the values of clean papers, such as the results published as part one of this series, the majority of forensic casework samples would have undergone some form of writing or printing process prior to examination. Understanding the effects of these processes on the $\delta^{13}\text{C}$ values of paper is essential for interpretation and comparison with clean samples, for example in cases where printed documents need to be compared to paper from an unprinted suspect ream.

This study was undertaken so that the source materials, the effects of the production process and the effects of printing and forensic testing could be observed with respect to 80 gsm white office papers. Samples were taken sequentially from the paper production facility at the Australian Paper Mill (Maryvale, VIC). These samples ranged from raw wood chips through the pulping, whitening and refinement steps to the final formed and packed paper. Cellulose was extracted from each sample to observe both fractionation and mixing steps and their effect on the $\delta^{13}\text{C}$ values. Overall, the mixing steps were observed to have a larger effect on the isotopic values of the bulk materials than any potential fractionation. Printing of papers using toner and inkjet printing processes and forensic testing were observed to have little effect on $\delta^{13}\text{C}$.

These experiments highlighted considerations for sampling and confirmed the need for a holistic understanding of sample history to inform the interpretation of results.

Crown Copyright © 2013 Published by Elsevier Ireland Ltd. All rights reserved.

[Production note: this paper is not included in this digital copy due to copyright restrictions.]

View/Download from: [Publisher's site](#)

* Corresponding author at: Forensic and Data Centres, Australian Federal Police, P.O. Box 401, Canberra, ACT 2601, Australia. Tel.: +61 2 62036078.

E-mail address: kylie.jones@afp.gov.au (K. Jones).