



Applied knowledge of Acceptance and Commitment Therapy (ACT): Developing and assessing the utility of a Situational Judgement Test (SJT)

Kathryn Jamison^{a,*}, David Curran^a, Ross White^a, Victoria Samuel^b

^a School of Psychology, Queen's University Belfast, David Keir Building, Malone Road, Northern Ireland, BT9 5BP, United Kingdom

^b School of Psychology, Cardiff University, Tower Building, 70 Park Place, Cardiff, CF10 3AT, United Kingdom

ARTICLE INFO

Keywords:

Acceptance and commitment therapy
Situational judgement test
ACT applied knowledge
ACT questionnaire
ACT training

ABSTRACT

Acceptance and Commitment Therapy (ACT) is an innovative form of cognitive behavioural therapy that aims to increase psychological flexibility. Currently, challenges exist with measuring practitioners' application and understanding of ACT. A situational judgement test (SJT) offers a novel approach to assessing a practitioner's ACT consistent knowledge and how it can be applied in practice. In the current research, two consecutive studies were completed to develop and evaluate the utility of the first SJT for assessing clinicians' applied ACT knowledge. First, expert consensus via three iterative rounds of Delphi methodology was used to develop the 10-item ACT SJT: 13 panellists participated in round one, 12 in round two, and 10 in round three. Involved experts specialised in the application of ACT in clinical practice and/or research for at least five years. A second study examined the utility of the developed ACT SJT, with significant pre-post changes in ACT SJT scores following ACT training. The development of the ACT SJT has potentially important implications for both clinical and research settings. It can be used by clinicians to self-evaluate their clinical application of ACT, by trainers evaluating the effectiveness of ACT introductory training, and by researchers investigating the application of ACT-consistent knowledge.

1. Introduction

Acceptance and Commitment Therapy (ACT; Hayes et al., 2011) is a "third wave" cognitive behavioural approach, which incorporates a range of strategies to enhance psychological flexibility (Hayes et al., 2004). Psychological flexibility refers to the acceptance of thoughts and emotions with minimal entanglement or judgement, and commitment to actions based on individual values (Boone et al., 2015). ACT advocates psychological flexibility via core principles that can be expressed within the "Triflex" model, which comprises of *openness* (i.e., cognitive defusion and acceptance of thoughts and feelings), *awareness* (i.e., being present and mindful), and *engagement* (i.e., doing what matters based on values and committed action) (Harris, 2009). Practitioners are trained in the use of ACT techniques, such as mindfulness, perspective-taking, and values elicitation - that aim to improve psychological flexibility, and to subsequently enhance wellbeing and encourage behaviour based on chosen values (Luoma & Vilardaga, 2013).

Theoretically, ACT is informed by a behavioural model of language and cognition called Relational Frame Theory (RFT; Hayes & Brownstein, 1985). RFT helps to explain why Triflex-related strategies such as mindful acceptance, defusion and values clarification can be effective in

helping people to cope better with life challenges (Barnes-Holmes et al., 2004). ACT has also been influenced by a pragmatic philosophy known as Functional Contextualism (Hayes et al., 2012), which explores the function of behaviours in particular contexts. Meta-analytic reviews have found that improved psychological flexibility is consistently associated with positive outcomes including improved mental health symptomology, enhanced subjective wellbeing, increased job satisfaction and performance, and increased quality of life (Gloster et al., 2020; Hayes et al., 2006; Ong et al., 2023). Systematic review outcomes have been promising for the application of ACT interventions within a range of settings, including with different mental health presentations (e.g., Bai et al., 2020; Swain et al., 2013), and with a variety of medical conditions (e.g., González-Fernández & Fernández-Rodríguez, 2019; Sakamoto et al., 2021). There is evidence for the broad applicability of ACT as a flexible, transdiagnostic treatment approach (Dindo et al., 2017).

A key objective of psychotherapy training is to enhance practitioner fidelity in delivering that therapy, as this is linked with improved client outcomes (Rakovshik & McManus, 2010). Treatment fidelity refers to the degree to which an intervention was implemented in line with the protocol or design (Proctor et al., 2011), while integrity represents part

* Corresponding author.

E-mail addresses: kjamison06@qub.ac.uk (K. Jamison), d.curran@qub.ac.uk (D. Curran), r.white@qub.ac.uk (R. White), SamuelV3@cardiff.ac.uk (V. Samuel).

of fidelity, focusing on the degree to which a treatment is implemented as intended (Borrelli, 2011). Competency refers to the general skills required to facilitate interventions (Kohrt et al., 2015). When practitioners are trained to deliver a therapeutic approach (e.g., ACT or Cognitive Behavioural Therapy [CBT]), it cannot be assumed that they will practice in line with their training, as it is common to 'drift' into other approaches (Waller & Turner, 2016). Treatment fidelity can be measured in research trials to ensure that the therapy delivered is in line with training and the treatment under study, thereby strengthening the quality of the conduct of trials (Akiba et al., 2022). Historically, there has been limited evaluation of treatment adherence within ACT research trials, which is a significant methodological concern (Graham et al., 2016; O'Neill et al., 2019). The development of the *ACT Fidelity Measure* (ACT-FM; O'Neill et al., 2019) has sought to remedy this. It requires an expert in ACT to listen to recorded sessions of the treatment delivered in the trial and then rate therapist fidelity using the ACT-FM questionnaire. In trials, however, ACT fidelity is often assessed retrospectively when the treatment has been completed. It may be helpful to notice if there are gains in practitioners' understanding of ACT-related principles and practices (henceforth referred to as *ACT knowledge*) immediately after training, so that additional training and appropriate supervision can be offered to practitioners, if required.

Currently, the effectiveness of ACT training can be measured via the *ACT Knowledge Questionnaire* (AKQ; Luoma & Vilardaga, 2013), which assesses knowledge about ACT-related concepts and processes. Previous research has found that the AKQ can be sensitive to knowledge gain following ACT training (Luoma & Vilardaga, 2013). However, AKQ may offer limited insight into the practical knowledge of ACT because it measures conceptual knowledge, ACT principles and definitions (such as psychological flexibility) and does not capture *how* practitioners might apply such concepts in response to clinical encounters. In addition, due to the high level of ACT-related terminology used within the AKQ, it's unlikely that newcomers to ACT would be sufficiently able to respond to all questionnaire items. To further assess the practical application of any ACT knowledge, behavioural intentions or gains from training, further psychometric measures are required. To help understand the impact of ACT training on applied knowledge of ACT for practitioners who are new to the ACT approach, we therefore developed an *ACT Situational Judgement Test* (SJT).

