

## Abstract

*Objective:* Cognitive-behavioral therapy (CBT) is effective for the treatment of anxiety and related disorders (ARDs). Despite this, the use of best-practice CBT in clinical practice is low. While training and assessment strategies have been developed to improve this science-practice gap, both within the educational and clinical training space, many of the assessment techniques developed to enhance the use of best practice CBT remain impractical to use in busy training settings and are prone to bias. *Method:* The current study presents a preliminary evaluation of the CBT Decision Making Questionnaire for Anxiety and Related Disorders (CDMQ-A). The CDMQ-A contains vignettes covering seven diagnostic categories, each followed by three questions, resulting in a 21-item questionnaire designed to assess CBT decision making in the treatment of ARDs in adult patients. A sample of expert (N = 7) ( $M_{\text{age}} = 42.14$ ;  $SD = 5.64$ ; 57.1% female) and provisionally registered psychologists (N = 104) ( $M_{\text{age}} = 30.76$ ;  $SD = 8.32$ ; 82.7% female) completed the measure. *Results:* Experts indicated that the vignettes demonstrated satisfactory face and ecological validity. Results indicated that the CDMQ-A can effectively discriminate between experts and provisionally registered psychologists with the expert sample scoring significantly higher than the provisionally registered psychologists  $t(10.63) = 6.9$ ,  $p = .01$ ;  $d = 1.74$ ). *Conclusions:* Implications for training and clinical practice are discussed.

*Keywords:* clinical decision making; anxiety; psychometric properties; dissemination; CBT; clinical training

**Key Points:**

1. Widespread dissemination of CBT has not necessarily led to effective implementation in clinical practice, and tools are required to enhance this use.
2. This study aims to examine the psychometric properties of the CBT Decision Making Questionnaire for Anxiety and Related Disorders.
3. Findings suggest that the CBT Decision Making Questionnaire can effectively discriminate between experts and provisionally registered psychologists in clinical decision making.

## A Preliminary Evaluation of the CBT Decision Making Questionnaire for Anxiety and Related Disorders (CDMQ-A)

Cognitive-behavioral therapy (CBT) is effective for the anxiety and related disorders (ARDs), demonstrated both in efficacy trials (Hunsley, Elliott, & Therrien, 2014; Norton & Price, 2007; Olatunji, Cisler, & Deacon, 2010) and in effectiveness studies (Hunsley et al., 2014; Stewart & Chambless, 2009). Outcomes from CBT are also durable with patients maintaining their improvements for many years post-treatment (Wootton, Bragdon, Steinman, & Tolin, 2015). As a result, several key international bodies recommend that CBT be the first line of treatment for the ARDs (National Institute for Health and Care Excellence, 2011; National Research Council, 2009) and significant public funding is available for the provision of CBT to those with ARDs both nationally (Barrington, 2006) and internationally (Clark, 2011). Finally, as an evidence-based psychological intervention, CBT has been identified as one of the key therapeutic modalities taught in professional training programs throughout the world (Barrington, 2006; Hipol & Deacon, 2013; Kazantzis & Munro, 2011; Weissman et al., 2006).

Widespread dissemination of CBT does not necessarily lead to effective implementation or adherence in clinical practice (Patel, Kaufman, & Arocha, 2002) and research demonstrates that the use of CBT in clinical practice is poor, both by registered practitioners in the community (Cook, Biyanova, Elhai, Schnurr, & Coyne, 2010; McCausland, Paparo, & Wootton, 2020; Robertson, Paparo, & Wootton, 2020; Young, Klap, Sherbourne, & Wells, 2001) and within training programs (Weissman et al., 2006). This observation has been made across mental health disorders more generally (Berry, Rosenfield, & Smits, 2009; Mussell et al., 2000) and ARDs more specifically (Ehlers, Gene-Cos, & Perrin, 2009; Goisman, Warshaw, & Keller, 1999; Wang et al., 2005). The consequence of

this is that many individuals with ARDs may not be receiving an evidence-based intervention (Cook et al., 2010; Goisman et al., 1999; Young et al., 2001), or are receiving an evidence-based intervention in a modified or suboptimal manner (such as through limited number of sessions or in a manner that fails to align with theoretical underpinnings or evidence-based guidelines), which has been shown to reduce the effectiveness of the intervention (Broman-Fulks, Berman, Rabian, & Webster, 2004; Deacon et al., 2012; Schmidt et al., 2000; Shafran et al., 2009; Smits, Berry, Tart, & Powers, 2008). In fact, when specifically considering ARDs, research suggests that when not receiving an evidence based intervention, clients are instead receiving such interventions as supportive counselling (Ehlers et al., 2009) or complementary and alternative treatments (Wang et al., 2005). There is a growing body of literature that suggests that barriers to dissemination are varied, but likely include such factors as clinicians beliefs and knowledge or training in evidence based practice (Shafran et al., 2009).

To ensure adherence to evidence-based intervention it is important to develop strategies to monitor effective dissemination and implementation of CBT in clinical practice (Beidas, Mehta, Atkins, Solomon, & Merz, 2013; McHugh & Barlow, 2010). This appears particularly important within the training setting, where developing clinicians are first acquiring the knowledge and skills to effectively assess and treat ARDs. Clinical judgement, or the ability to make decisions through the integration of empirical evidence, observation and client data, has been shown to be impacted by experience (Ruscio & Stern, 2006; Spengler et al., 2009), which trainee clinicians do not typically possess. However, this has been debated in the literature by some who suggest that judgement cannot be expected to improve given the ambiguous nature of activities undertaken by psychologists (Dawes, 1994) and that experience may in fact worsen clinical judgement (Wedding, 1991).

