

Counselling Skills and Competencies Tool: Scale development and preliminary data

pacja.org.au/2021/11/counselling-skills-and-competencies-tool-scale-development-and-preliminary-data

Jane L. Fowler, John G. O’Gorman, and Mark P. Lynch

The purpose of the study reported here was to develop an assessment tool for evaluating levels of competence reached by students completing a beginning-level course in counselling. These students were already working as counsellors, however had received no formal training in counselling prior to commencing the course. The assessment tool was developed to provide feedback to students on their progress towards the goals of the course and formed the basis of the summative assessment of their competence as counsellors.

The course was developed as part of a project directed at improving the counselling skills of staff working in human services in Papua New Guinea (PNG) (Fowler et al., 2021). An earlier review of the knowledge and skills of counselling practitioners in PNG by the Australian Department of Foreign Affairs and Trade (DFAT) indicated a great need for improvement in this area (DFAT, 2017). A key recommendation of the report was ensuring that core competencies were understood and practiced. The course was therefore directed towards developing a basic level of competence in counselling as it is practiced in Australia, characterised by Schofield (2013) as humanistic, client-centred, and experiential.

To meet its purpose as a summative assessment tool for the course, it was considered necessary that the CSCT demonstrate validity and reliability as per the Standards for Educational and Psychological Testing (AERA, 2014), that it be efficient in terms of the time and resources required for its implementation, and that it provide a concise summary of the judgements of the assessors to provide feedback to those assessed. As far as validity is concerned, content validity was considered the central issue, that is, whether the domain of counselling behaviours expected of a beginning student was adequately covered and that all critical behaviours were included. As for reliability, a number of approaches were recognised, but for a tool that relies on human judgement when assessing student performance, the important question was whether the scores provided by different assessors were comparable, that is, whether interrater agreement was high.

The steps followed in developing the tool were those used in the traditional development of assessment devices (Crocker & Algina, 2006; Davis, 1987; DeVellis, 2012): define the construct; review the relevant literature; choose a measurement model; write, trial, and edit items; check reliability and validity.

Counselling Competence Defined

The definition of counselling adopted here was the same definition outlined by Geldard et al. (2016): a process involving collaboration between counsellor and client so that the client might work through their problems and discover solutions. Foundational to this process are what have been termed micro-skills, such as attending to the client, questioning to open an issue, and reflecting on the meaning the client attaches to an experience. These micro-skills are observable and discrete aspects of the counsellor’s response to a client, and can be shaped by modelling, feedback, and repeated experience.

Although of continuing importance, competence as a counsellor involves more than mastering a set of micro-skills (Ridley et al., 2011). The definition of competence adopted was: “an individual’s capability and demonstrated ability to understand and do certain tasks in an appropriate and effective manner consistent with the expectations for a person qualified by education and training in a particular profession or speciality” (Kaslow, 2004, p. 775). Kaslow noted that the overall competence of a professional is not one thing, but is made up of proficiency in a number of more specific competencies. These are higher-level abilities to draw on knowledge, skills, and attitudes in meeting complex demands in a particular context. In counselling, these include the ability to show empathy for a client’s experience or to show positive regard towards a client. Thus, competence in counselling involves a set of competencies that involve micro-skills, but are not limited to them.

The level of competence that the course aimed to produce was defined as “basic” in the DFAT (2017) report. Basic counselling was defined as: “a set of primary and essential counselling skills... [including] empathy and unconditional positive regard (respect without judgement); supportive listening and questioning skills; the ability to establish trust, and explore issues and needs; and encouragement to make decisions” (DFAT, 2017, p.vi). Skilled or professional counselling was distinguished from basic counselling through the higher skill level of the counsellor, who would be able to “...reflect, make skilled observations, use effective questioning, facilitate coping mechanisms, summarise goals and priorities, conduct risk assessments, support the development of an action plan, and provide further guidance” (DFAT, 2017, p.vi).

Literature Review

A literature search was conducted to check if suitable evaluation tools were available before undertaking the development of a new scale. English language reports published in Canada, the United States, Great Britain, and Australia were searched from 1990 to 2018. The following narrative review is the result.

Traditionally, evaluation of competence in counselling has relied heavily on overall judgements or global impressions, usually by the supervisor or supervisors of the training program. Perceived limitations with these have led to the development of scales for more fine-grained assessment of the skills and competencies of the trainee. Loesch (1995) briefly reviewed methods used to assess the performance of counsellors, noting that while rating scales were the most frequently employed method, those making the ratings were often external to the counselling situation. He recognised that self-assessments were widely used because they assessed the important

Rogers' quality of emotional congruence between counsellor and client (Rogers, 1951), but the subjectivity of self-ratings had limited them to personal development rather than summative assessment. He argued that ratings by clients were useful, although the change process that might be initiated through counselling was best evaluated some time after counselling had commenced. Loesch considered the use of external raters as less intrusive, more flexible with respect to the counselling context, and more objective when ratings were based on video tapes of sessions. Ratings provided by supervisors, training peers, or professional colleagues were useful for providing feedback for development purposes, but in summative contexts ratings were typically provided by supervisors or colleagues with the assumed expertise to make such judgements.

Eriksen and McAuliffe (2003) reviewed the instruments developed to measure counselling competencies in the period 1960 to the time of their review. They set five criteria that a desirable instrument needed to meet, stating that measures must: "(a) be valid and reliable; (b) rely on observations of actual in session performance of counselling skills; (c) be accessible, that is, have face validity, be easy to use, and be relevant for students and instructors as a feedback device; (d) rely on ratings by expert judges, rather than on ones by students, clients, or peers; and (e) require qualitative judgments as to the contextual appropriateness of the use of particular skills." These criteria are consistent with Loesch's overview of methods and elaborate the three criteria considered initially to be important (see above).

Only three instruments in the Eriksen and McAuliffe review met at least some of these criteria, which led Eriksen and McAuliffe (2003) to develop an improved version of one of them, the Skilled Counseling Scale (SCS), originally developed by Urbani et al. (2002). Eriksen and McAuliffe replaced some of the items of the SCS and added new ones to produce a new 22-item scale. Additionally, they changed the rating method from recording the number of times a behaviour was observed (from 1 *not at all* to 5 *always*) to a more effective 5-point scale that ranged from *major adjustment needed* to *highly developed*.