SJTs measure knowledge and behavioural intentions by analysing responses to hypothetical case-based scenarios (Wolcott et al., 2020). Typically, a scenario will outline a problem or dilemma, with possible answers assessing the knowledge and other related skills of respondents (Olaru et al., 2019). SJTs explore and evoke current behaviour which can be used as an indicator of future behavioural responses (Whetzel & McDaniel, 2009). It has been suggested that SJTs represent a low-cost opportunity to evoke realistic responses to scenarios which closely portray possible real-world situations (Murase et al., 2019). While the concept of SJTs might be a novel approach within an ACT framework and have not been used in other therapeutic modalities at this stage, SJTs have been extensively used and evaluated within a range of other settings. Specifically, SJTs are commonly used to explore competence, knowledge, and skills within workplace settings, such as with candidate recruitment or when assessing job performance and professional competencies (Persich & Robinson, 2020; Wolcott et al., 2020). The validity and potential impact of SJTs has been demonstrated within previous studies, with a recent systematic review and meta-analysis highlighting the moderate level of incremental predictive validity of SJTs in the selection of candidates for medical training (Webster et al., 2020).

SJTs have been proposed as important for assessing training needs, as they can highlight training deficits for individuals who respond in a manner that does not demonstrate a cognitive awareness of the link between relevant situations and the target behaviours within SJTs (Murase et al., 2019). Thus, there is potential for the use of SJTs in the evaluation of therapeutic training. An ACT SJT could be a helpful assessment of peoples' response to ACT training or benchmarking for

training needs, which could enhance the quality of ACT delivered in routine clinical practice and research trials. It could also provide information regarding treatment integrity. Fidelity to treatment models has been emphasised within implementation research, to ensure outcomes can be linked with specific processes in the treatment model, and to help design future intervention trials (Bellg et al., 2004; Bhattacharyya et al., 2009; Plumb & Vilardaga, 2010). In addition, there is potential utility of SJTs in under researched areas, including assessing the relationship between ACT practitioner characteristics and client outcomes. As such, we set out to undertake a programme of research across two studies aimed at developing the first ACT SJT (Study 1) and conducting a preliminary evaluation of its utility (Study 2).

Study 1 aimed to utilise experts specialising in ACT (both in theory and applied practice) via Delphi methodology to reach consensus on the content, refinement, and inclusion/exclusion of items for the ACT SJT, with a focus on assessing ACT knowledge at a beginner level. Delphi methodology has been previously used to achieve group consensus around item inclusion within ACT-related measures, followed by validation studies (Francis et al., 2016), including with the ACT-FM (O'Neill et al., 2019).

Following the development of the ACT SJT, Study 2 was conducted to evaluate the utility (i.e., usefulness and practical value) of the ACT SJT via pre- and post-training administration, to assess the acceptability and sensitivity of the ACT SJT. Participants were attendees at introductory ACT training courses (i.e., professionals who are naive to, or relatively new to, ACT).

Study 1 Delphi study to develop the Acceptance and Commitment Therapy Situational Judgement Test (ACT SJT)

2. Methodology

2.1. Aim

Study 1 aimed to develop and refine the ACT SJT through Delphi methodology.

2.2. Design

Although there is no specifically agreed protocol on the application of a Delphi Methodology (Fink-Hafner et al., 2019), guidance was followed to promote adherence to the methodology, including anonymity for panellists to reduce social desirability and influence responding, controlled feedback between survey rounds and iterative discussions (as guided by Nasa et al., 2021). A total of three rounds were completed within the Delphi study. Ethical approval was obtained from Queen's University Belfast School of Psychology Ethics Committee (study number EPS 22_346) on the November 15, 2022.

2.3. Participants

The involvement of experts is a core component of Delphi methodology (Hsu & Sandford, 2007). Consequently, only health professionals who have consistently applied ACT for at least five years within a clinical or research role were eligible to participate. Purposeful recruitment was used initially to recruit people into round one, with contact made via email by the research team to established international ACT experts. Subsequently, an advert was also posted on social media, which included a Qualtrics link for ACT experts to register their interest in the study and to confirm they met the eligibility criteria for participation. Informed consent was obtained prior to beginning the study.

There were 12 full responses and one partial response in round one, 12 full responses in round two, and 10 full responses in round three. Demographics of participants are shown in Table 1. Gender identity options were available via a drop-down menu, including 'female', 'male', 'non-binary', 'transgender' and 'other', with an open-text choice

Table 1
Demographics of participants across all rounds of the Delphi study.

Characteristic	Round One	Round Two	Round Three
	n	n	n
Gender			
Female	4	4	3
Male	9	8	7
Ethnicity Gender			
White	12	11	9
Mixed/multiple ethnic groups	1	1	1
Profession Gender			
Mental Health Nurse	1	1	1
Clinical/Educational/ Occupational	10	10	8
Psychologist			
Psychotherapist	2	1	1
Role relevant to ACT Gender			
Practitioner	7	6	6
Both researcher and practitioner	4	4	2
Other	2	2	2
Country of Practice Gender			
UK and/or Ireland	8	7	5
Rest of Europe	2	2	2
North America	2	2	2
South America	1	1	1

Note. $n = 13$ in round one, $n = 12$ in round two, $n = 10$ in round three.

for participants to state their own preferred terminology to define their gender identity. On average, participants rated their current ACT knowledge as 6.08 (range: 5–7) on a scale from 1 (intermediate: some prior knowledge/experience) to 7 (expert: high level of previous knowledge/experience). The median years of ACT experience was 15 years (interquartile range [IQR] 9–18.5). Participants worked with a number of client groups, including adult mental health ($n = 4$) and adult mental and physical health ($n = 4$), child and adolescent mental health ($n = 1$), adult, child and adolescent mental health ($n = 1$), adult, child and adolescent mental and physical health ($n = 1$), adult, child and adolescent mental and physical health, paediatrics, psychosis and primary care populations ($n = 1$), child and adolescent mental health, and

parents and teachers ($n = 1$).

When reported, there was a range of expertise listed by participants, including concepts and interventions specifically relating to ACT and other contextual behavioural approaches ($n = 5$), behavioural medicine including chronic illness and/or chronic pain ($n = 2$), mental health and/or relational issues ($n = 2$), ACT applied to children and young people's contexts ($n = 2$), and physical health adjustment and loss ($n = 1$). Four participants had published one ACT-related study, and five participants had published 10 or more ACT-related studies. Six participants were recognised as Association of Contextual Behavioural Science (ACBS) Peer Reviewed ACT Trainers. A total of 11 participants provided ACT supervision within their role.