Assessments of practical understanding typically include the use of short answer clinical vignettes (Myles & Milne, 2004) and case reports (Barnfield, Mathieson, & Beaumont, 2007; Keen & Freeston, 2008; McManus et al., 2010). Of clinical vignette tasks, only one standardized task exists to the authors knowledge. The Video Assessment Task (Myles & Milne, 2004) presents short video clips and asks clinicians to answer questions about symptoms, problem assessment and treatment. Initial examinations suggest very good inter-rater reliability for identification of symptoms ( $r = 0.97$ ), problem identification ( $r = 1.0$ ) and naming of appropriate CBT strategies ( $r = 0.94$ ) (Myles & Milne, 2004). Importantly however, the use of videos and expert examination of answers makes this technique particularly time and resource intensive.

Assessments of practical application of knowledge typically include the use of objective structured clinical examinations (OSCEs) which are routinely used in medical training settings (Epstein, 2007; Muse & McManus, 2013; Sholomskas et al., 2005). In recent years, the use of OSCEs as an assessment of clinical competence has increased (Kaslow et al., 2009; Pachana et al., 2011; Roberts & Norris, 2020). Research within the medical setting has indicated that OSCEs are as reliable as assessment of interactions with real clients (Epstein, 2007) and its utility as an assessment tool within the field of psychology is likely significant (Fairburn & Cooper, 2011). However, such techniques remain time intensive, and therefore potentially impractical to use as an assessment of clinical decision making in busy training settings or programs with large cohorts of students. In psychology, limited research exists to support the reliability and validity of this technique, however preliminary assessments suggest that the use of this technique as an assessment tool is considered to be a valid and realistic measure of competence in psychology training by students and staff alike (Hung et al., 2012; Sheen et al, 2015; Yap et al., 2012). Within medical training settings,

these techniques have demonstrated good reliability (Tudiver, Rose, Banks, & Pfortmiller, 2009; Wass, Jones, & Van der Vleuten, 2001).

Clinical practice assessments include assessor-rated treatment sessions and therapists' self-assessments. Techniques that require observation of live or recorded sessions, such as the transdiagnostic Cognitive Therapy Scale-Revised (Blackburn et al., 2001) and the disorder specific Multicenter Collaborative Study for the Treatment of Panic Disorder-Global Competence Item (Huppert et al., 2001), are an effective means of assessing comprehension, knowledge application, and allow for the provision of feedback on clinician strengths and areas of development. However, these are also particularly time and resource intensive and therefore may be impractical to use in busy training settings or for assessment of larger scale dissemination efforts.

Self- and assessor-rated measures of CBT competence exist, such as the Cognitive Therapy Self-Rating Scale (Bennett-Levy & Beedie, 2007), the Cognitive Therapy Adherence and Competence Scale (Barber, Liese, & Abrams, 2003) and the Manual-Assisted Cognitive Behaviour Therapy Rating Scale (Davidson et al., 2004) exist, however are particularly prone to bias in the form of both under- and over-estimation of skill by trainees (Brosan, Reynolds, & Moore, 2008; McManus, Rakovshik, Kennerley, Fennell, & Westbrook, 2012) or through the use of profession specific 'buzz words,' such as 'cognitive restructuring', which may allow trainees to identify correct answers based on recognition without necessarily having an understanding of the theoretical underpinnings. Taken together, whilst important developments, there appears a need for a means of evaluating knowledge application or clinical decision making, particularly within the busy training setting, that is time efficient and not susceptible to bias.

In an effort to overcome issues of previously developed tools, Carpenter et al. (2016) developed and evaluated a 24-item questionnaire (the Assessment of Clinical decision-

making in Evidence-based treatment for Child Anxiety and Related and Disorders; ACE CARD). The ACE CARD was designed to assess comprehension and clinical reasoning ability utilized by trainees when working with children with anxiety disorders. The tool consists of 12 clinical vignettes separated into two parallel forms that training clinicians complete using a four item multiple choice response format. Vignettes describe typical therapeutic situations experienced at different time points when working with anxious youth. Participants are asked to select the multiple-choice response, using a vignette matching style approach, that most reflects the CBT model of treatment.

An initial psychometric evaluation of the ACE CARD found that this tool is sensitive to clinical experience (Carpenter et al., 2016). For example, experts performed significantly better than trainees and the questionnaire was able to accurately distinguish between these populations (Carpenter et al., 2016). This research marks a preliminary but significant step forward in our ability to understand clinical decision making when working with pediatric patients with ARDs. However, there is currently no research that has looked at the assessment of clinical decision making when working with adult patients using a brief and easy to administer format, representing a key gap in current dissemination and implementation efforts.

The availability of such a tool may allow training providers to assess CBT clinical decision making in trainees and allow them to implement appropriate remediation strategies early in the clinician's development. Therefore, the aim of this study was to develop and evaluate the initial psychometric properties of the CBT Decision Making Questionnaire for Anxiety and Related Disorders (CDMQ-A), a 21-item questionnaire that measures CBT decision making in the treatment of the ARDs. It is expected that the CDMQ-A will effectively distinguish between expert and trainee clinicians, demonstrate high levels of reliability and validity, and high levels of sensitivity and specificity.

## Method

### Participants

A sample of clinicians with expertise in the treatment of ARDs ( $N = 7$ ) ( $M_{\text{age}} = 42.14$ ;  $SD = 5.64$ ; 57.1% female) and provisionally registered psychologists ( $N = 104$ ) ( $M_{\text{age}} = 30.76$ ;  $SD = 8.32$ ; 82.7% female) completed the measure. To be considered an expert, participants were required to have 1) a qualification equal to or greater than a Master's degree in Clinical Psychology; 2) possess more than 5 years' post-graduate experience; 3) indicate their clinical specialization as being within the ARDs; and 4) indicate CBT as their main theoretical orientation. Experts were identified through national and international university clinical training programs and specialist CBT clinics for ARDs. Experts were then invited to participate in this study by email, sent by the authors. As this part of the study used a convenience sample, it is not possible to report participation rate. The expert clinician sample consisted of 4 women and 3 men with a mean age of 42.14 years ( $SD = 5.64$ ), the majority of whom worked in a specialized clinic treating ARDs. Of the sample, 5/7 (71.4%) reported their highest level of qualification as holding a PhD and 2/7 (28.6%) reported having a Master's degree in Clinical Psychology. The sample on average reported 12 years of post-qualification clinical experience ( $SD = 5.45$ ).