Eriksen and McAuliffe's (2003) resulting Counselling Skills Scale (CSS) was designed for rating by experts (educators) who had viewed a full counselling session. Raters average the scores for the items in each of six subscales: (1) Shows interest and appreciation, (2) Encourages exploration, (3) Deepens the session, (4) Encourages change, (5) Develops therapeutic relationship, and (6) Manages the session.

Each of the items is defined on the instrument with a brief statement of the skill in terms of observable behaviour. Eriksen and McAuliffe reported interrater agreement for a convenience sample of five counsellors (including themselves) of 76.8%, a percentage that was achieved only after a focus group discussion about differences between raters. A simple percentage agreement index, it should be noted, does not adjust for chance agreement in the way an intraclass correlation coefficient (ICC) does (e.g., Gisev et al., 2013). Eriksen and McAuliffe reported an item analysis of the scale using a sample of 29 counsellors in training that resulted in item-total correlations of .18 to .71, with an internal consistency (Cronbach's alpha) coefficient of .90. Comparisons of total score and scores on the six subscales from the beginning to the end of the course indicated statistically significant increases for all but one subscale, with an effect size (Cohen's *d*) for a total score of .80 and effect sizes of .20 to .83 for the subscales, evidence of construct validity for the scale. The CSS represents a significant improvement on previous instruments, particularly in the provision of rating scales and scoring, but data on content validity and reliability is sparse, and the instrument is not generally available.

Since the work of Eriksen and McAuliffe (2003), four instruments designed to measure a form of counselling effectiveness have been reported in the literature. The Counselor Activity Self-Efficacy Scales (CASES) (Lent et al., 2003) include items that measure basic helping skills (only five resemble basic counselling skills: Open Questions, Listening, Reflection of Feelings, Restatements, and Attending), session management (in the form of a checklist), and counselling challenges. CASES is a self-efficacy measure, asking respondents to report on their confidence in their "ability over the next week to...". It examines generic perceptions about skills rather than specific counselling performance.

The second instrument, the Motivational Interviewing Treatment Integrity scale (Moyers et al., 2005), was specifically designed to document the integrity of delivery of motivational interviewing and therefore is not appropriate for measuring basic counselling skills. The third, the Counsellor Effectiveness Scale (Oluseyi & Oreoluwa, 2014), includes 35 items describing an effective counsellor in terms of personality characteristics (e.g., emotionally stable), personal qualities (e.g., has integrity), and performance indicators (e.g., helps clients to discover themselves), and thus does not measure basic counselling skills.

The fourth and most recent instrument is the Counseling Competencies Scale – Revised (CCS-R) (Lambie et al., 2018). Content for the CCS-R was developed through a review of counselling literature, expert input, and psychometric evaluation. Exploratory Factor Analysis resulted in a 23-item instrument with two factors: Counselling Skills and Therapeutic Conditions (12 items); and Counselling Dispositions and Behaviours (11 items), with an internal consistency coefficient (Cronbach's alpha) of .96. A subsequent Confirmatory Factor Analysis, however, failed to show a satisfactory fit of the data to this two-factor model. The developers acknowledged that the statistical and theoretical need to remove nine of the original 32 items from the instrument may have resulted in the loss of key aspects when measuring a counsellor's competency. The measure has only 12 items that measure counselling skills and therapeutic conditions, and only eight of those measure skills (viz. nonverbal skills, encouragers, questions, reflecting-paraphrasing, reflecting-reflection of feelings, reflecting-summarising, advanced reflection: meaning, and confrontation).

Whereas some of the skills listed on the CCS-R are clear and singular, such as "tell me more about...", others are less clear and include multiple components (for example, body position, eye contact, posture, distance from client, voice tone, rate of speech, use of silence, etc.). Similarly, questions are presented collectively as "use of appropriate open and closed questioning (e.g., avoidance of double questions)". The second factor, Counselling Dispositions and Behaviours, includes record keeping and task completion items. They were included because the instrument's purpose was to assess counselling students completing practicum experiences in a community counselling setting, and were not relevant for the purpose of assessing basic counselling skills.

Counselling skills are assessed on the CCS-R by viewing a 10- to 15-minute segment of a videoed counselling session and rating each item on a 5-point scale ranging from 1 (*harmful*) to 5 (*exceeds expectations*), with the middle point being *near expectations*. Interrater consistency was assessed using a two-way, mixed, average-measure ICC on Factor 1 (.91), Factor 2 (.56), and total score (.84). The

relevance of the ICC information that the authors reported on the practical use of the test can be questioned. The mixed design implies that raters were considered a fixed effect, that is, all possible raters in the population of interest were employed, rather than, as seems more likely, a random sample from the population. A consistency index means that only the rank order of raters' scores was of interest and not the degree of agreement across raters. Such agreement is important when there are standards of proficiency employed that the counselling student is expected to meet. In the same way, where the test is to be used by only one rater, the average value across a number of raters is of less interest than the value for a single rater. Although decisions the authors made regarding the ICC may have suited the purposes of their research, the results they obtained are of limited interest to those who would want to use the CCS-R.

In summary, the literature search revealed that the CSS (Eriksen & McAuliffe, 2003) and CCS-R (Lambie et al., 2018) were the only instruments available for rating and scoring the counselling performance of trainees. In terms of the five criteria proposed by Eriksen and McAuliffe (2003) for the evaluation of tools for assessing counselling performance (see above), both the CSS and the CCS-R are based on observations of performance (Criterion b) and ratings by expert judges (Criterion d) that involve qualitative judgements of appropriateness (Criterion e). They both show face validity and appear relatively easy to use (aspects of Criterion c) but their relevance for feedback depends on their validity. Both were based on the counselling literature and involved the use of expert judges in their development. To this extent, they show a degree of content validity (an aspect of Criterion a).

However, in the case of the CCS-R, the coverage of the characteristics of competence for a beginning counsellor may be questioned, as the subscale relating to Counselling Skills and Therapeutic Conditions, the relevant one for assessing counselling competence, consists of only 12 items. Furthermore, several items, possibly relevant, were deleted in the interests of factor analytic purity. The CSS shows better coverage of the construct domain with 22 items, about two thirds of which showed item-total correlations in excess of .41. There is, however, a lack of clarity in some of the item descriptions, and some include multiple aspects of behaviour. Internal consistency reliability (Cronbach's alpha) is high for both instruments (.94 for the Counselling Skills subscale of the CCS-R and .90 for the CSS). For both scales there are few data for interrater agreement. Although on this analysis against the criteria, the CSS would be judged the superior instrument, the full scale was not published with the article on its development and could not be located. We had a practical need for an instrument that could be used in teaching and assessing the skills and competencies of entry-level counsellors, and found that none of those reported to date met our requirements. Therefore, we developed the CSCT with the aim of providing a scale for the summative assessment of beginning competence in counselling, and with the objective of meeting the criteria outlined by Eriksen and McAuliffe (2003).

