SPECIAL ISSUE

The Supervision Evaluation and Supervisory Competence Scale: Psychometric Validation

Craig J. Gonsalvez,[1] Geaty Hamid,[2] Nicole M. Savage[3], and Danielle Livni[3]

[1]School of Social Sciences & Psychology, Western Sydney University, [2]Department of Psychology, South Western Sydney and Sydney Local Health Districts Population Health, and [3]School of Psychology, University of Wollongong

Objectives: The last two decades has witnessed a strong endorsement of competency-based models for both practitioner training and professional supervision. The valid and reliable measurement of supervisee and supervisory competence is an essential step towards progress, yet currently there are few instruments that can claim to measure the range of supervisor competencies. The current study establishes the Supervision Evaluation and Supervisory Competence (SE-SC) scale as a new, psychometrically sound instrument.

Method: A total of 142 supervisees anonymously completed overall evaluations of supervision satisfaction and supervisor effectiveness and of specific supervisor competencies using the SE-SC instrument. The specific competencies were subjected to a hierarchical cluster analyses to determine the underlying structure of supervisory competence.

Results: The results supported a six-cluster solution that included (a) Openness, caring and support, (b) Supervisor's Knowledge and Expertise as Therapist, (c) Supervision Planning and Management, (d) Goal-Directed Supervision, (e) Restorative Competencies, and (f) Insight into and Management of Therapist-Client Dynamics and Reflective Practitioner Competencies. The results yielded excellent internal reliability, test–retest reliability, and concurrent validity for the six clusters, with high and meaningful correlations with subscales of the Supervisory Working Alliance Inventory (SWAI) and the Supervisory Styles Inventory (SSI). More importantly, the six clusters together better predicted overall scores on supervision satisfaction and effectiveness (85% of variance) than did subscales of the SWAI (56%) and the SSI (57%).

Conclusion: The SE-SC demonstrates good psychometric properties and is a useful scale to measure a supervisee's evaluation of supervisory competence.

Key words: competency scales; measurement of competencies; professional supervision; psychometric properties; supervision effectiveness; supervision evaluation; supervisory competence.

What is already known on this topic	**What this paper adds**
1 Evaluation of supervision satisfaction and effectiveness by supervisees are important components of supervision evaluation.	1 The development and psychometric validation of a new instrument, the Supervision Evaluation and Supervisory Competence (SE-SC) scale.
2 Available measures of supervision evaluation do not adequately represent the range of supervisor competencies.	2 Innovative cluster analyses to reveal the underlying structure of supervisory competence into six independent clusters.
3 Supervisory working alliance is a strong predictor of supervision satisfaction.	3 Evidence that competency clusters better predict supervision satisfaction and supervisor effectiveness than omnibus constructs such as supervisory alliance.

The last two decades has witnessed the rise and rise of competency-based models for practitioner training across the range of health professions within Australia and internationally, and the effects of this substantive paradigm shift have inexorably filtered into professional supervision training and practice. In the words of Milne and Watkins (2014) "… there has perhaps never been a more sweeping and all-pervasive singular change to so rapidly impact the entirety of the supervision enterprise than the advent of the competency movement itself" (p. 674). This emphasis is best exemplified by recent emphases on competencies in supervision guidelines recently formulated by major professional bodies (American Psychological Association, 2014; Association of State and Provincial Psychology Boards, 2015) including the recommendation of the Psychology Board of Australia (2013) that all accredited supervisor training be competency based.

Correspondence: Craig Gonsalvez, School of Social Sciences & Psychology, Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia.
Email: c.gonsalvez@westernsydney.edu.au

Competency Models of Supervision

For several decades, supervision models were classified under three broad categories: developmental models, psychotherapy-based supervision models, and process-based models of supervision (Bernard & Goodyear, 2014). More recently, competency-based models have emerged as a fourth category (e.g., Falender & Shafranske, 2004; Gonsalvez, Oades, & Freestone, 2002). Competency models have their origins in pedagogic theory, especially as they apply to education and training in medicine and other health related professions. Competency-based supervision has been defined as "a metatheoretical approach that explicitly identifies the knowledge, skills and attitudes that comprise clinical competencies, informs learning strategies and evaluation procedures, and meets criterion-referenced competence standards consistent with evidence-based practices (regulations), and the local/cultural clinical setting" (APA, 2014, p. 5; also see Falender & Shafranske, 2007).