The provisionally registered psychologist (trainee) sample consisted of 86 women and 18 men with a mean age of 30.76 years ( $SD = 8.32$ ). To be included in the trainee sample, participants were required to 1) be currently enrolled in a professional postgraduate psychology training program; 2) be located in Australia; 3) have received training in CBT as part of their coursework. Of the sample 32/104 (30.8%) were enrolled in a PhD or combined PhD and master's program, 62/104 (59.6%) were enrolled in a Masters of Clinical Psychology program and 10/104 (9.6%) were enrolled in a Masters of Professional



Psychology program. Participant characteristics for the expert sample and provisionally registered sample are outlined in Table 1.

As could be expected, results of an independent-sample  $t$  tests revealed that the two groups significantly differed in age [ $t(109) = -3.56, p = <.001$ ]. The trainee sample was significantly younger ( $M = 30.76, SD = 8.32$ ) than the expert sample ( $M = 42.14, SD = 5.64$ ), representing a large effect ( $d = 1.74$ ). One-way chi-square revealed no significant differences in gender  $\chi(1) = 2.79, p = .09$  between the two groups.

[Table 1 near here]

## Measures

**Demographics Questionnaire.** The demographic questionnaire asked participants to indicate their 1) gender; 2) age; 3) highest level of training; 4) registration status; 5) years of clinical experience; 6) theoretical orientation; and 7) clinical specialization (if fully registered as a psychologist).

**Generalized Anxiety Disorder – 7 item (GAD-7)** (Spitzer, Kroenke, Williams, & Löwe, 2006). The GAD-7 is a self-report scale designed to measure the severity of symptoms associated with generalized anxiety. All items are rated on a 4-point Likert scale and scores range from 0-21, with higher scores indicating more severe symptoms of anxiety. Internal reliability for this scale has been reported as excellent ( $\alpha = 0.92$ ) (Spitzer et al., 2006). In the current study, internal reliability was found to be  $\alpha = 0.89$ . For the purposes of this study, the GAD-7 was used as a measure of divergent validity.

**CBT Decision Making Questionnaire for Anxiety and Related Disorders (CDMQ-A).** The CDMQ-A was developed by both authors who have extensive experience in the assessment of ARDs, CBT for the ARDs, and the supervision of clinical psychology trainees delivering CBT. Both authors equally contributed to the development of the tool, with content drawn from typical presentations seen within university training clinics in Australia.

An initial set of eight vignettes, each reflecting one separate diagnostic category was developed. A set of three questions for each vignette were then developed in order to assess clinical decision-making skills, utilizing a four-item multiple choice response format (see Figure 1 for a sample vignette). Additional case information was included in each of the three questions that followed the initial case vignette. Vignettes and associated questions were designed to broadly assess decision making regarding assessment, cognitive and behavioral interventions. Additional questions were designed to assess knowledge of theoretical models, the importance of an evidence-based assessment, case formulation, intervention for subclinical symptoms, and stages of change. Together, questions were designed to assess key skills required to deliver a CBT intervention for common manifestations of ARD symptoms. Importantly, whilst diagnostic classification is utilized here as one factor that contributes to treatment decision making, it is acknowledged that clinicians must utilize a formulation driven approach to clinical decision making. Where possible, individual factors, such as cause and maintaining factors, are noted in the provided case information to reflect the importance of the case formulation approach.

Accurate responses were informed by scientific literature, evidence-based CBT treatment manuals and a CBT case formulation approach. Incorrect responses were derived from observations of common decision-making errors that trainees often make during their training, observed by the authors. The CDMQ-A underwent multiple reviews by authors and were then evaluated by six international clinicians with expertise in the ARDs, who were asked to rate vignettes and questions within each diagnostic category on face and ecological validity. Face validity was assessed by asking experts to identify, on a 10- point likert scale, how effective each measured what it purported to measure. Ecological validity was assessed by asking experts to identify, on a 10- point likert scale, the degree to which each case reflected clients seen in their clinical practice. Following expert feedback and review, one

vignette and the associated three questions were discarded due to concerns with measurement of decision making and ecological validity.

In the final version of the CDMQ-A, vignettes and questions related to 7 diagnostic categories were retained: 1) obsessive-compulsive disorder (OCD); 2) panic disorder (PD); 3) post-traumatic stress disorder (PTSD); 4) body dysmorphic disorder (BDD); 5) social anxiety disorder (SAD); 6) generalized anxiety disorder (GAD); and 7) agoraphobia (AG). Three-questions were presented for each vignette, resulting in a final 21-item measure. The questionnaire is scored by totaling correct responses to each diagnosis, with total scores ranging from 0 – 7. To achieve a perfect score, all three items linked to each vignette must be answered correctly. The CDMQ-A is available from the corresponding author upon request (see Figure 1 for a sample vignette).