2.4. Measures

2.4.1. Initial draft of the ACT SJT

As members of the current research team have specialised in the application of ACT in clinical practice and research (including an ACBS Peer Reviewed ACT Trainer), an initial draft of suggested items and content of the SJT was developed using the ACT knowledge and expertise of the research team, alongside relevant literature (see Fig. 1 for example items in the initial draft of the ACT SJT, as presented to panellists in round one of the Delphi study). Guidance supports the use of a structured questionnaire in the first round of a Delphi study if there is relevant evidence and literature that is accessible for the topic area (Hsu & Sandford, 2007). The ACT SJT built upon the recently developed 'Congruent ACT responses with Young People' (CoACT-Y; Samuel et al., n.d.). This questionnaire uses SJT to explore ACT consistent interpersonal behaviour, aimed at individuals working with children and young people. The initial item pool of 15 items (that is, 15 scenarios each with three response options) for the ACT SJT was drafted with an attempt to provide adequate coverage of the Triflex (Harris, 2009). For each scenario provided, one response option was designed to be "ACT consistent", with the other two "ACT inconsistent" (instead intended to derive from counselling-orientated and CBT-orientated approaches).

Situational Judgement Test of ACT Congruent Responses

Someone you are working with says, "the things I think are true – I am 'a failure' and that won't change no matter what I try". Do you:

- ☐ A. Respond by saying that it sounds like their mind gives them a hard time telling them that they are a failure. Ask them to notice and describe the feelings that come up when the words go through their mind.
- ☐ B. Reflect that it sounds like they feel that the thought will always stay with them, which makes sense in terms of the difficult experiences they've had.
- ☐ C. Gather evidence for and against their belief that they are a failure with questions such as, would someone else describe them as a failure? Why not? What would be a more balanced way of looking at this?

Someone you are working with says: "I don't want to talk about that, it's just going to make me feel worse". Do you:

- ☐ A. Say that it sounds like their mind is telling them that talking will bring up feelings that are too difficult. Ask if they can notice the tone of those thoughts, or what it brings up in their body.
- ☐ B. Say that it's understandable they don't want to talk about it. Ask if they can think of a time when they have talked about difficult things and not felt worse.
- ☐ C. Say that you can see how difficult it is for them to talk about this. Ask can they remember the first time they felt scared to speak about their experiences.

Fig. 1. Example items in the initial draft of the ACT SJT

2.5. Procedure

Participants were sent a link to the draft ACT SJT, presented on Qualtrics (survey software), which was used to host each round of the Delphi study. In round one, participants provided demographic information to support the collation of descriptive statistics about the sample. A brief introduction to the ACT SJT was then provided, to orientate participants to the upcoming task. Participants were then presented with suggested items and asked to rate if items should be included in the ACT SJT, based on a 7-point Likert scale (in which 1 represented the lowest level of agreement rating, and 7 signified the highest point). Space was provided for further feedback on each individual scenario and suggested response options within the ACT SJT, and space to suggest new items, response options, or add general comments. Participants were invited to rounds two and three if they initially participated in rounds one and two, respectively. Between each round, responses were analysed by the research team to decide on exclusion and inclusion of items based on agreement by the Delphi panel. Delphi responses were collected by the lead researcher, who anonymised responses prior to sharing with the research team. No identifiable information was provided to the wider team about specific contributions in each round. Feedback was provided to participants via summary documents after each round that consisted of anonymised panel comments and agreement ratings for item inclusion, and the research team's response, including decisions for items retained, edited, or removed.

In round two, a second iteration (amended based on feedback from round one) of the ACT SJT was sent to participants in the Delphi panel, shaped by the level of item agreement and suggestions from the previous round. It was agreed that the order that items were presented would be changed for rounds two and three, to account for order effects. Again, participants saw suggested items to rate, alongside space to comment on each item.

In round three, a prototype draft ACT SJT with scoring instructions was sent to participants for any concluding comments and considerations. This allowed participants to view the questionnaire in draft format, provide closing feedback on items, layout, and usability of the ACT SJT, provide comments on the introductory and scoring instructions, and add any other relevant comments or reflections about participation in the study. Each round took participants approximately 20–60 min to complete, with several weeks between each round to allow for analysis, and refinement of the ACT SJT.

2.6. Analysis

For the Delphi study, the level of agreement between expert ratings was used to reach a consensus for item inclusion, while qualitative information provided via prompt items guided the item edits. This resulted in the adaptation of the ACT SJT for further rounds resulting in a prototype version of the ACT SJT by round three. As recommended by similar previous Delphi studies (Francis et al., 2016; O'Neill et al., 2019) and relevant guidelines (Diamond et al., 2014), there was an 80 % cut-off point for the agreement score (at '5, 6 or 7' on the Likert scale) between experts when ranking items to be included in the ACT SJT.

Expert comments were explored for common themes to inform modifications to each item. The team looked for suggestions that were echoed by at least two participants, for amendments to be made. Suggestions made by one panel member were discussed with reference to relevant ACT literature, and changes made if evidence and discussion supported it.

3. Delphi study results

3.1. Round one ratings

In round one, five items received an agreement rating below 80 % (i.e., less than nine participants rated these items between 5 and 7 on the

agreement Likert scale for item inclusion) and were removed prior to round two. Of the items that were removed, the lowest rated item had an average agreement rating of 4.31 ($SD = 2.18$), with six participants (46.2 %) rating this between 5 and 7 on the Likert scale for inclusion. Of the retained items, three received a 100 % agreement rating by the panel for inclusion in the ACT SJT (that is, all participants scored items between 5 and 7 on the agreement Likert scale for item inclusion). [Supplementary Table 1](#) shows ratings at round one, and decisions from the research team regarding retained and removed items.

3.2. Round one amendments

3.2.1. Purpose and aim of the ACT SJT

Participants commented on the purpose of the ACT SJT and suggestions for developing the ACT SJT into a measure of ACT fluency. Changes were made by the research team to ensure the purpose of the ACT SJT was reinforced and made clear to panel members, with a re-orientation to the function of the ACT SJT as an introductory questionnaire of ACT consistent behaviour included in the summary feedback to panellists.

3.2.2. ACT SJT and overlap with the Triflex model

There were comments regarding the overlap of the ACT SJT with the Triflex model and ensuring adequate coverage of this across all included items. It was agreed that the target ACT Triflex process would be noted alongside each item for round two, to allow the panel to see the specific process that was targeted and to facilitate closer refinement to this within round two.

3.2.3. Length and detail of items

Comments were provided by the panel regarding the mixed length and brief detail of some items. The research team agreed that, as this is intended to be a brief measure, items did not require a high level of detail. Instead, the focus was to fine-tune responses to ensure they were consistent with the suggested edits when appropriate, with consideration given to equal length of response options, to ensure limited bias towards any detailed responses.