Measurement Model

The instruments reviewed above used the summated rating scale model, a common model in social science research (e.g., Spector, 1992). A number of dimensions, each considered to reflect a characteristic of the construct of interest, are rated on scales from 4 to 11 points, and the ratings on the scales combined, usually summed, form the overall rating. The justification for this is the demonstration that ratings on the scales are correlated with each other and with the overall rating, and that the average intercorrelation across the scales reaches some minimum threshold. A frequently used threshold is a Cronbach's alpha of .70. One drawback in using the model is that attempts to maximise the alpha coefficient can lead to deletion of scales that are important features of the construct (Boyle, 1990; Helms et al., 2006), as seems to have happened in the case of the CCS-R. More statistically elegant models are now available, but the summated rating scale has served well in developing assessment devices and was the one employed here.

A rubric format was used for the rating scales, with the performance criteria to be assessed presented as a chart with a number of items for each. Six-point scales were used to cover the expected range of performance, and indicators for the different rating levels were provided.

Item Writing and Editing

The skills and competencies included in the CSCT were drawn from three key sources. First, several well-known textbooks, commonly and internationally used in counselling education (viz., Corey, 2015; Geldard et al., 2016; Ivey et al., 2018) were reviewed to ascertain what are generally considered primary and essential counselling skills and competencies. From this review, a list of skills and competencies was compiled and grouped into five categories (Attending, Reflecting, Questioning, Therapeutic Alliance, and Core Counselling Conditions). Second, the list was presented to three experienced practicing counsellors, two of whom identified as women and one who identified as a man, aged from 44 to 58 years. They were known to the research team through teaching and counselling associations. The three counsellors had a total of 63 years' counselling experience between them (range 15 to 18 years), and 40 years' experience in supervising beginning counsellors. The experienced counsellors were invited to provide feedback on the list of skills and competencies and to identify any perceived omissions. Two of the three experienced counsellors recommended the inclusion of a sixth category, Facilitating the Session, as they felt that this set of competencies was insufficiently covered in counselling training, often leaving new counsellors without the ability to structure or guide a counselling session. As a third source, three beginning counsellors were asked what they considered to be the most significant inclusions and omissions in their skills training. All three beginning counsellors were between 12 and 24 months post-graduation from their Masters in Social Work program, which had included an applied counselling course. These three post-graduates were the first responders to an invitation sent to 10 post-graduates whom staff in the researchers' university department knew to be working in roles that involved counselling. Conversations with the beginning counsellors confirmed the list of skills and competencies that had been compiled, including the need for the category of Facilitating the Session.

Clear, specific, and precise definitions were written for each item. Each skill or competency was defined in a manner that: (a) did not presuppose counselling knowledge, (b) was considered by experienced counsellors to be a typical or acceptable definition for the skill or competency, and (c) was clearly distinguishable from every other skill or competency on the list. Meeting these requirements ensured that the definitions were relevant to, and understood by, beginning counsellors (e.g., trainees, teachers, and assessors) and that each skill and competency could be assessed independently of each other.

It is important to note that the three sources used to identify the skills and competencies for inclusion in the scale all assume an Anglosphere understanding of the experience and practice of dominant groups within mainstream Australian culture. Experiences and practices in minority cultures (e.g., people with disabilities, immigrant groups) may mean that these competencies and skills do not apply and are inappropriate for use in counselling with these groups.

The list of 23 skills and their associated definitions was provided to three people who were not involved in the counselling profession and three people who were experienced counsellors. The three counsellors identified as women, were aged from 38 to 61, and were a university lecturer with 18 years' experience teaching counselling and two practicing psychologists with a combined 32 years' counselling experience. The opinions of non-counsellors were also sought to help ensure that professional jargon was not being included in the definitions, so that feedback to beginning counsellors would be understandable. Each respondent was asked to rate each skill and its definition on a scale from 1 (*not at all*) to 7 (*a great extent*) for clarity, understandability (non-counsellors only), accuracy (counsellors only), and discriminability. The median score on all four 7-point scales was 7, with a range from 6 to 7. In accordance with feedback provided by the respondents, some minor adjustments were made. For example, changes were made to the wording of the two definitions for Respect and Unconditional Positive Regard, which the experienced counsellors had considered to be insufficiently discriminating. The items and definitions presented in Table 1 are the final list of counselling skills and competencies.

Table 1. Counselling Skills and Competencies Tool

	Skill or Competency	Definition	Well above expectations (5)	Above expectations (4)	Slightly above expectations (3)	Slightly below expectations (2)	Below expectations (1)
Attending	Eye contact	Direct gaze with occasional breaks, if culturally appropriate	Excellent eye contact for whole session	Good eye contact for majority of session	Adequate eye contact for majority of session	Adequate eye contact for some of session	Insufficient eye contact or looks away at inappropriate times
	Vocal quality	Tone of voice and speech rate communicates warmth and ease	Excellent vocal quality for whole session	Good vocal quality for majority of session	Adequate vocal quality for majority of session	Adequate vocal quality for some of session	Poor vocal quality, that is, too loud, soft, fast, or slow
	Body language	Open and relaxed posture, leaning forward, appropriate body gestures such as nods and facial expressions	Excellent body language for whole session	Good body language for majority of session	Adequate body language for majority of session	Adequate body language for some of session	Body language that is awkward; uncomfortable posture and gestures
	Verbal tracking	Minimal encouragers such as "uh-huh", "tell me more", and "what else?"; repeating key feeling and content words	Excellent verbal tracking for whole session	Good verbal tracking for majority of session	Adequate verbal tracking for majority of session	Adequate verbal tracking for some of session	Poor use of minimal encouragers; little encouragement for client to continue
Reflecting	Reflection of content	Paraphrase or reflect back using own words, what has been said	Excellent and accurate reflection of content throughout session	Good and accurate reflection of content throughout session	Adequate and accurate reflection of content throughout session	Adequate paraphrasing with some inaccuracies	Over or under use, direct repeating of client's words, or often inaccurate