After your comprehensive assessment you determine that John has a primary diagnosis of obsessive-compulsive disorder and a secondary diagnosis of major depressive disorder. John denies current suicidal ideation, intent or plan to harm himself. His obsessive compulsive symptoms include themes of contamination (with repetitive checking that items in his home are clean) and intrusive thoughts about harming others, specifically by stabbing them, which result in repetitive behaviours to cancel out the thoughts. Should you:

- Reassess because it sounds like John might have an autistic spectrum disorder
- Help John to understand that having bad thoughts is completely normal and everyone has them from time to time
- Do some values work with John where you can help him determine other positive attributes about himself
- Have John create a list of all items in his home that need to be checked for cleanliness. Once cleaned, he can check them off the list and thus prevent him from checking again

*Figure 1. Sample clinical vignette and one corresponding questions / response options from the CDMQ-A.*

## **Procedure**

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee (Western Sydney University Human Research Ethics Committee; H12072) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was

obtained from all individual participants included in the study. The study was conducted in two parts. Part one consisted of online questionnaire completion by a sample of expert clinicians. Experts were identified and invited to participate in the online questionnaire via direct email which included a link to the participant information sheet, consent form, demographics questionnaire, the CDMQ-A (including additional questions relating to face and ecological validity for each vignette) and the GAD-7. The online questionnaires appeared in fixed order and took approximately 30 minutes to complete.

Part two of the study consisted of online questionnaire completed by provisionally registered psychologists (i.e., trainees). Participants were identified and invited to participate via email through postgraduate psychology training programs and publically available list serves, and via advertisement on relevant social media websites. Emails and advertisements contained the study link which opened to the participant information sheet, consent form, brief demographic questionnaire, the CDMQ-A and the GAD-7. The online questionnaires appeared in fixed order and took approximately 20 minutes to complete.

### **Data analysis**

Assumption testing was undertaken prior to analyses. Due to small and unequal group sizes, non-parametric tests were used as appropriate. Independent-sample *t* tests and chi-square tests were conducted to assess group differences. Reliability of the CDMQ-A was assessed with Cronbach's alpha. Divergent validity was assessed via a Pearson correlation coefficient between the CDMQ and the GAD-7. Effect sizes were calculated and interpreted as suggested by Cohen (1992). Differences between group means were reported as: Small = .20, medium = .50, and large = .80 (Cohen, 1992). Effect size of correlations were interpreted as: Small = .10, medium = .30, and large = .50 (Cohen, 1992). Sensitivity and specificity were determined using a receiver operating characteristic (ROC) analysis. ROC analysis results were analyzed in line with recommendations by (Šimundić, 2009); good = 0.70-0.80,

very good = 0.80-0.90, excellent = 0.90=1.0. All analyses were conducted using IBM SPSS Statistical Software version 25.

## Results

### Validity

Consistent with prior research (Carpenter et al., 2016; Simpson et al., 2010), expert clinicians rated vignettes on both face and ecological validity on a 10-point Likert scale, as recommended by Nevo (1985). All vignettes included in the final questionnaire were reported as having acceptable face ( $M = 6.9$ ;  $SD = 2.31$ ; range: 0-10) and ecological validity ( $M = 7.41$ ;  $SD = 2.24$ ; range: 2-10), where acceptable was defined as receiving a score on this scale greater than six.

An independent-sample  $t$  test was conducted to investigate group differences on the total CDMQ-A between experts and trainees. Results indicated a statistically significant difference [ $t(10.63) = 6.9$ ,  $p = .01$ ]. The mean score on the CDMQ-A was higher in the expert sample ( $M = 6.29$ ,  $SD = .76$ ) than in the provisionally registered sample ( $M = 4.01$ ,  $SD = 1.69$ ), representing a large effect ( $d = 1.74$ ). Differences between correct and incorrect responses provided by the experts and trainees on each vignette were examined using chi-square. Results indicated that experts and trainees differed significantly on the vignettes related to OCD ( $\chi^2(1) = 6.03$ ,  $p = .01$ ), PD ( $\chi^2(1) = 8.18$ ,  $p = .004$ ), and GAD ( $\chi^2(1) = 5.29$ ,  $p = .02$ ). Overall, trainees were found to have made errors 61.5% of the time on OCD specific vignettes as compared to experts who made errors only 14.3% of the time; 55.8% of the time on PD specific vignettes as compared to experts who did not make any errors; and 44.2% of the time on GAD specific vignettes as compared to experts who did not make any errors.

To assess divergent validity, a Spearman's correlation coefficient was conducted between the total score on the CDMQ-A and the GAD-7. No significant relationship was found between these measures ( $r = .12$ ,  $p = .23$ ; small effect).

**Reliability**

Cronbach's alpha was .71, indicating acceptable internal consistency.

**Sensitivity/Specificity**

Receiver operating characteristic (ROC) analyses were conducted in order to determine the diagnostic sensitivity and specificity of the CDMQ-A. The area under the curve (AUC) was .89 (95% CI: .79 - .97), suggesting very good diagnostic accuracy. A cut-score of 6 provided the best balance between sensitivity and specificity (sensitivity = .86; specificity = .24). The positive predictive power was .55 and the negative predictive value was .99.

**Discussion**

The CDMQ-A was developed to evaluate clinical decision making in trainees relevant to psychological assessment and intervention for adults with ARDs. The aim of the present study was to evaluate the initial psychometric properties of the CDMQ-A in a sample of expert clinicians and provisionally registered psychologists. It was hypothesized that the CDMQ-A would effectively distinguish between expert and trainee clinicians, demonstrate high levels of reliability and validity, and high levels of sensitivity.

Expert clinicians reported acceptable face and ecological validity for the CDMQ-A items. Results found a statistically significant difference between results on the CDMQ-A by experts and provisionally registered psychologists and evidence of a significant association between expertise and total score on the CDMQ-A. Scores from the expert sample and trainee sample were significantly different on three of the seven disorder specific vignettes (GAD, OCD, and PD) and thus these vignettes may have the most utility in identifying competency in the treatment of ARDs, however further research is required.

