Misidentifications of Alphanumeric Characters in Serial Number Restorations

Waszczuk. J. R [ab]*, Peacock. L. [a], Chadwick. S. [a], Maynard. P. [a], Raymond. J. [c], Roux. C. [a]

[a] *Centre for Forensic Science, School of Mathematical and Physical Sciences, University of Technology Sydney, NSW*

[b] *Firearms Identification & Armoury Team, Australian Federal Police, Majura, ACT*

[c] *Science and Research Unit, Forensic Evidence & Technical Services Command, New South Wales Police Force*

*Joel R. Waszczuk, Forensic Firearm Examiner. Contact: joel.waszczuk@gmail.com

Misidentifications of Alphanumeric Characters in Serial Number Restorations

Chemical etching has long been established as a reliable means of recovering defaced (obliterated) characters such as serial numbers. Although successful restorations can present clearly defined characters, an assessment of 16 compiled proficiency test results show varied participant interpretations of samples marked with the same characters. Thus, a study was completed to determine if specific alphanumeric characters are more likely to be misidentified, comprising of a practical exercise and analysis of the aforementioned proficiency tests. The practical exercise involved 64 steel plates stamped with six randomised characters that were obliterated and chemically restored. Participants individually assessed the plates and provided their interpretations. A total of 39 participants of various backgrounds and experience in serial number restorations were used in this study. After comparing participant interpretations with the ground truth (actual stamped marking), characters more prone to be misidentified were determined. The proficiency test results showed the most frequently misidentified characters were the letters *B* and *N*, amongst a limited set. Whilst the practical assessment found a notable increase in misidentifications for the letters *G*, *Q* and *S*. Interestingly none of these characters were used in the proficiency tests. It is hoped that the results of this study raise awareness for forensic examiners undertaking restorations.

Keywords: Serial Number, Restoration, Interpretation, Misidentification, Alphanumeric Characters

1.1: Introduction

Serial number restoration is a practice often performed by forensic investigators such as firearm examiners and vehicle examiners, as both firearms and vehicles involved in criminal activity can have their identifiable markings (serial number or vehicle identification number) removed.

A serial number is an identifiable code used for the purposes of tracking and registration. This unique marking also provides an avenue for investigators to trace the items history, [1, 2, 3] such as previous owners. This information is commonly known by criminals and as such illicit items may be seized with a defaced serial number; often accomplished by grinding, filing, or milling [5, 11]. The efficient recovery and accurate interpretation of a defaced serial number is an important task for forensic investigators [4]. As the restoration on a metal surface using chemical etching involves the application of solid-state physics, chemistry, metallurgy and engineering, [1] there are a significant number of factors that influence the success and clarity of the restored markings. The visible result is influenced by factors such as the material hardness [3], carbon content [6, 7], material thickness [8], contrast [8], lighting [8, 9], surface contours [5], *ghosting*, available resources, and stamping type [10], thus the restored markings may not always present as apparent as their original stamped form. *Figure 1* shows an example incomplete stamping due to the of off-axis stamping application.

1.1.1: Application Methods

Serial numbers can be affixed to an item by various methods including laser etching, engraving, electrochemical marking and stamping [12]. Die stamping is a common practice, particularly on firearms [2, 9]. Although there have been further advances in the technology, a significant number of firearms are marked by die stamping [5, 13].

FIGURE 1: AN EXAMPLE OF OFF-AXIS SERIAL NUMBER MARKING FOUND ON A LITHGOW MANUFACTURED FIREARM

1.1.2: Die Stamping

The die stamping process involves applying a mirrored work piece (stamp) to the metal's surface, with sufficient force to deform the material [5]. With enough force the material undergoes plastic deformation, resulting in a permanent cavity in the shape of the stamp die [2], however, the true extent of the alteration runs deeper than is visible to the naked eye [9, 14]. The process alters the substrate area immediately around and below the stamp,[4] changing the materials hardness, strength, magnetic, electrical and chemical properties [13, 14]. {*see Figure 2*} Hence, even if the serial number is visibly removed, due to the disturbances in the subsurface structure around the indentation, important detail may be retained [15].

Die stamp characters can be differentiated into three classes based on their shape, being *'sharp faced'*, *'flat faced'* or *'open faced'* [5, 12]. Polk (1975) has also shown that the shape of the die impacts the depth of deformation (the area of available information for a successful recovery) to the material, determining that blunt faced dies produce subsurface deformation at greater depths. However, there appears to be no previous research into assessing if specific alphanumeric characters create a greater depth of deformation compared to others. Thus, it is not known if some characters are more likely to be successfully restored based on their shape alone.

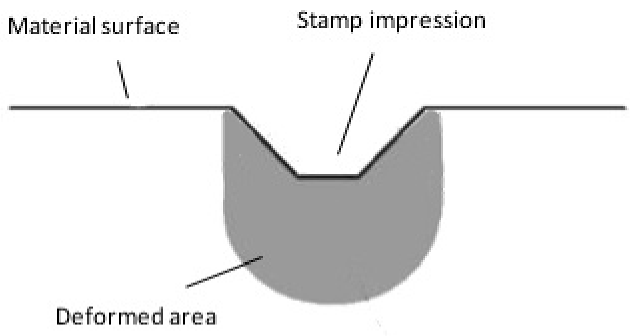


FIGURE 2: ILLUSTRATION OF SUBSURFACE DEFORMATION FROM SERIAL NUMBER STAMPING

1.1.3: Obliteration Methods

Serial numbers can be defaced (*often referred to as 'obliterated'*) as a means to prevent linking the item back to an individual crime or owner [14]. The *obliteration* process generally involves abrading the material surface at the marking until it is no longer visible [14]. Although there are many other obliteration methods such as centre punching, over stamping, heating and rusting, [14, 15], the most common methods encountered are grinding, filing or milling [5, 16].

1.1.4: Chemical Etching

Chemical etching is one of the most common methods utilised by investigators to restore a defaced serial number [1, 3, 13]. The process involves the application of a chemical agent such as an acid using the swabbing method [1]. The practice has been extensively studied throughout the years [1, 3, 9, 10, 15, 17, 18] with numerous chemical formulations tested [7, 10, 13, 15, 19, 20] and demonstrated to be capable of highlighting previously '*obliterated'* serial number.

Chemical etchants *restore* serial numbers by highlighting the difference between the compressed (*strained*) and non-compressed (*unstrained*) material [4, 21, 22], thus allowing for a visual representation of the material deformation. *{Refer to Figure 3}* These etchants work by exploiting the different chemical potential between the substrates deformed and undeformed regions [14, 15]. Bish (1979) states that deformed regions generally dissolve more readily in acids than undeformed, however there are exceptions [21].