	Reflection of feeling	Identify and acknowledge key emotions or feelings that have been expressed through verbal or non-verbal behaviour	Excellent and accurate reflection of feeling throughout session	Good and accurate reflection of feeling throughout session	Adequate and accurate reflection of feeling throughout session	Adequate reflection of feeling with some inaccuracies	Over or under use, misses key emotions, or often inaccurate
	Summarising	Draw together, recapture, and review what has been covered to a certain point, both during and at the end of the conversation	Excellent and accurate summarising throughout session	Good and accurate summarising throughout session	Adequate and accurate summarising throughout session	Adequate summarising with some inaccuracies	Over or under use, misses content or feelings, or too long
Questioning	Open questions	Elicit a response of more than a few words; usually start with "how" and "what"	Excellent and effective use of open questions throughout session	Good and effective use of open questions throughout session	Adequate and effective use of open questions throughout session	Adequate use of open questions, with limited effectiveness	Over or under use, irrelevant, or asks multiple open questions at one time
	Closed questions	Can be answered with a minimal response, often as little as "yes" or "no"	Excellent and effective use of closed questions throughout session	Good and effective use of closed questions throughout session	Adequate and effective use of closed questions throughout session	Adequate use of closed questions, with limited effectiveness	Over or under use, irrelevant, or asks multiple closed questions at one time
	Clarifying questions	Seek to understand what has been said, such as "Are you saying that...?"	Excellent and effective use of clarifying questions throughout session	Good and effective use of clarifying questions throughout session	Adequate and effective use of clarifying questions throughout session	Adequate use of clarifying questions, with limited effectiveness	Overuses or misses opportunity to ask clarifying questions
	Specifying questions	Seek concrete information and detail, such as "How long have you been...?"	Excellent and effective use of specifying questions throughout session	Good and effective use of specifying questions throughout session	Adequate and effective use of specifying questions throughout session	Adequate use of specifying questions, with limited effectiveness	Overuses or misses opportunity to ask specifying questions
	Elaborating questions	Seek further explanation or expansion of what is being discussed, such as "Could you tell me more about...?"	Excellent and effective use of elaborating questions throughout session	Good and effective use of elaborating questions throughout session	Adequate and effective use of elaborating questions throughout session	Adequate use of elaborating questions, with limited effectiveness	Overuses or misses opportunity to ask elaborating questions
Relationship & Rapport	Relationship	Counsellor is attentive, and actively engaging and bonding with client	Excellent attentiveness and engagement for whole session	Good attentiveness and engagement for majority of session	Adequate attentiveness and engagement for majority of session	Adequate attentiveness and engagement for some of session	Minimal attentiveness and engagement for majority of session
	Rapport	Counsellor and client are connected, in sync, and have shared understanding	Excellent connection, openness, and responsiveness for whole session	Good connection, openness, and responsiveness for majority of session	Adequate connection, openness, and responsiveness for majority of session	Adequate connection, openness, and responsiveness for some of session	Minimal connection, openness, and responsiveness for majority of session

Core Counselling Conditions	Congruence	Counsellor demonstrates genuine, authentic, and true self with client	Excellent display of authentic and true self for whole session	Good display of authentic and true self for majority of session	Adequate display of authentic and true self for majority of session	Adequate display of authentic and true self for some of session	Minimal display of authentic and true self for majority of session
	Unconditional positive regard	Counsellor accepts client completely, without judgement	Excellent display of care and non-judgement for whole session	Good display of care and non-judgement for majority of session	Adequate display of care and non-judgement for majority of session	Adequate display of care and non-judgement for some of session	Minimal display of care and non-judgement for majority of session
	Empathy	Counsellor communicates understanding of the challenges, thoughts and experiences of client	Excellent display of seeing client perspective for whole session	Good display of seeing client perspective for majority of session	Adequate display of seeing client perspective for majority of session	Adequate display of seeing client perspective for some of session	Minimal display of seeing client perspective for majority of session
	Respect	Counsellor values client as a person and treats them with dignity, consideration, and courtesy	Excellent display of treating client with dignity, consideration and courtesy for whole session	Good display of treating client with dignity, consideration and courtesy for majority of session	Adequate display of treating client with dignity, consideration and courtesy for majority of session	Adequate display of treating client with dignity, consideration and courtesy for some of session	Minimal display of treating client with dignity, consideration and courtesy for majority of session
Facilitating the Session	Open the session	Counsellor introduces the session and establishes guidelines and emotional safety	Excellent introduction that creates a safe space and clear guidelines in first few minutes of session	Good introduction that creates a safe space and clear guidelines in first few minutes of session	Adequate introduction with delayed creation of a safe space and clear guidelines	Adequate introduction with insufficient creation of a safe space and clear guidelines	Minimal introduction without creation of a safe space and clear guidelines
	Hear the story	Counsellor explores presenting issues with client by gathering information and drawing out stories, concerns, or problems	Excellent gathering of information and drawing out stories and concerns	Good gathering of information and drawing out stories and concerns	Adequate gathering of information and drawing out stories and concerns	Adequate gathering of information and drawing out stories and concerns; some key points missed	Minimal gathering of information and drawing out stories and concerns
	Prioritise primary concerns	Counsellor and client collaboratively assess and agree on key issues to be focussed and worked on	Excellent assessment and prioritisation of key issues to be worked on	Good assessment and prioritisation of key issues to be worked on	Adequate assessment and prioritisation of key issues to be worked on	Adequate identification of key issues; no assessment and prioritisation	Identify and focus on one issue at the expense of other issues
	Work on change	Counsellor applies methods and techniques of one or more therapeutic approaches to work with client to facilitate change or a solution to the problem	Excellent application of methods or techniques to facilitate change	Good application of methods or techniques to facilitate change	Adequate application of methods or techniques to facilitate change	Adequate application of methods or techniques; some overuse of advice-giving and problem-solving	Minimal application of methods or techniques for change, and overuse of advice-giving and problem-solving

Close the session	Counsellor brings conversation to end, allowing time to “change gear”, and ensures client is feeling psychologically safe to re-enter the world	Excellent closure of session, with client ready to re-enter the world	Good closure of session, with client ready to re-enter the world	Adequate closure of session, with client ready to re-enter the world	Adequate closure of session, but rushed and insufficient checking with client	Quickly and briefly ends session and insufficient checking with client
-------------------	-------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------	------------------------------------------------------------------	----------------------------------------------------------------------	-------------------------------------------------------------------------------	------------------------------------------------------------------------

To permit each skill or competency in the CSCT to be rated according to performance, rather than simply its absence or presence, 6-point Likert scales, from 0 (*harmful*) to 5 (*well above expectations*) were provided for each item. The inclusion of 0 (rather than 1) and its accompanying label of *harmful* as the lowest point of the scale was purposeful. Because the instrument was a measure of counselling performance, it was important that trainees and assessors were afforded the opportunity to acknowledge instances where a particular skill or competency was not just under- or over-utilised but performed in a manner that was detrimental to the client, hence yielding a score of 0. The descriptor *harmful* was adopted in recognition of counselling codes of conduct and ethics that state as one of their first principles, “do no harm” (e.g., American Psychological Association, 2017; Psychotherapy and Counselling Federation of Australia, 2017). Overarching categories of *highly competent* (scores 5 and 4), *competent* (3 and 2), and *not competent* (1 and 0) were included to help facilitate feedback about levels of competency.