3.2.4. Specific item changes

No additional items were suggested for inclusion in round two. The wording of an item scenario was changed based on participant feedback and the discussion by the research team ("I tried what you said" changed to "I did what you said"). With regards to ACT consistent responses for items, when advised by panel members, an additional focus of validation, curiosity and empathy was incorporated into responses. Further refinement was suggested for specific items, to ensure sufficient overlap with the targeted ACT processes. With regards to ACT inconsistent responses, panel comments focused on how to differentiate ACT target responses from other response options for each item. Vague or passive response options were edited when suggested. Two participants commented regarding the length and detail of items and to ensure that target responses were not noticeably longer than distractor response options.

3.3. Round two ratings

All ten items met the 80 % criteria and were retained for round three. Further details of ratings at round one is shown in [Supplementary Table 2](#).

3.4. Round two amendments

Key amendments focused on ensuring close alignment with targeted ACT concepts and better differentiation between ACT consistent and inconsistent response options. As identified from round one, balanced coverage across the Triflex model was achieved, with 'awareness' represented by items 2, 3 and 4, 'openness' targeted in items 1, 6 and 8, and

'engagement' focused on in items 5, 7 and 9. Item 10 covered both openness and engagement. No specific changes were suggested with regards to scenarios provided in round two, however an additional item was suggested for inclusion in round three, which was discussed by the research team, and it was agreed that although this item had potential, introducing a new item within the final round would not allow for the same level of scrutiny from the panel that other items within the questionnaire received.

3.5. Round three amendments

3.5.1. Item scenarios and response options

One participant agreed with suggested changes to an item based on feedback provided by the panel in previous rounds, with edits completed. Further refinement was made to an item to improve the consistency of ACT and ACT inconsistent response options. As suggested, the word 'upon' was added to an item scenario (becoming 'what would you say if someone you are working with did not complete a task which was agreed upon the last time you had met ...').

3.5.2. Introductory and scoring instructions

Changes were made to the introductory instructions, using the suggestions provided by two participants: to include the word 'evaluate' instead of 'explore', and to clarify that practitioner's responses are the target of ACT SJT, not client's behaviour. Changes were also made to the scoring instructions, based on the comments of three participants.

3.5.3. The layout and usability of the ACT SJT

Eight comments provided positive feedback regarding the layout, including references to the clarity and accessibility of the questionnaire. Two participants commented on the layout of the ACT SJT, with more space added to the questionnaire, particularly with the instructions.

Study 2 Testing the utility of the Acceptance and Commitment Therapy Situational Judgement Test (ACT SJT)

4. Methodology

4.1. Aim

Study 2 sought to explore the utility of the ACT SJT that was developed in Study 1 (freely available on request from the lead author; kjamison06@qub.ac.uk). It was hypothesised that there would be significant changes in scores of the ACT SJT pre- and post-training administrations; that the ACT SJT would demonstrate acceptable internal consistency; and that participants would provide positive feedback about the clarity of the ACT SJT.

4.2. Design

This was a longitudinal, within-groups design, as the SJT was completed by the same participants (non-clinical sample) both before and after ACT training.

4.3. Procedure

Convenience sampling was used to recruit attendees at Contextual Consulting ACT training courses. Contextual Consulting (<https://contextualconsulting.co.uk>) is a training provider specialising in ACT workshops, provided by specialists in the field. Four introductory-level workshops were targeted for recruitment. Two live workshops were delivered in March 2024, including the focused ACT foundations workshop. This covered ACT for brief interventions, with introductory material completed within the initial 6 h. The ACT with parents and teachers workshop lasted 2 h, and focused on the DNA-V (discoverer, noticer, advisor, values) approach. ACT for insomnia and sleep

difficulties was a live 6-h workshop facilitated in April 2024, that covered sleep, insomnia and the relationship to the ACT model. An introduction to the ACT model was a 1-h pre-recorded workshop that could be accessed on demand and focused on key aspects of the ACT model.

A recruitment advert was sent by Contextual Consulting to participants attending live ACT workshops, with a study link emailed to participants if they signed up to the pre-recorded ACT workshop. A recruitment advert was posted on social media to aid recruitment for the pre-recorded workshop. After consenting to participate, a Qualtrics link was sent to complete the ACT SJT and provide demographic information, prior to ACT workshops. After completing the training, all participants were invited to complete the ACT SJT and an acceptability questionnaire. For the pre-recorded workshop, participants confirmed they had completed the training prior to proceeding with post-training tasks.

4.4. Measures

4.4.1. The developed ACT SJT

The final version of the ACT SJT consisted of introductory instructions, followed by ten items (10 scenarios and three response options per item). Scoring instructions are provided, which advise that ACT SJT items are scored as '1' if the respondent provides the ACT-consistent approach and '0' for all other responses. A total score is calculated by totalling the scores for each item. Higher overall scores indicate that ACT concepts may exert more influence in how clinicians respond to people, while lower overall scores suggest that other therapeutic modalities may be more influential.

4.4.2. Acceptability questionnaire

A four-item acceptability questionnaire was developed and administered alongside the ACT SJT. This asked participants to rate on a scale from 1 (not at all) to 7 (extremely) how easy it was to understand the questionnaire (item 1) and differentiate between response options (item 2), and to provide qualitative information about any particularly difficult to understand scenarios or response options (item 3), and potential improvements to the ACT SJT (item 4). This measure was only administered post-training.

4.5. Participants

A sample size of 34 was calculated via G*power repeated-measures *t*-test analysis, based on a desired power level of 0.80, alpha of 0.05 and a medium effect size ($d = 0.50$). This was comparable to other studies testing SJTs pre- and post-training (e.g., [Peckler et al., 2012](#)). A total of 52 participants completed pre- and post-training questionnaires, with 6 more completing the pre-questionnaire. Demographics of participants are shown in [Table 2](#). There was a median of 5 years ($M = 9.2$, $SD = 11.2$) experience in their profession (ranging from under 1 year to 45 years). In terms of ACT knowledge, participants on average rated themselves as 2.02 ($SD = 1.61$; range: 0 to 6) on a scale from 0 (novice) to 7 (expert). A total of 24 out of 58 participants indicated that they had previous ACT training (41.4 %). Gender identity information was captured via an open-text box for participants to state their preferred terminology.

5. Study 2 results

Paired *t*-tests were used to assess SJT scores pre- and post-ACT training. Internal consistency was assessed via Cronbach's alpha of the SJT scores (pre- and post-ACT training). Likert responses about the acceptability of the ACT SJT were subject to descriptive statistics including frequency counts. There was a statistically significant difference in mean ACT SJT scores between pre-ACT training and post-ACT training, $t(51) = 4.82$, $p < .001$. The mean pre-training ACT SJT score

Table 2
Demographics of participants in study 2.