Polk (1975) detailed the interactions of chemical reagents on metals [5], this study explained how chemical etchants can be preferential to attacking the deformed regions due to their *higher chemical potential* creating an '*etching pit'*. The etchants generally react with the deformed regions faster than the undeformed, and hence as the deformed areas dissolve a groove in the material is created. The use of light to reflect off the etching pit grooves create shadows allowing for a visual representation of the restored characters [14, 16]. Turley (1987) explains that the deeper within the metal, the more diffuse the deformation [10]. Practical experiments will show that this results in less of a defined groove and thus less of a change in the surface profile. This makes the restored characters more difficult to distinguish or isolate from the background (*undeformed*) metal. The less light reflectivity and less shadowing the more difficult the optical assessment and interpretation. Thus, the less deformity interacted upon by the etchants, the less defined the restoration, and hence the more ambiguous the interpretation.

FIGURE 3 - RESTORED SERIAL NUMBER *9G5QEV*

1.2: Interpretation

The recognition of characters is a skill which has been continually developed from a young age [25]. This practice later becomes a subconscious exercise due to its frequency of use. The interpretation of alphanumeric characters may be segmented into the following processes (*Figure 4)*. These processes may also occur simultaneously:

(1) The interpreter distinguishes the character(s) from background noise. Cognitively filtering out any information deemed not to relate to the characters.

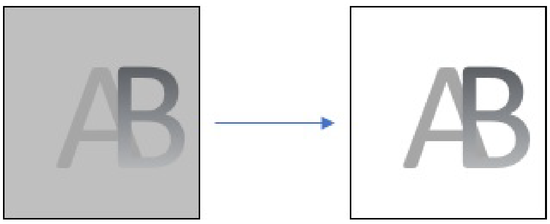
(2) Where multiple characters are present, the interpreter determines if any characters’ overlap, and both separates and isolates each character

(3) The interpreter recognising the features of characters, and further recognises where features are not present. For example, if a tail is present on the bottom right of a ‘O’ then the character is determined to be a capital ‘Q’.

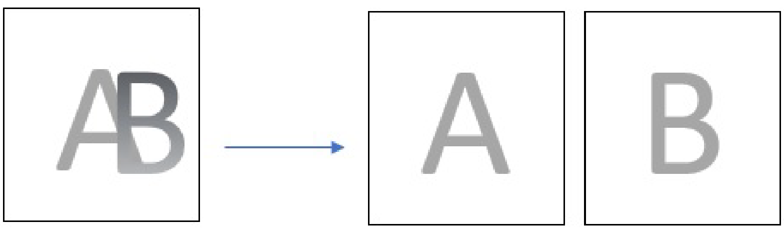
(4) A central point of the character is distinguished, and the relative positions of character features are oriented to the central point (zoning).

(5) The interpreter evaluates whether the features of the character conform to any pre-determined context and where relevant applies the appropriate context (e.g., if the character is determined to be from the English alphabet).

(6) The information recognised is applied to the context and the final classification of the character is determined.



1. Distinction of character from background noise



2. Character isolation/segregation

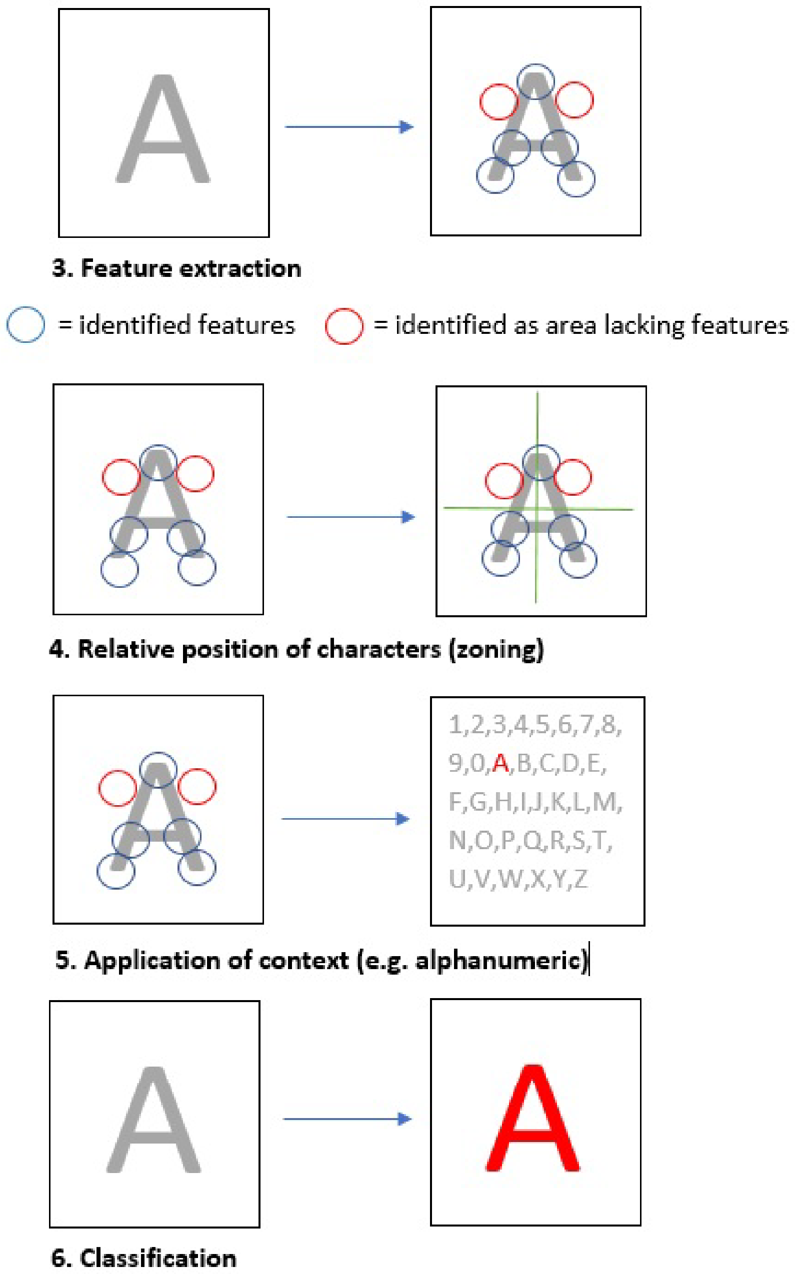


FIGURE 4: CHARACTER RECOGNITION PROCESS

(ADAPTED FROM OPTICAL CHARACTER RECOGNITION MODELS) [25, 26]

However, humans are imperfect beings, and as such human involved activities often involve a degree of error. The interpretation of characters and text by human eye is one such example. Previous literature has stated that the human eye makes approximately 4% error when reading without context [26].

A publication by Shastay (2015), later expanded on by Grissinger (2017), reported on misidentifications of alphanumeric symbols (characters) in handwritten and computer-generated data for medication orders (pharmaceutical scripts).