We acknowledge that *detrimental to the client* and *harmful* are often consensus judgements made by those without the lived experience of the person being counselled and may be underestimates of the true harm.

To assist raters in their determination of a trainee’s level of performance, in addition to the provision of the 6-point scale and overarching competency ranges, specific descriptors were provided for each skill and competency at each level of the scale. For example, Reflection of Feeling, which was defined as the ability to “identify and acknowledge the key emotions or feelings that have been expressed through verbal or non-verbal behaviour”, had “Judgemental, dismissing, unrelated, or absent” as its descriptor for a score of 0 or *harmful*. This combination of specific scores for each skill and competency, the overarching level of competency, and the specific descriptor for each point of the rating scale was intended to make it as simple as possible to distinguish ratings and to facilitate discussion between rater and trainee.

Overall scores on the scales are first obtained by calculating an average score (out of 5) across items for each of the six categories and then adding these to provide an overall score out of 30. The averaging process was expected to provide more reliable scores than using the item scores individually.

Checking Validity and Reliability

The extent to which the scales of the CSCT represent a balanced and adequate sampling of relevant dimensions, knowledge, and skills was considered in the initial stages of scale development. When items were selected, well-known counselling texts and experienced and beginning counsellors were consulted to ensure adequate coverage of counselling skills and competencies. When definitions were written, similar sources were drawn on to confirm that the concepts had been defined clearly, representatively, and discriminately. As a further check, the tool was given to a new sample of three experienced counsellors who had supervised beginning counsellors and taught counselling courses. As noted earlier, the sources that we relied on in developing the scale limit its validity to use with those from dominant cultural groups and cannot generalise necessarily to those from marginalised groups.

These highly experienced clinical supervisors and educators, who had not been involved in the previous stages of development, were asked to review the tool and respond, on a scale from 1 (*not at all*) to 7 (*a great extent*), to four questions about the balance and adequacy of the tool as a measure of counselling skills and competencies. Mean response ratings are presented in Table 2. On the basis of these ratings and previous feedback and ratings during item selection and definition, content validity was judged to be satisfactory.

Table 2. Results of Respondent (Experienced Counsellor) Ratings for Comprehensiveness and Adequacy of the CSCT (n=3)

To what extent does this tool...	M	SD	Mdn	Range
...include skills and competencies demonstrated by an effective counsellor?	7.00	0.00	7.00	7.00
...exclude skills and competencies demonstrated by an effective counsellor?	6.67	0.58	7.00	6 to 7
...represent a balanced sampling of counselling skills and competencies?	6.67	0.58	7.00	6 to 7
...represent an adequate sampling of counselling skills and competencies?	7.00	0.00	7.00	7.00

A second source of validity evidence is in the examination of the internal structure of an assessment instrument, typically by using exploratory or confirmatory factor analysis. These techniques, however, call for large sample sizes. A preliminary examination of internal structure is possible through item analysis, which was conducted on the evaluations of 25 of the first intake of PNG students. There were 20 women respondents (age range 27 to 66 years, $M = 42.2$ years) and five men respondents (age range 32 to 57 years, $M = 43.6$ years), with gender classified by respondent self-report. No respondents identified as non-binary or gender diverse. Respondents came from the Highlands (6), Islands (5), Momase (7), and Southern regions (7) of PNG. A total of 25 respondents had Bachelor-level degrees (seven related to counselling) and five had Master-level degrees (one related to counselling). They worked in a range of positions in the government (17), non-government/faith based (5), or mining sectors (2). One was in private practice as a counsellor. All were sighted (not blind or visually impaired). Data on participant culture/ethnicity and neurodiversity were not collected.

These students participated in video-recorded simulated counselling sessions. Each participant was a counsellor, using the same role-play situation with an actor as a client. Participants were asked to “conduct a 30-minute counselling session” and students were rated by a member of the teaching staff. The duration of the tapes ranged from approximately 10 to 26 minutes, with a mean of 18 minutes. Scores on each of the 23 items of the total scale were correlated with the total across all items. The total score was corrected for each correlation by removing the contribution of the item being correlated to the total score, so that the overlap of item and total did not lead to a spuriously high correlation, a standard procedure in item analysis. These item/total correlations are shown in Table 3, with the means and standard deviations for the items and the item/scale correlations. The latter is the correlation of the item with the total for the scale to which the item belongs. Again, the correlation was corrected for the item/scale overlap.

Table 3. Item Statistics

Item	Mean	SD	Item/Scale Item/Total	
			<i>r</i>	<i>r</i>
Eye contact	3.25	0.590	0.786	0.430
Vocal quality	3.10	0.707	0.820	0.609
Body language	3.02	0.651	0.692	0.484
Verbal tracking	2.96	0.896	0.769	0.678
Ref. content	2.29	0.846	0.693	0.738
Ref. feeling	1.98	0.891	0.682	0.802
Summarising	1.77	0.921	0.482	0.621
Open	2.42	0.776	0.915	0.877
Close	2.42	0.776	0.915	0.877
Clarifying	2.38	0.824	0.867	0.828
Specifying	2.19	0.734	0.916	0.895
Elaborating	1.96	0.820	0.780	0.870
Relationship	2.71	0.833	0.850	0.759
Rapport	2.50	0.737	0.850	0.835
Congruence	2.56	0.712	0.683	0.786
Uncon.Pos.Reg	2.88	0.784	0.754	0.686
Empathy	2.40	0.821	0.649	0.791
Respect	2.69	0.704	0.736	0.702

Open session	2.00	1.383	0.268	0.390
Hear story	2.21	1.103	0.840	0.834
Prioritise	1.77	0.921	0.706	0.671
Work change	2.00	0.967	0.764	0.800
Close session	2.15	0.972	0.650	0.702

According to the summated rating scale model that guides scale construction, individual scales should correlate with the total score if they are reflecting the same construct. Table 3 indicates that this is the case for the CSCT. Although some item/total correlations are larger than others, all are substantial, justifying the addition of individual scale scores into an aggregate score. Also shown are the correlations of the individual scales with the total scale score, which again are substantial in all cases. The average item/scale and item/total correlations are summarised by Cronbach's alpha, an estimate of the internal consistency of a scale. The Cronbach alpha coefficient, calculated over all 23 items for the sample of 25 students, was .96. Table 4 summarises the alphas for each of the subscales.