The CDMQ-A demonstrated acceptable reliability ( $\alpha = .71$ ), excellent divergent validity when compared with the GAD-7, and very good utility as a means of distinguishing between expert and provisionally registered psychologists (AUC = .89). While a finding of

high specificity suggests that the CDMQ-A was able to correctly identify those who engage in accurate clinical decision making, low specificity suggests that the tool had greater difficulty identifying incorrect clinical decision making. **TOOL CAUTIONS** However, it will be important to examine the convergent validity of the measure in future studies. Future studies may also wish to examine divergent validity using measures designed to assess a similar but still distinct construct. As this is the first study to investigate the psychometric properties of the CDMQ-A, it is important that future research replicate these findings. However, the results of the current study are consistent with studies that have used a similar methodological approach, when applied to a child anxiety population. Here, Carpenter et al. (2016), found that the ACE CARD was able to distinguish between an expert and trainee sample, had good internal consistency, and moderate utility as a means of distinguishing between expert and provisionally registered psychologists. Taken together, these studies provide preliminary support for the use of a vignette based tool to assess clinical decision making in clinical psychology trainees. To the authors' knowledge, the CDMQ-A is the first tool of its kind to assess clinical decision making with an adult anxiety population using a vignette approach.

The CDMQ-A represents an important step forward in the evaluation of clinical decision making in psychology trainees specifically when working with adult patients with ARDs. Administration of the CDMQ-A allows training institutions to evaluate the application of knowledge to intervention. When deficits are identified through incorrect responses, additional training and resources can be provided specific to these domains, with the aim of further supporting the development of clinical decision-making skills in a closely supervised setting. Given that research has highlighted that clinical decision making improves with experience (Ruscio & Stern, 2006; Spengler et al., 2009), which training clinicians do not as yet possess, a tool which has the ability to identify and remediate errors in clinical decision

making early in training is likely to further support effective implementation efforts. While the results are preliminary, the tool has several important implications for dissemination and implementation of evidence-based practice for the ARDs, identified as an area of need (Beidas et al., 2013; McHugh & Barlow, 2010). First, given that this tool is brief, easy to administer (paper and pencil format or online) and score, and requires minimal staff input, the CDMQ-A is likely more practical to use in busy training settings, meeting a key area of identified need (Muse & McManus, 2013). Second, by requiring the application of knowledge to assessment and treatment decision making, this tool is designed to extend upon existing measures of knowledge assessment. Third, through the utilized question response format and removal of ‘buzz words’ (such as ‘exposure hierarchy’ or ‘cognitive restructuring’), the tool arguably provides a better assessment of comprehension of theory underpinning evidence-based clinical decision making, and thus may not be prone to the bias experienced by other similar measures (Brosan et al., 2008; McManus et al., 2012). Fourth, while it requires further investigation, the tool may be able to evaluate the effectiveness of these training efforts by examining change after coursework devoted to the treatment of ARDs. Finally, the CDMQ-A may be used as a template for the development of decision-making tools for other common adult psychological disorders, such as depressive disorders, suggested as an area of need more broadly (McHugh & Barlow, 2010; Muse & McManus, 2013).

While overall findings of the present study provide preliminary support for the psychometric properties of the CDMQ-A, a number of limitations should be noted. Firstly, generalizability of the findings may be limited due to the sample including only provisionally registered psychologists practicing in Australia. It remains unknown whether cultural variations would impact the utility of this tool in other English-speaking countries. Future research may wish to further investigate the psychometric properties of this tool in both



Australian and international contexts. Secondly, this study used only a small expert sample. Whilst past research has also utilized a similar sample size (Simpson et al., 2010), future studies may wish to administer this questionnaire to large samples of training, generalist, and expert clinicians in the future. Thirdly, clinical decision making is complex and the CDMQ-A acts only as one brief measure of clinical decision making and does not evaluate the application of this clinical decision making on assessment or intervention in clinical practice with real patients. The goal of the tool is to identify trainees who may be particularly lacking specific skills in CBT decision making, which may require remediation, rather identifying those whose skills are superior. Future research may wish to compare the CDMQ-A to other practice based measures of clinical decision making such as the Therapy Process Observational Coding System (McLeod, Islam, & Wheat, 2013; McLeod, Smith, Southam-Gerow, Weisz, & Kendall, 2015). Finally, as only one measure of clinical decision making was developed in this study, administration of this questionnaire to the same population at various time points may yield practice effects. Importantly, future research should seek to examine the questionnaires sensitivity to change following training and potentially develop equivalent questionnaires measuring the same or similar clinical decision-making competencies.

Overall, findings of the present study build on the developing literature on the importance of clinical decision making in psychological practice and provides preliminary support for the psychometric properties of the CDMQ-A. The development of tools that are brief, easy to administer, score and require minimal staff input have the potential to have significant implications for the dissemination and implementation of evidence-based practice, particularly within busy educational settings. Future research efforts should be made to continue to refine such tools, including with other clinical presentations, and evaluate its use in various settings and diverse samples.

### References

- Barber, J. P., Liese, B. S., & Abrams, M. J. (2003). Development of the cognitive therapy adherence and competence scale. *Psychotherapy Research, 13*(2), 205-221.  
<https://doi.org/10.1093/ptr/kpg019>
- Barnfield, T. V., Mathieson, F. M., & Beaumont, G. R. (2007). Assessing the development of competence during postgraduate cognitive-behavioural therapy training. *Journal of Cognitive Psychotherapy, 21*(2), 140-147.  
<https://doi.org/10.1891/088983907780851586>
- Barrington, J. (2006). Cognitive behaviour therapy: Standards for training and clinical practice. *Behaviour Change, 23*(4), 227-238. <https://doi.org/10.1375/bech.23.4.227>
- Beidas, R. S., Mehta, T., Atkins, M., Solomon, B., & Merz, J. (2013). Dissemination and implementation science: Research models and methods. In J. S. Comer & P. C. Kendall (Eds.), *The Oxford Handbook of Research Strategies for Clinical Psychology* (p. 62–86). Oxford University Press.  
<https://doi.org/10.1093/oxfordhb/9780199793549.001.0001>
- Bennett-Levy, J., & Beedie, A. (2007). The ups and downs of cognitive therapy training: What happens to trainees' perception of their competence during a cognitive therapy training course? *Behavioural and Cognitive Psychotherapy, 35*, 61-75.  
<https://doi.org/10.1017/s1352465806003110>
- Berry, A. C., Rosenfield, D., & Smits, J. A. J. (2009). Extinction retention predicts improvement in social anxiety symptoms following exposure therapy. *Depression and Anxiety, 26*(1), 22–27. <https://doi.org/10.1002/da.20511>
- Blackburn, I. M., James, I. A., Milne, D. L., Baker, C., Standart, S., Garland, A., & Reichelt, F. K. (2001). The revised cognitive therapy scale (CTS-R): Psychometric properties.