Misidentifications were found to occur due to the similarities in the appearance of characters, for example a lowercase *l* can be interpreted as a *1* depending on the font. The authors provided further examples of misidentified characters including mistaking an uppercase letter *Z* as a numeral *2*; and a numeral *0* mistaken for a *4*, *9*, and *6* [27, 28]. Distinguishing between alphanumeric characters becomes more challenging as the similarities between the pair increase [25, 28]. By design of the English characters alone there exists higher rates of misinterpretations between characters. For example, a number *0* would be expected to be relatively easily distinguished compared to the number *4*, however more difficult to distinguish compared to the letter *O*, because of the similarities in discernible factors between the two characters. Due to these similarities the interpretation is also reliant on the features not presented. For example, the letter *O* is distinguished from the letter *Q* singularly by the tail towards the base of the *Q*. Although the above literature relates to handwritten characters the same errors can relate to the interpretation of stamped characters.

1.2.1: Serial Number Interpretation

Although there have been extensive studies on restoring defaced (obliterated) numeric and alphanumeric characters using various chemicals [1, 7, 10, 15, 18, 19], there appears to be limited published literature that assesses the errors for individual characters. Even considering it has been many decades since it was documented that misinterpretations of characters may occur [15]. NASA's 1977 report on serial number restoration by Treptow identifies specific issues with the chemical etching process that may lead to misrepresentations of the restored markings [15]. The report specifically mentions how the number *1* with a serif can be misidentified as an incomplete number *4*. Likewise, with a character appearing to be the number *3* being an incomplete number *8*.

A document produced as an identification guide for Michigan State Police [29] provides specific examples of common misidentifications of serial numbers on firearms.

For example, the guide explains that the forward slash ('/') which forms part of the serial number for JP Sauer & Sohn manufacture revolvers is '*frequently… confused for the number one*'. Similarly, the *Q* prefix commonly found in the serial number of Rohm manufacture RG31 model revolvers is frequently misreported as the letter *0*. Whilst the capital letter *I* (uppercase '*i*') in Taurus manufactured firearms is often mistaken for a *1* as the stamp for the letter *I* is a plain vertical line [29].

A paper by Rak (2019) discusses transcribing errors of vehicle identification numbers onto vehicle registers and police information systems. The paper identifies 'visual confusion' as an avenue for misidentification of the vehicles identifiers, listing a number of categories such as insufficient lighting, lack of knowledge of VIN structure, accessibility and visual impairment. Rak furthermore specifies some frequently confused characters to be the letter *U* for a *V;* a *H* instead of an *M*, the letter *O* for the number *0*, and the number *5* being misinterpreted to be the letter *S* [30].

If the abovementioned misidentifications are common for unaltered markings, the potential exists for these same to occur after these markings have been defaced and chemically restored, yet there appears no previous research assessing the frequency of these errors, their origins, or any mitigation strategies.

1.3: Proficiency Testing

Proficiency tests in serial number restorations are regularly completed by forensic examiners [23]. Collaborative Testing Services. Inc. (CTS) is a primary supplier for these proficiency tests. CTS disseminates two serial number proficiencies per year; the tests are completed using each laboratory's preferred method of recovery before the results are sent back. CTS collate the data and anonymously publicises participant results on their website [24]. By comparing the individual proficiency test results with the ground truth (actual characters used) a determination can be made as to which serial number characters are more likely to be misidentified, and what they are likely to be misidentified as.

Previous studies [1, 3, 9, 10, 15, 17, 18] have shown that etching agents can recover defaced markings such as serial numbers, however the ability of individuals to correctly classify these markings to their previously stamped characters has not been specifically assessed. When looking at the English characters (*A* to *Z*, *0* to *9*) it is apparent that some characters display a larger amount of similar features compared to others. One would expect that these similar characters would likely be more prone to misidentification during interpretation. As it is known that human error exists in any visual examination, an assessment of restored alphanumeric character interpretations was proposed to determine if specific characters were more prone to being misidentified. By contrasting each character's frequency of use against its misidentifications, interpretation trends could be distinguished.

2: Materials & Method

To determine common character misidentifications this study comprised of two parts; analysis of CTS compiled data of 16 proficiency test results, and a separate practical study in which serial numbers were created and obliterated on steel plates. The practical study involved participants that had been trained in serial number restorations as well as those that had not (novices) as a comparison. Ethics approval was completed through the University of Technology Sydney (ETH21-6002) and participants were required to consent to the collection of their responses and a brief background questionnaire.

2.1.1: Proficiency testing data collection

CTS proficiency test results of serial number restorations from 2014 onwards were collected (16 proficiency tests in total). Each proficiency involved the restoration of a random six-digit serial number on a steel or zinc bar. These were CTS reference numbers *14-5250, 14-5251, 15-5250, 15-5251, 16-5250, 16-5251, 17-5250, 17-5251, 18-5250, 18-5251, 19-5250, 19-5251, 20-5250, 20-5251, 21-5250* and *21-5251*. These results show each individual response which can be compared against the known (actual pre-stamped serial number). The results were summarised onto an excel sheet and the misidentifications of characters extracted from the data.

The 16 CTS serial number proficiency tests involved a total of up to 3870 participants (approximately 242 per proficiency). The number of participant's interpretations per character was also calculated as well as each character's representation (frequency) across the 16 tests. By comparing each character's representation against the average, a scaled percentage was also calculated.

2.1.2: Practical assessment

Steel sample plates were constructed using 4140 medium carbon steel and '300 Plus' mild carbon steel *(Table 1 and Table 2).* A circular cutter was used to create each sample plate to be approximately 50x50 mm with 7 mm thickness. One face of each plate (later to become the stamped face) was smoothed to a relatively uniform surface finish using a belt sander with 240 grit sandpaper. Each plate was then hand engraved with a reference marking and the plate thickness measured at five separate points using a digital micrometre gauge. The thickness and reference markings were recorded on an Excel spreadsheet.