Table 4. Alpha Coefficients for the Subscales of the CSCT

Scale	Alpha
Attending	.89
Reflecting	.78
Questioning	.96
Relating & rapport	.92
Core conditions	.86
Facilitating	.82

In evaluating these coefficients, Nunnally's (1967) recommendations are widely used. He proposed that in the course of developing an assessment device, alphas of .70 to .80 were satisfactory; for use in research, alphas of .80 to .90 were appropriate; but for assessments in the individual case, alphas better than .90 were required. The alphas for the total score and for both the Questioning and the Relating and Rapport scales meet the standard for individual use, with those for the remaining scales not far below the standard (the Reflecting scale being the one exception).

Comparison of the item/scale correlations with the item/total correlations in Table 3 indicates that some items are better than others in that they show stronger item/scale than item/total correlations. Each of the six scales should reflect a component of the construct assessed by total score, but it should, if it is a component in its own right, differ to some extent from the total. Where an individual scale has more in common with the total of all other items than it does with the items of its own scale, the discriminant validity of the individual scale is suspect. This is the case in nine of the 23 items, which, given the small sample size involved, must be a matter for further investigation. Currently, the total scale demonstrates the internal consistency to be expected if a single construct, such as counselling competence, were being assessed.

As a further check on the construct validity of the CSCT, scores at the commencement of the course for the 25 students were compared with their scores at the end of the course. A valid measure of beginning counsellor competence was expected to be sensitive to the effects of the course. Means and standard deviations for each of the subscales and for the total scale score are presented in Table 5. The table reveals there was an increase in all measured means from pre- to post-course, and further, there was a reduction in variance over that period (i.e., a decrease in the magnitudes of the standard deviations). *F* tests on the difference in variances indicate the change was statistically significant in all cases. Because the variances were unequal, a Welch's *t* test was used to compare the means, rather than a matched-pairs *t* test. The Welch's *t* test does not take into account the correlation between pre- and post-course scores and therefore is not as sensitive as the matched-pairs *t*. However, because of the unequal variances it was the preferred statistical test of the difference. Again, in all cases, the *t* values were statistically significant. These additional statistics are shown in Table 5, along with an estimate of the effect size in all cases.

Table 5. Change from Pre- to Post-Course for each of the Six Scales and Total Score (n = 24)

Scale	Mean Pre	Pre	Mean Post	SD Post	F	Welch's <i>t</i>	<i>d</i>
Attending	3.08	0.621	3.58	0.465	1.78*	3.13**	0.91
Reflecting	2.01	0.737	2.51	0.520	2.01*	2.72**	0.78
Questioning	2.27	0.725	2.64	0.495	2.14**	2.05*	0.60
Relating & rapport	2.60	0.755	3.03	0.507	2.22**	2.32*	0.67
Core conditions	2.63	0.634	3.03	0.435	2.12**	2.55*	0.74
Facilitating	2.03	0.822	2.68	0.470	3.06**	3.36**	0.97
Total	2.44	0.626	2.90	0.397	2.49**	3.04**	0.88

Notes: *F* value for comparison of Pre and Post variances; *df* = 23/23 in all cases

t is Welch's *t* value for comparison of Pre and Post means; *df* = 23 in all cases

** *p* < .01, * *p* < .05

A decrease in variance in performance as a result of a training course is not surprising as behaviour comes to approximate the ideal, and is as important an effect as an increase in the mean. The magnitude of the effect in all cases is within or beyond the range Cohen (1988) described as medium to large ($d = .5$ to $.8$), and is well beyond the threshold for a practical effect in education ($d = .4$) according to Hattie (2008), based on his extensive reviews of the educational research literature.

Reliability is the consistency with which scores are replicable across items or contexts, or, as in the present research, raters. Some level of variability is to be expected when different judges observe and rate candidates on behaviour scales. The question, however, is whether that variability is so great as to negate the value of the ratings provided. In an extreme case, if the rating assigned to a candidate depends on the individual rater and does not agree from one rater to another, the ratings cannot be used in all fairness to make decisions about candidates. The degree to which raters concur is one source of evidence about reliability. When there are multiple raters, agreement is typically estimated using the ICC across the ratings they provide. The ICC can be used to provide evidence for interrater consistency by determining if the scores given by the different raters produce the same rank order of candidates from, say, best to worst performing.

Alternatively, the ICC can be used to provide evidence of agreement or accuracy, that is, the extent to which the raters assign the same or highly similar scores. For example, two raters may assign scores so that the rank order of candidates is preserved, but one may typically assign higher scores than the other. Their ratings are consistent but not accurate. Whether the higher or the lower score is the more accurate cannot be determined from the ratings themselves, but a lack of accuracy does indicate a problem of reliability. Ideally, consistency and interrater agreement should both be high.

To assess interrater agreement, four raters individually assessed six participant videos on each of the 23 skills and competencies included in the tool. Next, mean scores were calculated for each rater and video on the six dimensions. For example, the mean score for Attending was generated from the Eye contact, Vocal quality, Body language, and Verbal tracking scales. Then, ICCs for absolute agreement were calculated on the mean scores for each of the six dimensions, using SPSS Statistics, Version 26.0 (IBM Corp, Armonk, NY). A two-way random effects model was assumed.

The ICCs for the Attending, Reflecting, Questioning, Therapeutic Alliance, Core Counselling Conditions, and Facilitating the Session dimensions were, respectively, .841, .822, .776, .759, .751, and .551 for the average across raters, and .628, .664, .736, .412, .575, and .627 for a single rater. The *F* value in all cases was statistically significant ($p < .01$). In interpreting ICCs, Rosner (2006) proposed the following guidelines: "ICC < 0.4 indicates poor reliability, $0.4 \leq \text{ICC} < 0.75$ as fair to good reliability, and $\text{ICC} \geq 0.75$ as excellent reliability." The findings for a single rater thus fall into the category of "fair to good reliability". We are currently working with variations on the instructions given to raters to determine if those values can be improved. Considering the subjectivity involved in rating individual demonstrations of counselling skills and competencies, results are promising for the developmental stage of the instrument.