	Focused ACT foundations	ACT with parents and teachers	Introduction to the ACT model	ACT with insomnia and sleep difficulties
	n	n	n	n
Gender				
Female	1	5	39	5
Male	0	1	5	2
Profession				
Psychologist (assistant, trainee ... or qualified)	1	3	29	5
Coach	0	2	1	1
Psychotherapist	0	1	1	1
Mental health professional or ...counsellor	0	0	4	0
Medical, nursing, or allied ...health professional	0	0	5	0
Teacher	0	0	4	0
Country of Practice				
UK	1	4	40	1
Ireland	0	0	3	0
Rest of Europe	0	0	0	1
Australasia	0	2	1	2
USA	0	0	0	3
Retention				
Completed pre & post	1	5	41	5
Did not complete post	0	1	3	2

Note. $N = 58$ overall ($n = 1$ in 'focused ACT foundations', $n = 6$ in ACT with parents and teachers, $n = 44$ in an introduction to the ACT model, $n = 7$ in ACT for insomnia and sleep difficulties).

was 5.38 ($SD = 2.81$), with an increase to 7.29 ($SD = 1.98$) post-training. In addition, there was a strong statistically significant correlation between pre-training scores and previous ACT knowledge $r(58) = 0.63$, $p < .001$. Those with previous ACT training had higher pre-training ACT SJT scores ($M = 7.63$; $SD = 2.3$) than did those with no prior ACT training ($M = 4.09$; $SD = 2.17$). A mixed-methods ANOVA was performed to evaluate the effects of previous ACT training on time (that is, scores obtained pre-and post-ACT training). The results indicated a significant main effect for previous ACT training, $F(1, 56) = 36.27$, $p < 0.001$, $\eta_p^2 = .393$, and a significant main effect of time, $F(1, 56) = 28.45$, $p < 0.001$, $\eta_p^2 = .337$, with a non-significant interaction between previous ACT training and time, $F(1, 56) = 6.26$, $p = .015$, $\eta_p^2 = .101$. Cronbach's alphas demonstrated acceptable to good reliability for the ten-item ACT SJT at pre-training ($\alpha = .787$) and post-training ($\alpha = .619$). Typically, SJTs do not have strong internal consistencies, potentially due to the heterogeneity between scenarios (Gregory et al., 2017; Ployhart & Ehrhart, 2003; Prewett et al., 2013).

Responses biases were explored for each item, with two items correctly identified by the majority of participants pre-ACT training: item 4 (45 out of 58 participants; 77.59 %) and item 9 (41 out of 58 participants; 70.69 %). For those who had completed previous ACT training, a response bias was clear for item 4 (23 out of 24 participants; 96 %) and present for those who had not completed prior ACT training (22 out of 34; 65 %). This was the highest correctly scored item in both groups. Similarly, the ACT response in item 9 was correctly identified by most participants who had attended previous ACT training (21 out of 24; 88 %) and by over half who had not attended previous ACT training (20 out of 34; 59 %). Participants who had prior ACT training scored highly for correctly identifying item 10 (21 out of 24; 88 %), however this response pattern was not replicated by participants who had not previously attended ACT training (14 out of 34; 41 %). Similarly, item 2 was correctly identified by most participants who had previously attended

ACT training (20 out of 24; 83 %), however this was lower amongst participants who had no prior ACT training (16 out of 34; 47 %).

5.1. Acceptability feedback

After completing the post-training ACT SJT, participants were asked to rate two questions on a Likert scale from 1 (not at all easy) to 7 (extremely easy) regarding the acceptability of the ACT SJT. Average ratings are shown in Table 3.

Qualitatively, participants were asked if there were any scenarios or response options that were particularly difficult to understand. Most respondents (42 out of 52; 80.77 %) said no or that it was easy to understand. One participant stated that response options would depend on the context of the client they were working with. Three participants said that it was hard to choose one answer due to little variation between options, while one participant said that the questionnaire was easy to understand, however selecting a response option was challenging due to limited ACT knowledge. Another commented that while none of the items were difficult to understand, not all response options precisely reflected their preferential response. Specific items were commented on. For example, one participant said that the phrase "working with someone" could refer to a colleague and a work task. Another said that one of the response options in the final item was difficult to understand.

When considering improvements to the questionnaire for future use, one participant suggested using a mixture of written and video-based scenarios, while another suggested that some wording could be made more concise. Two participants commented on the use of ACT language, and how this could influence responses for those who are familiar with ACT. Two participants also suggested using different items in pre-test and post-test questionnaires.

6. Discussion

The number of professionals practising ACT is growing across the globe, however there are current challenges in how to assess practitioners' applied knowledge of ACT. SJTs potentially represent a new way to measure therapeutic knowledge by exploring behavioural intentions when responding to realistic scenarios (Murase et al., 2019). A SJT could be used in several ways; to measure training gains, detect any skill deficits, or for self-evaluating applied knowledge. The current research undertook two consecutive studies that aimed to develop and evaluate the utility of a novel questionnaire that assesses ACT consistent knowledge. Within the Delphi study, a 10-item SJT was created via expert consensus that specifically focused on behavioural intentions relating to ACT; the ACT SJT. Preliminary data were collected via Study 2 that subsequently examined the utility (i.e., the usefulness and practical value) of the ACT SJT.

From the Delphi study, a total of ten items were selected for inclusion in the final version of the ACT SJT, with refinements made as necessary, and five items removed from the initial draft of the ACT SJT. Of the ten final items, a total of 23 response options were amended throughout the three rounds of the Delphi study, with seven response options remaining the same as the initial draft (that is, four CBT-orientated response options, and three counselling-orientated response options). From the initial draft to the final version of the ACT SJT, all ACT consistent response options were edited for each item.

With regards to Study 2, the internal consistency of the ACT SJT was acceptable to good. There was also a significant pre-post change in ACT

Table 3
Average acceptability ratings by participants.

Acceptability questions	Mean	SD
1. How easy was this questionnaire to understand?	6.1	0.82
2. How easy was it to tell the difference between each response option?	5.3	0.94

SJT scores following ACT training. Convergent validity is supported by the positive correlation between previous ACT knowledge and pre-training scores on the ACT SJT. This could also indicate support for the ACT SJT as a questionnaire best designed to assess the gains of ACT beginners, as also indicated by two respondents who suggested that ACT language could influence responses patterns, for those who are familiar with ACT. However, both prior ACT training and time (pre-versus post-training) had significant and large effects on scores, with the lack of significant interaction suggesting that the level of improvement in scores from pre-to post-training was similar for those who had previously accessed ACT training and those who had not.