Chemical Analysis (%)		**Characteristics**	
Carbon	0.40%	Tensile Strength (MPa)	900
Silicon	0.25%		
Magnesium	0.80%		
Chromium	0.90%		
Molybdenum	0.20%		

TABLE 1: 4140 STEEL PROPERTIES[1]

[1] Edcon Steel 4140 Round Bar Data Sheet - https://www.edconsteel.com.au/store/storage/pdf/4140.pdf

Chemical Analysis (%)		Characteristics	
Carbon	0.25%	Tensile Strength (MPa)	440-480
Silicon	0.50%		
Phosphorus	0.04%		
Sulfur	0.04%		

TABLE 2: 300 PLUS STEEL PROPERTIES[2]

Each plate was stamped with six random alphanumeric (serial number) characters, consisting of *A* to *Z*, *1* to *9*. Randomisation was accomplished using an online random alphanumeric generator. The distribution of characters used amongst all the sample plates were monitored on an Excel spreadsheet to ensure a consist frequency of representation. Each stamped plate was also engraved with a reference number which was correlated against the six-character serial number on the Excel spreadsheet. A hydraulic press and a sharp-faced style die stamp using Arial-type font was used to press stamp all sample plates *{see Figure 5 and 6}*. The 4140 steel plates were stamped at 300 kg/cm^2 (29.42 MPa), whilst the 300 Carbon Plus plates, being a softer material, were stamped at 150kg/cm^2 (14.7 MPa). This provided a visually similar depth of character impression. This was later measured using an Alicona G5 focus variation instrument. The sample plates were then run under a grinding wheel until the serial numbers are no longer visible. The plates were then re-measured using the digital micrometre, recorded and the difference between the initial depth and the ground depth calculated.

[2] Edcon Steel 300 Plus Square Bar Data Sheet - https://www.edconsteel.com.au/store/storage/pdf/300plus.pdf

FIGURE 5: EXAMPLE OF STAMPED SERIAL NUMBER PLATE

Literature on serial number restorations emphasise preparing the obliterated area by polishing [5, 10, 13, 14] to achieve the best results, however this is not always possible due to the obliteration method and depth. Therefore, to best replicate these scenarios some sample plates were polished prior to chemical etching whilst others were left with the gross grinding marks from the obliteration process. The polished samples were worked to a mirror-like finishing using Brasso® polish and sandpaper of increasing grit (400, 800 and 1200).

FIGURE 6: DIE-STAMP WITH CHARACTERS ABC123

Each obliterated plate was cleaned using Acetone then chemically etched to restore the serial numbers. Chemical restoration was completed using Copper Chloride Working Solution[3] with sparse amounts of 10% Nitric Acid Solution (primarily for contrast).

---

[3] *Note: Copper Chloride Working Solution is a variant of Fry's solution, comprising of 36% w/v Copper (II) Chloride, 48% v/v Hydrochloric Acid and 40% v/v Water. Traditional Fry's Reagent comprises of Copper (II) chloride 90 g, Hydrochloric acid 120 mL, Deionised Water 100 mL.*

As per previous literature [1, 6, 8, 14] chemical etching was completed by wiping a cotton swab dipped in etchant gently across the defaced surface (the *swabbing technique*), under a fume hood. A hand-held multi-phase light source was used to provide oblique light to the surface for inspection. Each plate was treated with the etchants until a clear and complete six-character marking was identified, or to the point where the restorer perceives that the markings are beginning to deteriorate. Where no characters were appearing, the etching process was continued for up to two hours. If no result was visible after two hours the etching process was terminated.

After a sample plate was successfully restored, it was dabbed dry. Each plate was then assessed by up to six participants (including the restorer) who had no prior knowledge of the six-character serial number that had been obliterated. Participants were provided an information sheet prior to undertaking any interpretations which included information about the stamping, style and format of each serial number sample, and encouraged to use a handheld LED torch utilising various angles of light when completing their inspections. Participants requiring prescription glasses were encouraged to wear them for the plate inspections however no further optics (e.g. microscopes) were permitted. Each participant was instructed to compare the visual result produced after restoration to a reference sheet illustrating the font of each character used in the study. This process was to emulate a laboratory restoration were the examiner would seek knowledge of known serial numbers to assist with their assessment. Each participant's interpretation of the restored markings was then recorded on a results page.

The completed interpretation result forms were digitised onto an Excel spreadsheet, documenting each individual result and any discrepancies between the known truth (actual stamped marking for each plate). A misidentification is counted where the participant records a result which differs from the known pre-obliterated serial number. For example, the serial number plate originally stamped with *395HRZ* was identified by one participant as *395H87*. Each misidentification was recorded and totalled so that commonalities could be identified.

3: Results & Discussion

3.1: CTS Proficiency test assessment

Proficiency test results show a level of overall competency with restored character interpretations, however, there exists some degree of variations in participant results. The summarised results are shown in *Table 3* below. The accuracy of interpretations is given as a percentage for each character. The number of each misidentification is shown in brackets. These results are illustrated in Figure 7, with the highest misidentified characters represented in red.

TABLE 3: CTS CHARACTER MISIDENTIFICATIONS

Character	**Representation in proficiency tests**	**Misidentification (scaled %)**	**Misidentification (Most frequently identified as)**	**%**	**Also Misidentified as**	**%**	**Further Misidentified as**	**%**
A	5.25%	0.63%	4 *(4)*	50%	9, 0, B, D	12.50%	-	-
B	3.90%	6.42%	8 *(35)*	78%	5 *(3)*	7%	3, C, 6, 2	2%
C	3.36%	0.77%	0 *(2)*	50%	E, 2	25%	-	-
D	6.94%	0.94%	0 *(21)*	100%	O	-	-	-
E	3.02%	0.71%	7, 8 or C *(1 each)*	33%	-	-	-	-
F	5.88%	1.13%	E *(10)*	56%	1 *(4)*	22%	2, 6, 7, P	5%
H	3.33%	0.20%	M *(2)*	50%	U	25%	11	25%
J	6.31%	1.42%	3 *(15)*	58%	O *(7)*	27%	C	8%
K	6.93%	0.23%	2, 8, E, R or H *(1 each)*	20%	-	-	-	-
N	4.11%	11.78%	H *(58)*	63%	5 *(17)*	18%	K *(7)*	7%
0	2.55%	0.33%	9 *(1)*	100%	*'O' (3 times)***	-	-	-
1	5.80%	0.06%	4 *(1)*	100%	*'I' (3 times)***	-	-	-
2	5.19%	0.56%	8 *(6)*	86%	O	14%	-	-
3	6.24%	0.22%	8 *(3)*	50%	6, 9, or 5	17%	-	-
4	7.03%	0%	-	-	5	-	-	-

5	5.07%	0.34%	B *(3)*	75%	G	25%	-	-
6	3.27%	0.20%	8 *(1)*	100%	-	-	-	-
7	9.77%	0.05%	2 *(2)*	100%	-	-	-	-
8	0.65%	0%	-	-	-	-	-	-
9	5.38%	0.52%	0 *(2)*	29%	D, G, 3, 8, B	14%	-	-