Test-retest reliability was assessed by having two of the raters re-review and rate the same six videos four weeks after their initial rating. Product moment correlations were calculated for each rater on the mean scores for each dimension between the two time periods, because interest was in the preservation of rank order rather than absolute agreement. Correlations for the first of the two raters varied from .85 to .96 across the six subscales, and for the second from .73 to .99. In both cases, four of the six correlations were in excess of .90, the threshold for use of a test for individual assessment (Nunnally, 1967), and all were statistically significant ($p < .05$).

Discussion

This project was based on the need for an assessment tool to evaluate the level of competence reached by students completing a beginning-level course in counselling. A review of the available options for this purpose, using the criteria for desirable measurement outlined by Eriksen and McAuliffe (2003), indicated two possible instruments, the CSS and the CCS-R. However, the CSS was not

generally available and the brevity of the counselling component of the CCS-R raised concerns as to its content validity. Therefore, we undertook the development of a new counselling assessment tool for the summative assessment of students completing a beginning-level course in counselling, using a standard method of educational test development. The CSCT was the result.

The CSCT uses ratings by experienced teaching staff of videos of student counselling sessions, the preferred source for summative assessments (Loesch, 1995). The 23 items were based on a review of the skills and competencies of counsellors as described in widely used counselling texts in Australia and North America. The items were checked for relevance and comprehensiveness by three experienced counsellors and three beginning counsellors, and for intelligibility by a group of three lay people not involved in counselling. The selection of sources and checking by experts and students were designed to ensure content validity. Scale points were anchored by statements that helped assessors evaluate students' levels of skill performance in domains relevant to counselling. The statements also assisted teaching staff to provide feedback to students on areas for improvement. Lambie and Ascher (2016) stressed the importance of assessment instruments that not only evaluated skills but could be used to facilitate discussion about strengths and areas of growth.

The validity and reliability of the CSCT were examined by having content validity checked by an independent panel of three expert judges and by examining the internal structure of the scales using item analysis. The Cronbach alpha indicated the internal consistency of the CSCT at a standard expected for use of a test in making individual decisions. Construct validity was then checked by comparing the performance of students as assessed by the scales before and after completion of the counselling course. The sensitivity of the CSCT to the manipulation of counselling skills and competence was demonstrated through statistically significant and meaningful changes in scores. Test-retest reliability based on the data for two expert judges who made evaluations was better than satisfactory, and interrater agreement based on ICCs was good when the scales were used by individual raters.

When compared with the options available at the outset of the project, the CSCT showed better content validity than the CCS-R, which has only 12 rather than 23 items and thus lacks the same coverage of the domain of beginning counselling competence. It has at least comparable, if not better, internal consistency (Cronbach alpha of .96 versus .90) than the CSS, and comparable, if not better, construct validity in terms of sensitivity to the effects of a counselling course (Cohen's $d = .86$ versus $d = .80$). In terms of interrater agreement, examination of the CSCT using the ICC methodology suggests a superiority to the CSS that used uncorrected interrater agreement. The CSCT thus fulfilled our requirement for a summative assessment tool which improved upon the CSS and CCS-R, and also conformed to the criteria outlined by Eriksen and McAuliffe (2003).

The principal limitation of the present study was that it used a small and unrepresentative sample to develop the CSCT. Twelve judges were used in developing the scale. Of these, six were expert counsellors, three were beginning counsellors, and three were judges with no prior knowledge of counselling (the latter to ensure a jargon-free tool). A sample of 25 beginning counselling students from PNG was used for validation of the scale. As all were convenience samples, generalisation of the current findings will depend on further research. There are no strong reasons, however, to expect a failure of the scale to generalise to student cohorts being introduced to a humanistic, client-centred, and experiential approach to counselling, as the constructs on which it was based were drawn from counselling texts widely used in Australia and North America. Although there are cultural differences between Australia and PNG, students had completed high school and university in PNG, where the language of instruction is English.

Sample size is an important limitation. Sophisticated psychometric analysis, such as factor analysis, was precluded by the small size of the sample and future studies are needed to confirm the internal structure of the scale in terms of the number and item composition of the subscales. The six scales and the items for each were developed on logical and conceptual grounds to ensure content validity. It is important to examine this structure empirically, using for example confirmatory factor analysis to check the distinctiveness of the scales in practice. The outcome of such analysis would not necessarily lead to deleting items, because we maintain the view that content validity is best ensured by an adequate conceptual analysis of a domain of interest. It would, however, provide insights into the logic of scoring the instrument in terms of the subscales. Further evidence on reliability of scores over time is required.

Future work, as well as examining the psychometrics of the CSCT, needs to examine the value of the feedback the subscales provide from the point of view of counsellors in training, and to consider counsellor and client cultural diversity, disability, and neurodiversity (e.g., use of eye contact). We attempted to make the points on the scales clear and discriminating, but whether this was achieved when students were advised about their performance has yet to be determined. Although developed primarily for summative assessment, a useful instrument is one that also allows students to learn from the assessment of their performance. As Chow and Miller (2018) argued, immediate and ongoing feedback is a critical factor in improving the effectiveness and reliability of psychotherapy and counselling, when added to deliberate practice. Deliberate practice involves a "tight focus on repetitively practicing specific skills until they become routine" (Rousmaniere et al., p.8). The CSCT may help target skills and competencies for deliberate practice in the course of training.

Finally, the CSCT is a product of a particular theory of counselling and the social context in which counselling has traditionally been practised in North American and Anglo Australian contexts. This has informed the choice of constructs to be assessed, the wording of items, and the methodology of scale construction. This potentially limits the reach of the scale to those who are educated in these cultural contexts. Its utility to wider populations across a range of cultural diversity, disability, and neurodiversity remains to be tested.

Ethics Approval

All procedures performed in studies involving human participants were in accordance with the University's ethical standards (Human Ethics Protocol 2018/924) and the 1964 Helsinki Declaration and its later amendments, or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Authors

Jane L. Fowler, Associate Professor, School of Health Sciences and Social Work, Griffith University (Logan Campus), Meadowbrook, QLD.

John G. O’Gorman, School of Applied Psychology, Griffith University, Meadowbrook, QLD.