Overall, there was little evidence of a ceiling effect, that is, questions correctly identified as ACT by participants at baseline were not noticeably high. However, there were exceptions to this, with two items correctly identified by the majority of participants pre-ACT training (item 4 and item 9). Acceptability ratings suggest that participants found the questionnaire acceptable and useable, with some suggestions for how to develop the ACT SJT going forward.

There are important clinical and research implications that derive from the development and assessment of the utility of the ACT SJT, including use by practitioners to self-assess their application of ACT, to evaluate the effectiveness of introductory-level ACT training (as demonstrated by the current study), or to detect any training needs. For example, the ACT SJT could be used to the identify individuals who, post-ACT training (and pre intervention commencement in research trials or in clinical work), have perhaps not fully understood the key ACT concepts. As a result, additional supervision or training could be offered for any identified knowledge gaps.

Going forward, the ACT SJT could also be a valuable resource in future research, particularly with exploring aspects that may impact on applied ACT knowledge. For example, dose-response, and if a certain amount of training is required to result in a sufficient level of ACT knowledge. Research is generally limited regarding the optimal amount of training needed for clinicians in evidence-based therapies, to obtain knowledge gains and ultimately impact on client changes (Frank et al., 2020). It is hoped that the ACT SJT offers an alternative method – to assess ACT congruent knowledge or gains from training – besides the existing main measures of ACT fidelity and ACT knowledge, namely the ACT-FM and AKQ.

Although beyond the scope of the current study, the relationship between practitioners' applied knowledge of ACT and psychological flexibility would benefit from further exploration, and the ACT SJT could be used to support with this. As such, the ACT SJT attempts to provide adequate coverage of the Triflex model, which aims to enhance psychological flexibility (Harris, 2009). This represented an important aspect of face validity, with three items targeting engagement (items 5, 7 and 9), three targeting awareness (items 2, 3 and 4), three targeting openness (items 1, 6 and 8), and one item linked to both openness and engagement (item 10). Further exploration of applied knowledge and psychological flexibility would be particularly relevant given the gap in literature regarding the real-world application of knowledge acquirement following ACT training, if this is linked with any changes in therapist psychological flexibility and any relationship with client treatment outcomes (Luoma & Vilardaga, 2013).

In addition, the ACT SJT could also be used to explore other factors that might influence applied ACT knowledge and training gains, including training format, exercises within training, personal practice, and supervision (Beidas & Kendall, 2010). Although not possible in the current study 2 due to a comparatively small sample size, future research could complete a factor analysis to determine whether the ACT SJT items load on to a single factor. In addition, further research could explore any association between ACT SJT scores and ACT fidelity scores, as this would provide useful information regarding validation (that is, if knowledge predicts fidelity). The assessment of fidelity is crucial within research trials, as it examines the therapy delivered and its consistency with training provided (Akiba et al., 2022). As the ACT SJT is aimed for

use with practitioners who are new to the ACT approach, it may be useful in the assessment of adherence and integrity associated with newly gained therapeutic knowledge post-training. In addition to developing a SJT for use with relative newcomers to ACT, there may be future benefits from developing a SJT for assessing ACT fluency with more experienced ACT practitioners.

There are several limitations within the current study that are important to acknowledge and consider for future research. Firstly, anonymised Delphi responses were shared with the wider research team and panellists; however, it is acknowledged that the lead researcher was able to identify comments by participants, which may have introduced unintended bias when reviewing feedback. No identifiable information was shared with the wider team by the lead researcher to link participants with any specific contributions. In study 2, we were unable to compare across professional groups and different ACT training sessions, due to the sample size. As such, it was not possible to explore if any group of professionals are more likely to endorse ACT congruent responses, even without training. Although demographic information was collected in study 2, including country of practice, details about participant's ethnicity was not obtained. Therefore, the diversity and representativeness of the sample is unknown. In addition, the sample within study 1 was quite homogeneous in terms of ethnicity (mostly white) and country of practice (mainly high-income), while participants in study 2 were also demographically homogeneous (mainly UK). As such, the applicability of the ACT SJT across cultural contexts remains unclear, and future research would benefit from including a more representative and diverse sample when testing the utility of the ACT SJT, cultural adaptations (including translations of the questionnaire) as well as assessing its applicability across different populations.

In addition, as with other SJTs (e.g., Prewett et al., 2013), there is the potential for faking and social desirability to impact on the response patterns of participants, and this was not specifically explored in study 2. A measure of social desirability (such as the Marlow-Crowne Social Desirability Scale; Crowne & Marlow, 1960) could be included in future research to detect any social approval biases in responses. Additionally, as with all SJTs, the responses needed to be interpreted cautiously as hypothetical scenarios lack the detailed contextual factors associated with real world events.

An additional issue that could be explored further is whether the ACT SJT targets different constructs, including response preference, behavioural intentions, and applied knowledge. Relatedly, there has been debate about the construct validity of SJTs, that is, what they truly assess (Patterson et al., 2016; Tiffin et al., 2019). Research has generally endorsed the hypothesis that SJTs assess knowledge about effective behaviour within specific scenarios, based on theoretical underpinnings and personal values (Patterson et al., 2016), with SJTs implemented and validated in a range of settings, targeting various concepts. The current study aimed to tap into applied knowledge of a novel construct via the ACT SJT; and, as far as we are aware, is the first tool to specifically use a SJT approach to target a therapeutic orientation. It is possible that SJTs targeting different therapeutic approaches could be developed and used to easily (representing a quick, low-cost option) assess treatment fidelity, by exploring if the therapy delivered is consistent with the training and treatment in research trials.

In conclusion, a ten-item questionnaire was developed via expert consensus in a Delphi study, with preliminary findings highlighting the promise of the ACT SJT. The development of the ACT SJT provides an important platform for future research that may help refine the questionnaire. The ACT SJT has potential utility for professionals across a range of settings, for self-learning, in research trials, and to detect any changes following ACT training, particularly for those new to the ACT modality. It is anticipated that the ACT SJT can continue to be refined, amended, and improved over time, based on further use within clinical practice, ACT training and research settings.

CRedit authorship contribution statement

Kathryn Jamison: Writing – original draft, Project administration, Methodology, Investigation, Formal analysis. **David Curran:** Writing – review & editing, Supervision. **Ross White:** Writing – review & editing, Supervision. **Victoria Samuel:** Writing – review & editing, Supervision.

Author note

Data is available upon reasonable request. The ACT SJT is freely available on request from the lead author; email kjamison06@qub.ac.uk.