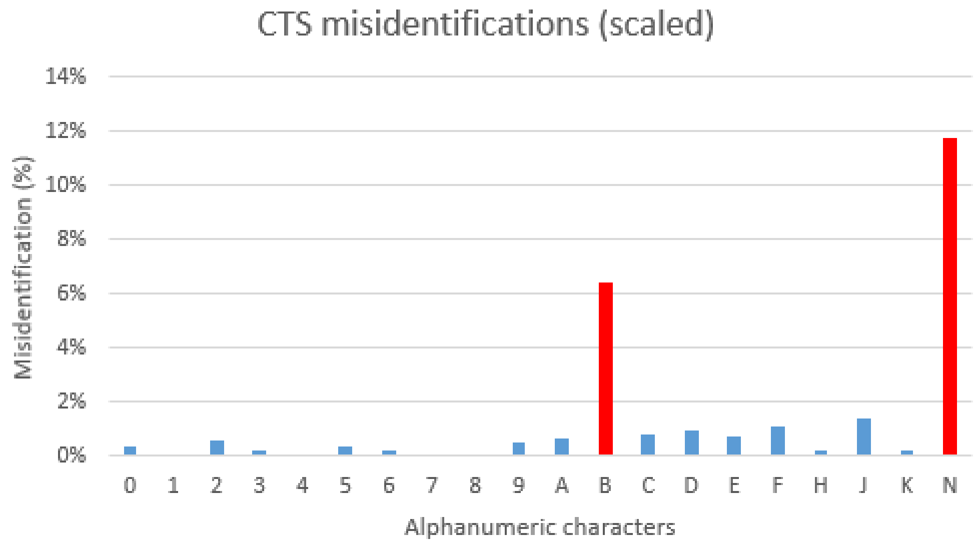


FIGURE 7: CTS CHARACTER MISIDENTIFICATIONS GRAPH

The data analysis in this study varies from that completed by CTS with regards to the following: The CTS proficiency results include instances where participants do not enter a result, or enter a character indicating the result is unknown, e.g., a '*', '?', '-', or '/'. These results are factored into the error by CTS. However, in this study, they are not recorded as a misidentification but rather taken to be a nil result. Where a participant recorded an either/or answer, for example., '*8* or *0*', only the first response was used in this study. Furthermore, results that appeared to be administrative errors were also not considered a misidentification. For example, in proficiency test 19-5251 the serial number stamping was known to be *5F3C7K*, however participant recorded the result as *5FC37K*. Participant results that were completely incorrect (all six characters) were also not counted as misidentifications. For example, also in proficiency test 19-5251, a participant result was recorded *3E42B8*.

The collation of results showed that the number *8* was used the least across the 16 proficiencies. The character was only used once in *test 16-5251*. This provided a representation figure of 0.65% which may be a reason why the character was not misidentified by any participant. Conversely the number *4* had one of the highest representation rates at 7.03%, however it also was not misidentified with any other character across its use in six proficiency tests (15-5250, 16-5250, 17-5250, 18-5251, 19-5250 and 20-5251).

Of the up to 3870 participant's results collected, the letters *B*, and *N* had the highest misidentification rate, with 6.42%, and 11.78% (scaled). The letter *B* was most often misidentified as an *8* (78%), or *5* (7%) but had also been misidentified as a *2*, *3*, *6* or *C*, (2%). Whilst the letter *N* was misidentified as either a *H* (63%), *5* (18%), or *K* (7%).

The evaluation of the CTS proficiencies was useful in that it provided specific examples of restored character misidentifications across many jurisdictions and participants. The CTS results show that most participants preferred chemical etching although other methods such as magnetic particle inspection were occasionally used. The comparison of results shows a large discrepancy between each proficiency test. For example, *test 21-5250* has a low of 60.4% of participants correctly identified the 4th character as an *N*. The overall results for this proficiency gave a 72% correct result (an average for all six characters). Whilst test 15-5250 showed 100% of participant results correctly identified all six characters. *Test 15-5250* also included the letter *N* in the proficiency. A probable reason for this disparity may be the depth of character stamping or depth of removal (obliteration). These depth measurements are not publicised by CTS.

A further point of interest involves CTS *test 18-5250*. This proficiency relates to the serial number restoration of a steel plate previously stamped with *K1J9D6*, before it was obliterated (defaced) by a milling machine. Prior to being disseminated to participants the plates were sent to three verification laboratories and 'restored'. Two of the verification laboratories successfully restored the characters and correctly recorded the previously obliterated characters *K1J9D6*. However, the third verification laboratory recorded the characters as *K1J9<u>0</u>6*. The test was subsequently disseminated and completed by 296 participants using a range of restoration techniques and chemicals. An assessment of the participant responses found the large majority restored the serial number using chemical etching, particularly Fry's reagent (or similar) [24]. The misidentification of the fifth character *D* by the verification laboratory was a common occurrence among the 296 participant responses. This character was misidentified as a *0* eighteen times in the proficiency test.

** In the table above shows instances where the number *0* was marked by CTS as an error if the participant result was perceived to be the letter *0*. Similarly, with the letter *I* for the number *1*. The CTS proficiencies do not include the letters *O* or *I*, therefore the result designation was strict to this criterion and thus no discretions were provided to allocate these results to the numerical counterparts. As participant results are anonymous it cannot be determined if the participants intended to record these results as letters rather the numbers.

3.2: Practical study assessment

Of the 64 plates chemically restored, 54 were constructed of 4140 steel, and 10 were constructed using the 300 Plus steel. 57 plates had all six pre-stamped characters identified by the restorer; whilst two plates had only five characters identified; one plate had only four characters identified; two plates had only three characters identified, one plate had only two characters identified and remaining plate had only one character identified during restoration.

Each plate was assessed on average by five individual participants (including the primary restorer undertaking the restoration), to a total of 322 plate assessments. As each plate contained six alphanumeric characters, this meant that 1932 individual assessments of chemically etched characters could be completed (64 plates x 6 characters x 5.03 interpretations), by a total of 39 participants of varied training and experience in serial number restorations. 18 of which had some degree of training in serial number restorations whilst 21 had no training.

A percentage of misidentifications per character was calculated by comparing the number of incorrect participant results against the total number of interpretations completed for that character. For example, the number *2* was assessed a total of 81 times in the study and misidentified 4 times (4.9%). Each character (*1* to *9* and *A* to *Z*) was used on average 11 times across the 64 plates. As there was on average five individual participant assessments per plate, this meant that each character was assessed approximately 55 times in the study. However, as the number of interpretations varied between each character based on its frequency of use and the availability of participants, the results were scaled.

Table 4 shows the scaled percentage of error for each alphanumeric character along with its three most frequent misidentifications. The number of each misidentification is shown in brackets. Where multiple characters share the same number of misidentifications, the result has been grouped together. For example, the number *8* was misidentified four times. The misidentifications were the characters *3*, *2*, *6*, and *B*. Thus, they each present 25% of the total misidentifications for the number *8*. These results identified an increase of misidentifications of certain characters; particularly the letters *G* (62%), *Q* (50%), and *S* (50%). Furthermore, this study found no misidentifications of the numbers *6*, *7*, and the letters *H* and *V*. *Figure 8* below further illustrates these results, with the highest misidentified characters shown in red.