*Behavioural and Cognitive Psychotherapy*, 29(4), 431-446.

<https://doi.org/10.1017/s1352465801004040>

Broman-Fulks, J. J., Berman, M. E., Rabian, B. A., & Webster, M. J. (2004). Effects of aerobic exercise on anxiety sensitivity. *Behaviour Research and Therapy*, 42(2), 125-136. [https://doi.org/10.1016/s0005-7967\(03\)00103-7](https://doi.org/10.1016/s0005-7967(03)00103-7)

Brosan, L., Reynolds, S., & Moore, R. G. (2008). Self-evaluation of cognitive therapy performance: Do therapists know how competent they are? *Behavioural and Cognitive Psychotherapy*, 36(5), 581-587.

<https://doi.org/10.1017/s1352465808004438>

Carpenter, A. L., Pincus, D. B., Conklin, P. H., Wyszynski, C. M., Chu, B. C., & Comer, J. S. (2016). Assessing cognitive-behavioural clinical decision-making among trainees in the treatment of childhood anxiety. *Training and Education in Professional Psychology*, 10(2), 109-116. <https://doi.org/10.1037/tep0000111>

Clark, D. M. (2011). Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: The IAPT experience. *International Review of Psychiatry*, 23(4), 318-327. <https://doi.org/10.3109/09540261.2011.606803>

Cook, J. M., Biyanova, T., Elhai, J., Schnurr, P. P., & Coyne, J. C. (2010). What do psychotherapists really do in practice? An Internet study of over 2,000 practitioners. *Psychotherapy: Theory, Research, Practice, Training*, 47(2), 260-267.

<https://doi.org/10.1037/a0019788>

Davidson, K., Scott, J., Schmidt, U., Tata, P., Thornton, S., & Tyrer, P. (2004). Therapist competence and clinical outcome in the Prevention of Parasuicide by Manual Assisted Cognitive Behaviour Therapy Trial: the POPMACT study. *Psychological Medicine*, 34(5), 855-863. <https://doi.org/10.1017/s0033291703001855>

- Dawes, R. (1994). *Psychotherapy: The myth of expertise*. Free Press
- Deacon, B., Lickel, J. J., Possis, E. A., Abramowitz, J. S., Mahaffey, B., & Wolitzky-Taylor, K. (2012). Do cognitive reappraisal and diaphragmatic breathing augment interoceptive exposure for anxiety sensitivity? *Journal of Cognitive Psychotherapy*, 26(3), 257-269. <https://doi.org/10.1891/0889-8391.26.3.257>
- Ehlers, A., Gene-Cos, N., & Perrin, S. (2009). Low recognition of post-traumatic stress disorder in primary care. *London Journal of Primary Care*, 2(1), 36-42. <https://doi.org/10.1080/17571472.2009.11493240>
- Epstein, R. M. (2007). Assessment in medical education. *New England Journal of Medicine*, 356(4), 387-396. <https://doi.org/10.1056/NEJMra054784>
- Fairburn, C. G., & Cooper, Z. (2011). Therapist competence, therapy quality, and therapist training. *Behaviour Research and Therapy*, 49(6-7), 373-378. <https://doi.org/10.1016/j.brat.2011.03.005>
- Goisman, R. M., Warshaw, M. G., & Keller, M. B. (1999). Psychosocial treatment prescriptions for generalised anxiety disorder, panic disorder, and social phobia, 1991–1996. *American Journal of Psychiatry*, 156(11), 1819-1821.
- Hipol, L. J., & Deacon, B. (2013). Dissemination of evidence-based practices for anxiety disorders in Wyoming: A survey of practicing psychotherapists. *Behaviour Modification*, 37(2), 170-188. <https://doi.org/10.1177/0145445512458794>
- Hunsley, J., Elliott, K., & Therrien, Z. (2014). The efficacy and effectiveness of psychological treatments for mood, anxiety, and related disorders. *Canadian Psychology/Psychologie Canadienne*, 55(3), 161. <https://doi.org/10.1037/a0036933>
- Hung, E. K., Fordwood, S. R., & Cramer, R. J. (2012). A method for evaluating competency in assessment and management of suicide risk. *Academic Psychiatry*, 36, 23–28. [doi:10.1176/appi.ap.10110160](https://doi.org/10.1176/appi.ap.10110160).