Declaration of competing interest

The authors have no conflict of interest to declare.

Acknowledgements

We want to thank everyone who took part in either of our studies. Of the ACT experts who participated in the Delphi study and consented to being named, we want to acknowledge (in alphabetical order by first name): Professor Craig White, Dr Elizabeth Burns, Helen McGillivray, Jim Lucas, Dr Kirk Strosahl, Professor Lance McCracken, Lou Lasprugato MFT, Margot Osorio, Professor Nuno Ferreira, Dr Ray Owen, and Ross McIntosh. We also want to thank Professor Chris Graham for his contribution to the conceptualisation of the research that led to the development of the ACT SJT, and Dr Joe Oliver for his support and access to Contextual Consulting workshops for our field study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jcbs.2025.100949>.

References

- Akiba, C. F., Powell, B. J., Pence, B. W., Nguyen, M. X., Golin, C., & Go, V. (2022). The case for prioritizing implementation strategy fidelity measurement: Benefits and challenges. *Translational Behavioral Medicine*, 12(2), 335–342. <https://doi.org/10.1093/tbm/ibab138>
- Bai, Z., Luo, S., Zhang, L., Wu, S., & Chi, I. (2020). Acceptance and commitment therapy (ACT) to reduce depression: A systematic review and meta-analysis. *Journal of Affective Disorders*, 260, 728–737. <https://doi.org/10.1016/j.jad.2019.09.040>
- Barnes-Holmes, Y., Barnes-Holmes, D., McHugh, L., & Hayes, S. C. (2004). Relational frame theory: Some implications for understanding and treating human psychopathology. *International Journal of Psychology and Psychological Therapy*, 4, 355–375.
- Beidas, R. S., & Kendall, P. C. (2010). Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice*, 17(1), 1–30. <https://doi.org/10.1111/j.1468-2850.2009.01187.x>
- Bell, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., Ogedegbe, G., Orwig, D., Ernst, D., Czajkowski, S., & Treatment Fidelity Workgroup of the NIH Behavior Change Consortium. (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH behavior change consortium. *Health Psychology*, 23(5), 443–451. <https://doi.org/10.1037/0278-6133.23.5.443>
- Bhattacharyya, O., Reeves, S., & Zwarenstein, M. (2009). What is implementation research? Rationale, concepts, and practices. *Research on Social Work Practice*, 19(5), 491–502. <https://doi.org/10.1177/1049731509335528>
- Boone, M. S., Mundy, B., Morrissey Stahl, K., & Genrich, B. E. (2015). Acceptance and commitment therapy, functional contextualism, and clinical social work. *Journal of Human Behavior in the Social Environment*, 25(6), 643–656. <https://doi.org/10.1080/10911359.2015.1011255>
- Borrelli, B. (2011). The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *Journal of Public Health Dentistry*, 71, S52–S63. <https://doi.org/10.1111/j.1752-7325.2011.00233.x>
- Crowne, D. P., & Marlow, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4), 349. <https://psycnet.apa.org/doi/10.1037/h0047358>
- Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., & Wales, P. W. (2014). Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology*, 67(4), 401–409. <https://doi.org/10.1016/j.jclinepi.2013.12.002>
- Dindo, L., Van Liew, J. R., & Arch, J. J. (2017). Acceptance and commitment therapy: A transdiagnostic behavioral intervention for mental health and medical conditions. *Neurotherapeutics*, 14(3), 546–553. <https://doi.org/10.1007/s13311-017-0521-3>
- Fink-Hafner, D., Dagen, T., Doušák, M., Novak, M., & Hafner-Fink, M. (2019). Delphi method: Strengths and weaknesses. *Advances in Methodology and Statistics*, 16(2), 1–19. <https://doi.org/10.51936/fcfm6982>
- Francis, A. W., Dawson, D. L., & Goljani-Moghaddam, N. (2016). The development and validation of the Comprehensive assessment of Acceptance and Commitment Therapy processes (CompACT). *Journal of Contextual Behavioral Science*, 5(3), 134–145. <https://doi.org/10.1016/j.jcbs.2016.05.003>
- Frank, H. E., Becker-Haimes, E. M., & Kendall, P. C. (2020). Therapist training in evidence-based interventions for mental health: A systematic review of training approaches and outcomes. *Clinical Psychology: Science and Practice*, 27(3), 20. <https://doi.org/10.1111/cpsp.12330>
- Gloster, A. T., Walder, N., Levin, M. E., Twohig, M. P., & Karekla, M. (2020). The empirical status of acceptance and commitment therapy: A review of meta-analyses. *Journal of Contextual Behavioral Science*, 18, 181–192. <https://doi.org/10.1016/j.jcbs.2020.09.009>
- González-Fernández, S., & Fernández-Rodríguez, C. (2019). Acceptance and commitment therapy in cancer: Review of applications and findings. *Behavioral Medicine*, 45(3), 255–269. <https://doi.org/10.1080/08964289.2018.1452713>
- Graham, C. D., Gouick, J., Krahé, C., & Gillanders, D. (2016). A systematic review of the use of Acceptance and Commitment Therapy (ACT) in chronic disease and long-term conditions. *Clinical Psychology Review*, 46, 46–58. <https://doi.org/10.1016/j.cpr.2016.04.009>
- Gregory, M. E., Sonesh, S. C., Feitosa, J., Benishek, L. E., Hughes, A. M., & Salas, E. (2017). Decision making on the labor and delivery unit: An investigation of influencing factors. *Human Factors*, 59(6), 937–955. <https://doi.org/10.1177/0018720817703740>
- Harris, R. (2009). *ACT made simple: An easy-to-read primer on acceptance and commitment therapy*. New Harbinger Publications.
- Hayes, S. C., Barnes-Holmes, D., & Wilson, K. G. (2012). Contextual behavioral science: Creating a science more adequate to the challenge of the human condition. *Journal of Contextual Behavioral Science*, 1(1–2), 1–16. <https://doi.org/10.1016/j.jcbs.2012.09.004>
- Hayes, S. C., & Brownstein, A. J. (1985). Verbal behavior, equivalence classes, and rules: New definitions, data, and directions [Invited address presentation]. In *Meeting of the association for behavior analysis*. Columbus, OH, United States.
- Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behaviour Research and Therapy*, 44(1), 1–25. <https://doi.org/10.1016/j.brat.2005.06.006>
- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2011). *Acceptance and commitment therapy: The process and practice of mindful change*. Guilford press.
- Hayes, S. C., Strosahl, K. D., Wilson, K. G., Bissett, R. T., Pistorello, J., Toarmino, D., Polusny, M. A., Dykstra, T. A., Batten, S. V., Bergan, J., Stewart, S. H., Zvolensky, M. J., Eifert, G. H., Bond, F. W., Forsyth, J. P., Karekla, M., & McCurry, S. M. (2004). Measuring experiential avoidance: A preliminary test of a working model. *Psychological Record*, 54, 553–578. <https://doi.org/10.1007/BF03395492>
- Hsu, C. C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research and Evaluation*, 12(1). <https://doi.org/10.7275/pdz9-th90>
- Kohrt, B. A., Jordans, M. J., Rai, S., Shrestha, P., Luitel, N. P., Ramaiya, M. K., Singla, D. R., & Patel, V. (2015). Therapist competence in global mental health: Development of the ENhancing assessment of common therapeutic factors (ENACT) rating scale. *Behaviour Research and Therapy*, 69, 11–21. <https://doi.org/10.1016/j.brat.2015.03.009>
- Luoma, J. B., & Vilardaga, J. P. (2013). Improving therapist psychological flexibility while training acceptance and commitment therapy: A pilot study. *Cognitive Behaviour Therapy*, 42(1), 1–8.
- Murase, T., Roebuck, A., & Takahashi, K. (2019). Development and validation of a situational judgement test of Japanese leadership knowledge. *Asia Pacific Business Review*, 25(2), 227–250. <https://doi.org/10.1080/13602381.2018.1548543>
- Nasa, P., Jain, R., & Juneja, D. (2021). Delphi methodology in healthcare research: How to decide its appropriateness. *World Journal of Methodology*, 11(4), 116.
- O'Neill, L., Latchford, G., McCracken, L. M., & Graham, C. D. (2019). The development of the Acceptance and Commitment Therapy Fidelity Measure (ACT-FM): A Delphi study and field test. *Journal of Contextual Behavioral Science*, 14, 111–118. <https://doi.org/10.1016/j.jcbs.2019.08.008>
- Olaru, G., Burrus, J., MacCann, C., Zaromb, F. M., Wilhelm, O., & Roberts, R. D. (2019). Situational judgment tests as a method for measuring personality: Development and validity evidence for a test of dependability. *PLoS One*, 14(2), Article e0211884. <https://doi.org/10.1371/journal.pone.0211884>
- Ong, C. W., Barthel, A. L., & Hofmann, S. G. (2023). The relationship between psychological inflexibility and well-being in adults: A meta-analysis of the Acceptance and Action Questionnaire (AAQ). *Behavior Therapy*. <https://doi.org/10.1016/j.beth.2023.05.007>
- Patterson, F., Zibarras, L., & Ashworth, V. (2016). Situational judgement tests in medical education and training: Research, theory and practice: AMEE Guide No. 100. *Medical Teacher*, 38(1), 3–17. <https://doi.org/10.3109/0142159X.2015.1072619>
- Peckler, B., Prewett, M. S., Campbell, T., & Brannick, M. (2012). Teamwork in the trauma room evaluation of a multimodal team training program. *Journal of Emergencies, Trauma, and Shock*, 5(1), 23–27. <https://doi.org/10.4103/0974-2700.93106>
- Persich, M. R., & Robinson, M. D. (2020). The person in context: Using situational judgment to gain insight into social functioning. *Journal of Social and Personal Relationships*, 37(1), 302–324. <https://doi.org/10.1177/0265407519864447>