TABLE 4: PRACTICAL STUDY CHARACTER MISIDENTIFICATIONS

Character	Representation (%)	Misidentification % Scaled	Misidentification (Most frequently identified as)	%	Also Misidentified as	%	Further Misidentified as	%
0	4.20%	21%	G *(4)*	*23.50%*	Q *(3)*	*18%*	6 *(2)*	*12%*
1	3.40%	12%	4 *(4)*	*36%*	I *(4)*	*36%*	T, U, O *(1)*	*9%*
2	4.20%	3%	R, 5, 4, 8 *(1)*	25%				
3	4.70%	13%	8 *(7)*	50%	8 *(3)*	21%	2, S, 5, R *(1)*	7%
4	3.40%	10%	Z *(3)*	50%	A, 8, I *(1)*	17%		
5	3.60%	3%	Z, 2 *(1)*	50%				
6	3.90%	0%						
7	3.90%	0%						
8	4.40%	3%	3, 2, 6, B *(1)*	25%				
9	3.60%	11%	8 *(4)*	44%	S, 3, R, 6, 0 *(1)*	11%		
A	2.10%	9%	4, M *(1)*	50%				
B	2.60%	19%	8 *(2)*	25%	3 *(2)*	25%	6 *(2)*	25%
C	1.60%	36%	0 *(2)*	40%	6 *(2)*	40%	G *(1)*	20%
D	2.90%	8%	0 *(4)*	100%				
E	2.60%	7%	B *(2)*	50%	8 *(1)*	25%	F *(1)*	25%
F	3.10%	8%	E *(2)*	33%	M, P, 7, H *(1)*	17%		
G	3.10%	62%	C *(9)*	35%	6 *(5)*	19%	0 *(4)*	15%
H	2.30%	0%						
I	1.80%	41%	F *(3)*	43%	T *(3)*	43%	P *(1)*	14%
J	2.30%	21%	0 *(4)*	36%	U *(2)*	18%	3 *(2)*	18%
K	2.30%	3%	A *(1)*	100%				
L	2.90%	7%	I *(1)*	33%	O, 1 *(1)*	33%		
M	2.30%	10%	H *(2)*	50%	7, 4 *(1)*	25%		
N	2.90%	10%	A *(3)*	50%	M, Y, P *(1)*	17%		
P	3.60%	5%	B, R, E, D, O *(1)*	20%				
Q	2.30%	50%	0 *(19)*	83%	G *(3)*	13%	8 *(1)*	4%
R	2.10%	33%	8 *(4)*	36%	P *(2)*	18%	9 *(2)*	18%
S	2.30%	50%	8 *(7)*	54%	6 *(3)*	23%	3 *(2)*	15%
T	1.80%	7%	1, I *(1)*	50%				
U	2.30%	9%	7, 6, M *(1)*	33%				
V	1.60%	0%						
W	1.60%	25%	Y, 2, K, O *(1)*	25%				
X	3.10%	8%	Y *(3)*	60%	Z, 1 *(1)*	20%		
Y	2.60%	3%	7 *(2)*	100%				
Z	2.30%	9%	7 *(4)*	80%	5 *(1)*	20%		

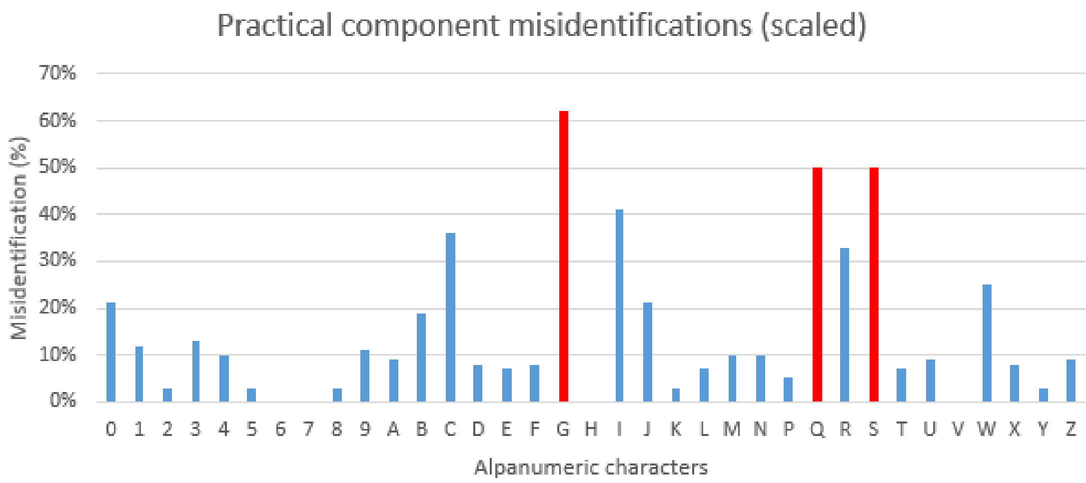


FIGURE 8: PRACTICAL STUDY CHARACTER MISIDENTIFICATIONS GRAPH

It was hypothesised that characters bearing the most similarities of their identifiable features would likely have the highest rate of misidentifications. For example, the number *0* and the letter *Q* are separated by only the tail of the *Q*. Thus, it was predicted that the results of this study would show elevated error rates between these two characters. Whilst characters bearing the least similar features would be unlikely to be mistaken. For example, the number *4* from the letter *W*. The results of this study illustrate this for almost all characters. Interestingly some characters were correctly interpreted in every instance. These were the numbers *6* and *7* and the letters *H* and *V*. Considering the similarities these characters' bear to others this was not an expected result. For example, a *6* shares significant similarities to the letter *C*, whilst the number *7* shares similarities to the letter *Z*.

To determine if the elevated misidentification rates of the characters *G*, *Q* and *S* was a result of interpretation or the obliteration process, the depths of obliteration of the plates containing these characters was assessed. This is shown in *Figure 9*. The average depth of obliteration for the 64 plates was found to be 0.348 mm.