- Huppert, J. D., Bufka, L. F., Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2001). Therapists, therapist variables, and cognitive-behavioural therapy outcome in a multicenter trial for panic disorder. *Journal of Consulting and Clinical Psychology, 69*(5), 747. <https://doi.org/10.1037/0022-006X.69.5.747>
- Kaslow, N. J., Grus, C. L., Campbell, L. F., Fouad, N. A., Hatcher, R. L., & Rodolfa, E. R. (2009). Competency assessment toolkit for professional psychology. *Training and Education in Professional Psychology, 3*(4), S27–S45. <https://doi.org/10.1037/a0015833>
- Kazantzis, N., & Munro, M. (2011). The emphasis on cognitive-behavioural therapy within clinical psychology training at Australian and New Zealand universities: A survey of program directors. *Australian Psychologist, 46*(1), 49-54. <https://doi.org/10.1111/j.1742-9544.2010.00011.x>
- Keen, A. J., & Freeston, M. H. (2008). Assessing competence in cognitive-behavioural therapy. *The British Journal of Psychiatry, 193*(1), 60-64. <https://doi.org/10.1192/bjp.bp.107.038588>
- McHugh, R. K., & Barlow, D. H. (2010). The dissemination and implementation of evidence-based psychological treatments: A review of current efforts. *American Psychologist, 65*(2), 73. <https://doi.org/10.1037/a0018121>
- McLeod, B. D., Islam, N., & Wheat, E. (2013). Designing, conducting, and evaluating therapy process research. In J. S. Comer & P. C. Kendall (Eds.), *The Oxford Handbook of Research Strategies for Clinical Psychology* (p. 142–164). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199793549.013.0009>
- McLeod, B. D., Smith, M. M., Southam-Gerow, M. A., Weisz, J. R., & Kendall, P. C. (2015). Measuring treatment differentiation for implementation research: The Therapy

- Process Observational Coding System for Child Psychotherapy Revised Strategies Scale. *Psychological Assessment*, 27(1), 314–325. <https://doi.org/10.1037/pas0000037>
- McManus, F., Rakovshik, S., Kennerley, H., Fennell, M., & Westbrook, D. (2012). An investigation of the accuracy of therapists' self-assessment of cognitive-behaviour therapy skills. *British Journal of Clinical Psychology*, 51(3), 292-306. <https://doi.org/10.1111/j.2044-8260.2011.02028.x>
- McManus, F., Westbrook, D., Vazquez-Montes, M., Fennell, M., & Kennerley, H. (2010). An evaluation of the effectiveness of diploma-level training in cognitive behaviour therapy. *Behaviour Research and Therapy*, 48(11), 1123-1132. <https://doi.org/10.1016/j.brat.2010.08.002>
- Muse, K., & McManus, F. (2013). A systematic review of methods for assessing competence in cognitive-behavioural therapy. *Clinical Psychology Review*, 33(3), 484-499. <https://doi.org/10.1016/j.cpr.2013.01.010>
- Mussell, M. P., Crosby, R. D., Crow, S. J., Knopke, A. J., Peterson, C. B., Wonderlich, S. A., & Mitchell, J. E. (2000). Utilisation of empirically supported psychotherapy treatments for individuals with eating disorders: A survey of psychologists. *International Journal of Eating Disorders*, 27(2), 230-237. [https://doi.org/10.1002/\(SICI\)1098-108X\(200003\)27:2<230::AID-EAT11>3.0.CO;2-0](https://doi.org/10.1002/(SICI)1098-108X(200003)27:2<230::AID-EAT11>3.0.CO;2-0)
- Myles, P., & Milne, D. (2004). Outcome evaluation of a brief shared learning programme in cognitive behavioural therapy. *Behavioural and Cognitive Psychotherapy*, 32(2), 177-188. <https://doi.org/10.1017/s1352465804001183>
- National Institute for Health and Care Excellence. (2011). Common Mental Health Disorders: The NICE Guideline on Identification and Pathways to Care. National Clinical

Guideline Number 123. Retrieved April, 2021, from,  
<https://www.nice.org.uk/guidance/cg123>.

National Research Council and Institute of Medicine. (2009). *Depression in parents, parenting, and children: Opportunities to improve identification, treatment, and prevention*. National Academies Press. <https://doi.org/10.17226/12565>

Nevo, B. (1985). Face Validity Revisited. *Journal of Educational Measurement*, 22(4), 287–293. <https://doi.org/10.1111/j.1745-3984.1985.tb01065.x>

Norton, P. J., & Price, E. C. (2007). A meta-analytic review of adult cognitive-behavioral treatment outcome across the anxiety disorders. *The Journal of Nervous and Mental Disease*, 195(6), 521-531. <https://doi.org/10.1097/01.nmd.0000253843.70149.9a>

Olatunji, B. O., Cisler, J. M., & Deacon, B. J. (2010). Efficacy of cognitive behavioral therapy for anxiety disorders: A review of meta-analytic findings. *Psychiatric Clinics*, 33(3), 557-577. <https://doi.org/10.1016/j.psc.2010.04.002>

Pachana, N. A., Sofronoff, K., Scott, T., & Helmes, E. (2011). Attainment of competencies in clinical psychology training: Ways forward in the Australian context. *Australian Psychologist*, 46(2), 67–76. <https://doi.org/10.1111/j.1742-9544.2011.00029.x>

Patel, V. L., Kaufman, D. R., & Arocha, J. F. (2002). Emerging paradigms of cognition in medical decision-making. *Journal of Biomedical Informatics*, 35(1), 52-75. [https://doi.org/10.1016/S1532-0464\(02\)00009-6](https://doi.org/10.1016/S1532-0464(02)00009-6)

Roberts, R., & Norris, K. (2020). Using objective structured clinical examinations for selection into and progress through postgraduate training. In G. J. Rich, A. P. Lopez, L. Ebersohn, J. Taylor, & S. Morrissey (Eds.), *Teaching psychology around the world* (Vol. 5, pp. 242–252). Cambridge Scholars Publishing.