- Ployhart, R. E., & Ehrhart, M. G. (2003). Be careful what you ask for: Effects of response instructions on the construct validity and reliability of situational judgment tests. *International Journal of Selection and Assessment*, 11(1), 1–16. <https://doi.org/10.1111/1468-2389.00222>
- Plumb, J. C., & Vilardaga, R. (2010). Assessing treatment integrity in acceptance and commitment therapy: Strategies and suggestions. *International Journal of Behavioral Consultation and Therapy*, 6(3), 263–295. <https://doi.org/10.1037/h0100912>
- Prewett, M. S., Brannick, M. T., & Peckler, B. (2013). Training teamwork in medicine: An active approach using role play and feedback. *Journal of Applied Social Psychology*, 43(2), 316–328. <https://doi.org/10.1111/j.1559-1816.2012.01001.x>
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(2), 65–76. <https://doi.org/10.1007/s10488-010-0319-7>
- Rakovshik, S. G., & McManus, F. (2010). Establishing evidence-based training in cognitive behavioral therapy: A review of current empirical findings and theoretical guidance. *Clinical Psychology Review*, 30(5), 496–516. <https://doi.org/10.1016/j.cpr.2010.03.004>
- Sakamoto, R., Ohtake, Y., Kataoka, Y., Matsuda, Y., Hata, T., Otonari, J., Yamane, A., Matsuoaka, H., & Yoshiuchi, K. (2021). Efficacy of acceptance and commitment therapy for people with type 2 diabetes: Systematic review and meta-analysis. *Journal of Diabetes Investigation*. <https://doi.org/10.1111/jdi.13658>
- Samuel, V., Knight, L., & Roberts, A. (n.d.). Situational judgement questionnaire: Congruent ACT responses with young people (CoACT-Y). [unpublished questionnaire].
- Swain, J., Hancock, K., Hainsworth, C., & Bowman, J. (2013). Acceptance and commitment therapy in the treatment of anxiety: A systematic review. *Clinical Psychology Review*, 33(8), 965–978. <https://doi.org/10.1016/j.cpr.2013.07.002>
- Tiffin, P. A., Paton, L. W., O'Mara, D., MacCann, C., Lang, J. W., & Lievens, F. (2019). Situational judgement tests for selection: Traditional vs construct-driven approaches. *Medical Education*, 54(2), 105–115. <https://doi.org/10.1111/medu.14011>
- Waller, G., & Turner, H. (2016). Therapist drift redux: Why well-meaning clinicians fail to deliver evidence-based therapy, and how to get back on track. *Behaviour Research and Therapy*, 77, 129–137. <https://doi.org/10.1016/j.brat.2015.12.005>
- Webster, E. S., Paton, L. W., Crampton, P. E., & Tiffin, P. A. (2020). Situational judgement test validity for selection: A systematic review and meta-analysis. *Medical Education*, 54(10), 888–902. <https://doi.org/10.1111/medu.14201>
- Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review*, 19(3), 188–202. <https://doi.org/10.1016/j.hrmr.2009.03.007>
- Wolcott, M. D., Lobczowski, N. G., Zeeman, J. M., & McLaughlin, J. E. (2020). Situational judgment test validity: An exploratory model of the participant response process using cognitive and think-aloud interviews. *BMC Medical Education*, 20(1), 1–12. <https://doi.org/10.1186/s12909-020-02410-z>