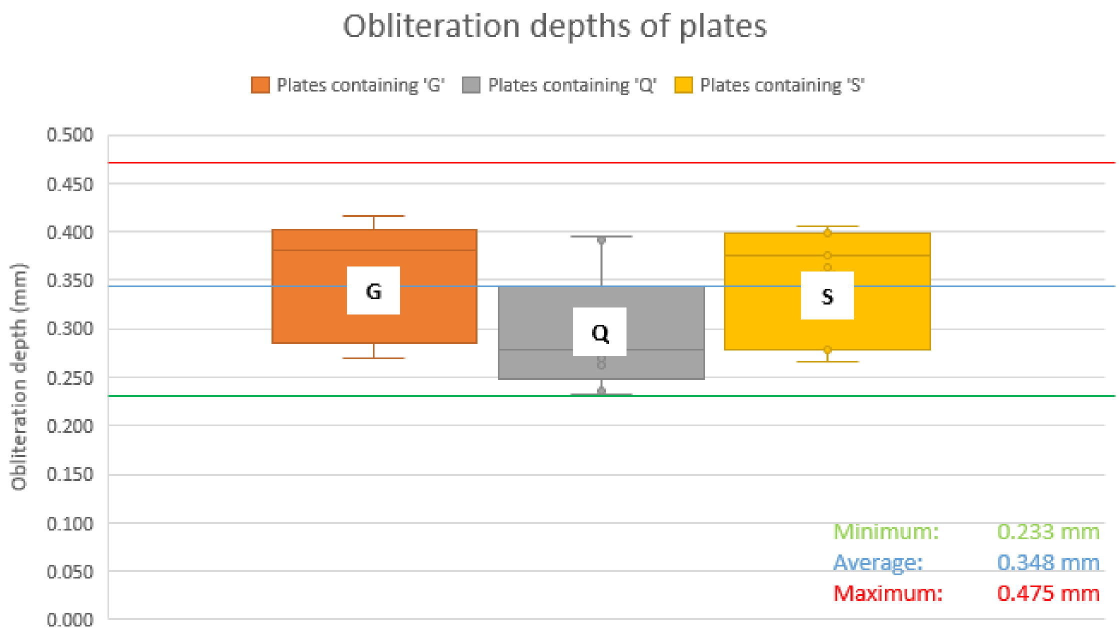


FIGURE 9: OBLITERATION DEPTHS COMPARED TO PLATES CONTAINING G, Q & S

The above results show that the depth of obliteration for the plates containing the letters *G, Q* and *S* were well within the average depth used in this study, yet each group had a good range of obliteration depth. (for example., the plates containing the letter *Q* ranged in obliteration depth of between 0.233mm and 0.391mm). It is also noted that the plates containing the letter *Q* only had three plates that were obliterated deeper than the average, which may have suggested that they would present clear restored markings, yet the misidentifications were prevalent.

The die stamp characters used in this study closely resembles an Arial-type font in that the characters predominately don't have serifs or tails. In this study the number *0* and the letter *O* were identical. The results simply list the character as the number *0*. Participants were advised that both characters were accepted to be the same. Therefore, in this study the characters bearing the most similarities were the *0* character and the letter *Q*. It should be noted that serial number fonts vary at the discretion of the manufacturer and that fonts containing characters with serifs and/or tails may be more readily distinguished from each other.

The number *0* die-stamp character was used 16 times (4.2% frequency); the restored character was inspected and interpreted 66 times over the course of the study. Of the 16 times, eight repeats were misidentified, but provided 17 individual interpretations. This gave a total error rate of the character of approximately 26% (unscaled). The number *0* was most frequently misidentified as the letter *G* (24%), the letter *Q* (18%) or number *6* (12%).

The letter *Q* die-stamp was used a total of nine times (2.3% frequency); the restored character was inspected and interpreted 49 times over the course of the study. Of the nine times, eight of which were misidentified, providing 23 individual interpretations of the character. This gave an error rate of the character of approximately 47% (unscaled). The letter *Q* was most frequently misidentified as the number *0* (83%), the letter *G* (13%) or the number *8* (4%).

The study found that overall, the letter *G* was most likely to be misidentified at 62% scaled, followed by the letter *Q* (50% scaled) and the letter *S* (50% scaled). Although it was hypothesised that the *Q* would be the most frequently misidentified character, due to the similarities between the letter *G* and other characters (such as *Q, 0, and C.*) this result was not surprising.

Although the study was successful in identifying misidentification rates of individual characters, there is dissimilarities between the study and common laboratory practice. For example, in the forensic laboratory if an examiner in unsure of the restored character, the results may be recorded as '*unknown*' or as a choice of characters, for example: '*The serial number was partly recovered and noted to be C275?. where the last character is either a 3 or an 8.'* Participants in this study were given an option to record what they believed each character to be, if they believed there was sufficient information to form an opinion. Where participants could not decide and recorded two potential options for a single character, only the first recorded option was noted in the study. This differs to restorations conducted in a forensic laboratory which would require higher levels of certainty before a determination is made. In future, a further study may be completed which assesses the accuracy of each participant's alternative response.

Forensic examiners may not always be privy to having a reference for the font, size, and format of the serial number they are attempting to restore. Participants in both the CTS tests and the practical component were also provided with a template to reference the size, shape and positioning of the characters. Having this information available is not always possible in laboratory conditions and thus the results may vary further because of this.

The practical component of this study required that participant interpretations remain confidential; even between the primary restorer and those subsequent participants. In most laboratories it is common practice for the primary restorer to form an opinion of the restored characters before seeking peer review from a second qualified examiner. The second examiner will then form a second independent opinion. On occasion where the results differ, the two interpretations are discussed. This creates an avenue for the primary restorer to explain how they came to their decision, which may include using photographs of characters that were clearer earlier in the restoration process. In situations such as this, the second examiner may reassess their opinion of the restored characters, thus reducing the overall error rate. Therefore, in this study an increased error rate may be present as this avenue for peer review was removed. Due to these differences the above studies are not an assessment of error rates for participants however is intended to provide insight to commonalities for misidentifications so that higher awareness is provided for examiners undertaking restorations.

There are several notable differences between the CTS assessments and the practical component. Firstly, the CTS proficiencies are aimed at assessing if the individual can undertake the serial number recovery procedure, whilst the practical component in this study was primarily assessing participant interpretations. Participants undertaking the chemical etching either already knew or were shown the procedure whilst the subsequent participants were only required to inspect and interpret each restored plate. There are also further dissimilarities. CTS sample plates only use the numbers *0* to *9* and letters *A-F, H, J, K* and *N* [24]. Interestingly the proficiencies do not include the letters *G, Q* or *S* which were the highest misidentified characters in the practical study. The restrictive use of characters by CTS is inconsistent with those found on manufactured firearms or Vehicle ID's. It is not known how CTS determined which characters would be included or excluded from their proficiency tests.

A misidentification was counted for any result that differed from the known serial number character, excluding nil results or occasions where an administrative error could be identified. For example, in proficiency *Test 19-5251* the serial number stamping was known to be *5F3C7K*, however participant recorded the result as *5FC37K*. This was not included as a misidentification. However, other administrative errors such as accidently recording the wrong character (e.g. recording a *B* instead of a *C*) could not be accounted for and may have been captured in this study as a misidentification. It is not known how many CTS participant results, or practical interpretation results were due to these errors.