- Ruscio, J., & Stern, A. R. (2006). The consistency and accuracy of holistic judgment: Clinical decision making with a minimally complex task. *Scientific Review of Mental Health Practice, 4*(2), 52-65.
- Schmidt, N. B., Woolaway-Bickel, K., Trakowski, J., Santiago, H., Storey, J., Koselka, M., & Cook, J. (2000). Dismantling cognitive-behavioural treatment for panic disorder: Questioning the utility of breathing retraining. *Journal of Consulting and Clinical Psychology, 68*(3), 417-424. <https://doi.org/10.1037/0022-006x.68.3.417>
- Shafran, R., Clark, D., Fairburn, C., Arntz, A., Barlow, D., Ehlers, A., Freeston, M., Garety, P., Hollon, S., Ost, L., Salkovskis, P., Williams, J., & Wilson, G. (2009). Mind the gap: Improving the dissemination of CBT. *Behaviour Research and Therapy, 47*(11), 902-909. <https://doi.org/10.1016/j.brat.2009.07.003>
- Sheen et al. 2015. Assessing the clinical competence of psychology students through Objective Structured Clinical Examinations (OSCEs): Student and staff views. *Australian Psychologist, 50*(1), 51-59. <https://doi.org/10.1111/ap.12086>
- Sholomskas, D. E., Syracuse-Siewert, G., Rounsaville, B. J., Ball, S. A., Nuro, K. F., Carroll, K. M. (2005). We don't train in vain: A dissemination trial of three strategies of training clinicians in cognitive-behavioural therapy. *Journal of Consulting and Clinical Psychology, 73*(1), 106-115. <https://doi.org/10.1037/0022-006X.73.1.106>
- Simpson, H. B., Maher, M., Page, J. R., Gibbons, C. J., Franklin, M. E., & Foa, E. B. J. B. t. (2010). Development of a patient adherence scale for exposure and response prevention therapy. *41*(1), 30-37. <https://doi.org/10.1016/j.beth.2008.12.002>.
- Šimundić, A. M. (2009). Measures of diagnostic accuracy: Basic definitions. *Ejifcc, 19*(4), 203.



- Smits, J. A., Berry, A. C., Tart, C. D., & Powers, M. B. (2008). The efficacy of cognitive-behavioral interventions for reducing anxiety sensitivity: A meta-analytic review. *Behaviour Research and Therapy, 46*(9), 1047-1054.  
<https://doi.org/10.1016/j.brat.2008.06.010>
- Spengler, P. M., White, M. J., Ægisdóttir, S., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G. R., & Rush, J. D. (2009). The meta-analysis of clinical judgment project: Effects of experience on judgment accuracy. *The Counseling Psychologist, 37*(3), 350–399.  
<https://doi.org/10.1177/0011000006295149>
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder. *Archives of Internal Medicine, 166*(10), 1092. <https://doi.org/10.1001/archinte.166.10.1092>
- Stewart, R. E., & Chambless, D. L. (2007). Does psychotherapy research inform treatment decisions in private practice? *Journal of Clinical Psychology, 63*(3), 267-281.  
<https://doi.org/10.1002/jclp.20347>
- Stewart, R. E., & Chambless, D. L. (2009). Cognitive-behavioural therapy for adult anxiety disorders in clinical practice: A meta-analysis of effectiveness studies. *Journal of Consulting and Clinical Psychology, 77*(4), 595. <https://doi.org/10.1037/a0016032>
- Tudiver, F., Rose, D., Banks, B., & Pfortmiller, D. (2009). Reliability and validity testing of an evidence-based medicine OSCE station. *Fam Med, 41*(2), 89-91.
- Wang, P. S., Lane, M., Olfson, M., Pincus, H. A., Wells, K. B., & Kessler, R. C. (2005). Twelve-month use of mental health services in the United States: Results from the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 629-640. <https://doi.org/10.1001/archpsyc.62.6.629>

- Wass, V., Jones, R., & Van der Vleuten, C. (2001). Standardized or real patients to test clinical competence? The long case revisited. *Medical Education*, *35*(4), 321-325. <https://doi.org/10.1046/j.1365-2923.2001.00928.x>
- Wedding, D. (1991). Clinical judgment in forensic neuropsychology: A comment on the risks of claiming more than can be delivered. *Neuropsychology Review*, *2*(3), 233-239. <https://doi.org/10.1007/BF01109046>
- Weissman, M. M., Verdeli, H., Gameroff, M. J., Bledsoe, S. E., Betts, K., Mufson, L., Fitterling, H., & Wickramaratne, P. (2006). National Survey of Psychotherapy Training in Psychiatry, Psychology, and Social Work. *Archives of General Psychiatry*, *63*(8), 925. <https://doi.org/10.1001/archpsyc.63.8.925>
- Wootton, B. M., Bragdon, L. B., Steinman, S. A., & Tolin, D. F. (2015). Three-year outcomes of adults with anxiety and related disorders following cognitive-behavioral therapy in a non-research clinical setting. *Journal of Anxiety Disorders*, *31*, 28-31. <https://doi.org/10.1016/j.janxdis.2015.01.007>
- Yap, K., Bearman, M., Thomas, N., & Hay, M. (2012). Clinical psychology students experience of a pilot objective structured clinical examination. *Australian Psychologist*, *47*, 165–173. doi:10.1111/j.1742-9544.2012.00078.x.
- Young, A. S., Klap, R., Sherbourne, C. D., & Wells, K. B. (2001). The quality of care for depressive and anxiety disorders in the United States. *Archives of General Psychiatry*, *58*(1), 55-61. <https://doi.org/10.1001/archpsyc.58.1.55>

Table 1.

*Participant Characteristics for Expert Sample (N = 7) and Trainee Sample (N = 104)*

	Expert sample		Trainee sample	
Age (M, <i>SD</i> )	42.1 (5.6)		30.8 (8.3)	
Gender (Female) (n, %)	4	57.1%	86	82.7%
Highest level of completed training (n, %)				
PhD	5	71.4%	-	-
Masters	2	28.6%	-	-
Bachelor's degree	-	-	104	100%
Course enrolled in (n, %)				
PhD or PhD / Masters combined	-	-	32	30.8%
Master of Psychology (Clinical)	-	-	62	59.6%
Master of Professional Psychology	-	-	10	9.6%
Years of experience (M, <i>SD</i> )	12 (5.4)		-	