Mark P. Lynch, Simulated Learning Environment (SLE) Practice Learning Centre, School of Human Services & Social Work (Logan Campus), Griffith University, Meadowbrook, QLD.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association. <https://www.apa.org/science/programs/testing/standards>
- American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct*. American Psychological Association. <https://doi.org/10.1037/0003-066X.57.12.1060>
- Australia’s Department of Foreign Affairs and Trade (DFAT). (2017). *Review of counselling services in the Pacific*. Final report. Pacific Women Shaping Pacific Development. <https://www.dfat.gov.au/sites/default/files/pacific-women-shaping-pacific-development-support-unit-final-report.pdf>
- Boyle, G. J. (1991). Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? *Personality and Individual Differences*, 7, 305–310. [https://doi.org/10.1016/0191-8869\(91\)90115-R](https://doi.org/10.1016/0191-8869(91)90115-R)
- Chow, D., & Miller, S. D. (2018). The question of expertise in psychotherapy. *Journal of Expertise*, 1(2), 1–9. https://www.journalofexpertise.org/articles/JoE_2018_1_2_MillerHubbardChow_earlyview.pdf
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge. <https://doi.org/10.4324/9780203771587>
- Corey, G. (2015). *Theory and practice of counseling and psychotherapy* (10th ed.). Brooks/Cole.
- Crocker, L. M., & Algina, J. (2006). *Introduction to classical and modern test theory*. Holt, Rinehart, & Winston.
- Dawis, R. V. (1987). Scale construction. *Journal of Counseling Psychology*, 34(4), 481–489. <https://doi.org/10.1037/0022-0167.34.4.481>
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage.
- Eriksen, K., & McAuliffe, G. (2003). A measure of counselor competency. *Counselor Education and Supervision*, 43(2), 120–133. <https://doi.org/10.1002/j.1556-6978.2003.tb01836.x>
- Fowler, J. L., Lynch, M. P., & Larsen, J. (2021). Counselling knowledge and skills in Papua New Guinea: Identifying the gaps. *International Journal for the Advancement of Counselling*, 43, 164–178 (2021). <https://doi.org/10.1007/s10447-021-09422-4>
- Geldard, D., Geldard, K., & Yin Foo, R. (2016). *Basic personal counselling: A training manual for counsellors* (8th ed.). Cengage Learning.
- Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: Key concepts, approaches, and applications. *Research in Social and Administrative Pharmacy*, 9, 330–338. <https://doi.org/10.1016/j.sapharm.2012.04.004>
- Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge. <https://doi.org/10.4324/9780203887332>
- Helms, J. E., Henze, K. T., Sass, T. L., & Mifsud, V. A. (2006). Treating Cronbach’s alpha reliability coefficients as data in counseling research. *The Counseling Psychologist*, 34(5), 630–660. <https://doi.org/10.1177/0011000006288308>
- Ivey, A., Ivey, M., & Zalaquett, C. (2018). *Intentional interviewing and counseling: Facilitating client development in a multicultural society* (9th ed.). Cengage Learning.
- Kaslow, N. J. (2004). Competencies in professional psychology. *American Psychologist*, 59(8), 774–781. <https://doi.org/10.1037/0003-066X.59.8.774>
- Lambie, G., & Ascher, D. (2016). A qualitative evaluation of the Counseling Competencies Scale with clinical supervisors and their supervisees. *The Clinical Supervisor*, 35(1), 98–116. <https://doi.org/10.1080/07325223.2015.1132398>
- Lambie, G., Mullen, P., Swank, J., & Blount, A. (2018). The Counseling Competencies Scale: Validation and refinement. *Measurement and Evaluation in Counseling and Development*, 51(1), 1–15. <https://doi.org/10.1080/07481756.2017.1358964>
- Lent, R. W., Hill, C. E., & Hoffman, M. A. (2003). Development and validation of the counselor activity self-efficacy scales. *Journal of Counseling Psychology*, 50(1), 97–108. <https://doi.org/10.1037/0022-0167.50.1.97>
- Loesch, L. C. (1995). *Assessment of counselor performance* (ED388886). ERIC. <https://files.eric.ed.gov/fulltext/ED388886.pdf>
- Mitchell, M. L., & Jolley, J. M. (2007). *Research design explained* (7th ed.). Wadsworth.
- Moyers, T. B., Martin, T., Manuel, J. K., Hendrickson, S. M. L., & Miller, W. R. (2005). Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment*, 28, 19–26. <https://doi.org/10.1016/j.jsat.2004.11.001>
- Nunnally, J. C. (1967). *Psychometric theory* (1st ed.). McGraw-Hill.

Oluseyi, A. E., & Oreoluwa, S. V. (2014). Factorial composition of counsellor effectiveness scale. *World Journal of Education*, 4(4), 61–69. <https://doi.org/10.5430/wje.v4n4p61>

Psychotherapy and Counselling Federation of Australia. (2017). *PACFA: Code of ethics*. <https://www.pacfa.org.au/Portal/Prac-Res/Code-of-Ethics.aspx>.

Ridley, C. R., Mollen, D., & Kelly, S. M. (2011). Beyond microskills: Toward a model of counseling competence. *The Counseling Psychologist*, 39(6), 825–864. <https://doi.org/10.1177/0011000010378440>

Rogers, C. (1951). *Client-centered therapy: Its current practice, implications and theory*. Constable.

Rosner, B. (2006). *Fundamentals of biostatistics* (8th ed.). Cengage Learning. <https://au.cengage.com/c/isbn/9781305268920/>

Rousmaniere, T., Goodyear, R. K., Miller, S. D., & Wampold, B. E. (2017). *The cycle of excellence: Using deliberate practice to improve supervision and training*. Wiley. <https://doi.org/10.1002/9781119165590>

Schofield, M. J. (2013). Counseling in Australia. In T. H. Hohenshil, N. E. Amundson, & S. G. Niles (Eds.), *Counseling around the world: An international handbook* (pp. 335-348). American Counseling Association.

Spector, P. E. (1992). *Summated rating scale construction: An introduction*. Sage. <https://doi.org/10.4135/9781412986038>

Urbani, S., Smith, M. R., Maddux, C. D., Smaby, M. H., Torres-Rivera, E., & Crews, J. (2002). A measure of counselor competency. *Counselor Education and Supervision*, 42, 92–106. <https://doi.org/10.1002/j.1556-6978.2002.tb01802.x>

[Return to Articles](#)