4: Conclusion

Previous studies have proven the ability for etching agents to recover defaced markings such as serial numbers, however the ability of individuals in correctly classifying these markings to their pre-obliterated characters does not appear to have been previously assessed. These studies were successful in determining that certain characters are more likely to be misidentified. An assessment of CTS proficiency test results showed the most frequently misidentified characters were the letters *B*, and *N*, amongst a limited set. Whilst the practical component (which included an assessment of all alphanumeric characters) found the letters *G*, *Q* and *S* most likely to be misidentified. Interestingly CTS proficiencies do not include the letters *G, Q* or *S*. The letter *G* was found to be most frequently misidentified as the letter *C* (35%) or the number *6* (19%). The letter *Q* was most frequently misidentified as an *0* (83%) or the letter *G* (13%); and the letter *S* most frequently misidentified as the numbers *8* (54%) or *6* (23%).

It is hoped that the results of this study raise awareness for forensic examiners undertaking restorations, and that greater caution is used when chemically restoring items known to use these characters in their serial number. It is further hoped that this research is expanded on to cover other restoration procedures, metal compositions and character fonts and sizes to provide a more extensive assessment of alphanumeric misidentifications.

5: References

1. Richa LK, Gupta AK, Mishra KM. Development of new reagent for restoration of erased serial number on metal plates. Egyptian Journal of Forensic Sciences. vol. 3, no. 1, pp. 26-34, 2013.
2. Wightman G, Matthew J. Restoration of stamp marks on steel components. Forensic Sci Int. vol. 180, no. 1, pp. 32-6, 2008.
3. du Preez GT. Manual to Ballistics and Related Identification Methods. Institute of Criminology, University of South Africa, Pretoria. 2003.
4. Matthews J Howard – Firearms Identification Vol I. The University of Wisconsin Press, Madison, pp. 77-80, 1962.
5. Polk DE, Giessen B. Metallurgical aspects of serial number recovery. AFTE Journal Vol. 7, No. 2, 1975.
6. Yin, Kuppuswamy R. On the sensitivity of some common metallographic reagents to restoring obliterated marks on medium carbon (0.31% C) steel surfaces. Forensic Science International, 183(1), 50–53. 2008. https://doi.org/10.1016/j.forsciint.2008.10.009
7. Abdul Wahab, Mohamad Ghani. NI, Kuppuswamy R. An investigation into the suitability of some etching reagents to restoring obliterated stamped numbers on cast iron engine blocks of cars, Forensic Sci Int, vol. 223, no. 1-3, pp. 53-63, 2012.
8. Baharum Kuppuswamy R, Rahman AA. Recovering obliterated engraved marks on aluminium surfaces by etching technique. Forensic Sci Int, 177(2), 221–227. 2008. https://doi.org/10.1016/j.forsciint.2008.01.004.
9. Massiah E. A Compilation of Techniques and Chemical Formula Used in the Restoration of Obliterated Markings. AFTE Journal Vol 8, No 2, pp. 26-62, 1976.

10. Turley DM. Restoration of Stamp Marks on Steel Components by Etching and Magnetic Techniques. Journal of Forensic Sciences, JFSCA, Vol. 32, No. 3. 1987.
11. Young SG. The restoration of obliterated stamped serial numbers by ultrasonically induced cavitation in water. NASA technical memorandum. 1973.
12. Collins JM. Modern Marking and Serial Number Methods. AFTE Journal Vol. 31. No. 3, p. 311. 1999.
13. Fortini A. Restoration of Obliterated Numbers on 40NiCrMo4 Steel by Etching Method: Metallurgical and Statistical Approaches. Journal of Forensic Sciences Vol. 61. Issue 1. 2016.
14. Katterwe H. – Restoration of serial numbers. In: Stauffer, E., & Bonfanti, M. Forensic investigation of stolen-recovered and other crime-related vehicles. Chapter 7 Restoration of Serial Numbers. Elsevier. 2006.
15. Treptow RS. Methods for the Restoration of Obliterated Serial Numbers. NASA. 1977.
16. New South Wales Police Force website. About Firearms Registry. https://www.police.nsw.gov.au/online_services/firearms/about_us. Accessed 20-03-22
17. Baharum, Kuppuswamy R. Rahman AA. Recovering Obliterated Engraved Marks on Aluminium Surfaces by Etching Technique. Forensic Sci Int. 177(2), 221–227. 2008. https://doi.org/10.1016/j.forsciint.2008.01.004
18. Schultheis C. Serial Number Restoration on Titanium Firearms. AFTE Journal Vol. 48. No. 4. 2016.
19. Mohd A, Kuppuswamy R. Restoration of Engraved Marks on Steel Surfaces by Etching Technique. Forensic Sci Int, Vol. 171. Issue 1. pp. 27-32. 2007.
20. Peacock L, Waszczuk JR. Investigating the effectiveness of chemical etchants on medium carbon steel. AFTE Journal Vol 54. No. 1. 2022.

21. Bish RL. The action of Fry’s reagent on steel. Metallography. Vol 12(2), pp. 147–151. 1979.

22. Heard BJ. Restoration of Erased Numbers: Handbook of Firearms and Ballistics, Chichester John Wiley 12-96. pp. 214. 1997.

23. Individual responses from National Institute of Forensic Sciences (NIFS) survey on serial number restorations. 2021

24. Collaborative Testing Services, Inc. Proficiency Testing Summary Reports, 2021- https://cts-forensics.com/program-3.php - Accessed 05.04.2022

25. de Freitas A, Oliveira L, Aires S, Bortolozzi F. Zoning and Metaclasses for Character Recognition. Proceedings of the 2007 ACM Symposium on Applied Computing, 632–636. https://doi.org/10.1145/1244002.1244146. 2007.

26. Mantis J. An Overview of Character Recognition Methodologies. Pattern Recognition Vol. 19. No. 6. pp. 425-430. 1986.

27. Shastay A. Misidentification of Alphanumeric Symbols in Both Handwritten and Computer-Generated Information. Wolters Kluwer Health, Inc. Vol. 33. No. 6. pp. 338-339. 2015.

28. Grissinger M. Misidentification of Alphanumeric Symbols Plays a Role in Errors. P&T (Lawrenceville, N.J.), 42(10), pp. 604–606, 2017.

29. Michigan State Police. Criminal Justice Information Centre. Firearms Identification Field Guide. https://www.michigan.gov/documents/Firearms_Guide_98674_7.pdf - Accessed 29-3-22

30. Rak R. Vehicle identification number - anatomy of error occurrence. Journal of Physics. Conference Series, 1303(1), 12146. 2019. https://doi.org/10.1088/1742-6596/1303/1/012146

Special thanks to Eric Murray for his ongoing commitment to the development of science in Australia; Jen and the Science & Research Unit, Forensic Evidence & Technical Services Command, New South Wales Police Force; the Australian & New Zealand Forensic Science Society, and the University of Technology Sydney, for all your support with this research.