It is useful to view competency-based supervision within the context of recent developments in the definition, description and classification of *practitioner* competence. Practitioner competence is conceptualised as a multi-dimensional construct comprising a range of independent competency domains (e.g., assessment, intervention, professionalism) and competency types (e.g., knowledge, knowledge-application, skills, relationship, and attitudes) acquired across a developmental trajectory from novice to expert (Fouad et al., 2009; Gonsalvez & Calvert, 2014; Rodolfa et al., 2013). Further, a recent empirical study demonstrates that individual competencies cluster into meaningful domains that, to a large extent, validate a priori theoretical models of competence (Gonsalvez et al., 2015). It is obvious that models of practitioner competence have applications for conceptualisations of *supervisory competence*. In the first instance, the multifaceted nature of practitioner competence demands that supervisors be capable of assessing, building and enhancing supervisee competence across these multiple domains and stages. Pilling and Roth (2014) who espouse a competence-perspective on supervision suggest four broad categories of supervisory competence: generic, specific, specific applications, and metacompetencies. Applied to role-based models of supervision, a competency approach would require supervisors to demonstrate competence in multiple supervisor roles (teachers, counsellors, consultant, assessor, coach, and mentor), and comprehensive assessment would require credible evaluation of performance in each role. A recent thematic analysis of supervision guidelines posits supervisory competence as comprising nine broad themes including ethics and professional practice, knowledge of the profession, diversity, reflective practice, supervisory alliance, structuring supervision, facilitating learning, supervision research, and evaluation (Olds & Hawkins, 2014).

Evaluation of Supervision and Supervisory Competence

If competencies and competency frameworks have become the very Zeitgeist of contemporary supervision practice (Holloway, 2012), competence assessment becomes a top-priority issue. This is because valid and reliable measurement of both supervisee and supervisory competence is both fundamental and central to the competency paradigm. Against such a theoretical backdrop, a question that assumes significance is: How well do questionnaires and supervision evaluation measures adequately represent supervisory competence? Further, how is assessment of supervisory competence different from assessment of supervision conducted within other models of supervision? There have been excellent recent reviews of instruments used in supervision. In a fairly extensive review, a comprehensive list comprising 49 instruments was subjected to close scrutiny before a smaller set of measures was identified (Wheeler & Barkham, 2014). Based on predetermined criteria (e.g., psychometric properties, applicability across theoretical approaches, brevity), six measures were recommended for both routine and occasional use: the Brief Supervisory Alliance Scale (Rønnestad & Lundquist, 2009), The Development of Psychotherapist Common Core Questionnaire (Orlinsky & Rønnestad, 2005), Role Conflict and Role Ambiguity Inventory (Olk & Friedlander, 1992), The Helpful Aspects of Supervision Questionnaire (Llewelyn, 1988), and the Supervisory Working Alliance Inventory (SWAI; Efstation, Patton, & Kardash, 1990). A second review of instruments identified eight commonly used measures including the Supervisory Relationship Questionnaire (SRQ, Palomo, Beinart, & Cooper, 2010), the Supervisory Relationship Measure (Pearce, Beinart, Clohessy, & Cooper, 2013), the Supervision Attitude Scale (Kavanagh et al., 1997), the Supervisory Styles Inventory (SSI; Friedlander & Ward, 1984), the Evaluation Process within Supervision Inventory (Lehrman-Waterman & Ladany, 2001), and two other instruments short-listed in the previous review: the SWAI (Efstation et al., 1990) and the Role Conflict and Role Ambiguity Inventory (Olk & Friedlander, 1992).

A comprehensive review of instruments used in supervision is both unnecessary and beyond the scope of the current article. However, several important points related to the questions posed may be deduced from the extant literature.

Supervision satisfaction versus supervision effectiveness

First, a large number of instruments fail to differentiate between satisfaction with supervision and supervision effectiveness (e.g., the SSI, Friedlander & Ward, 1984; the Evaluation Process within Supervision Inventory, Lehrman-Waterman & Ladany, 2001; Supervisee Levels Questionnaire, McNeill, Stoltenberg, & Pierce, 1985). The emergence of competency-based approaches highlights the need for conceptual clarity, not just between supervision satisfaction and effectiveness as demanded by O'Donovan and Kavanagh (2014), but also between the two concepts and supervisory competence. Supervision satisfaction measures the supervisee's subjective perception of how well or badly supervision progressed, and is therefore likely to be influenced by supervisee-factors (knowledge, sophistication, expectations, and developmental stage), supervisor factors (e.g., caring and support) and context-factors (e.g., whether supervision was freely provided or paid for; individual versus group supervision; see Edwards et al., 2005). Supervision effectiveness is "the effect of clinical supervision on their [the supervisee's] professional development, improvement in skills, time

for reflection, and the quality of the supervisory relationship" (Edwards et al., 2005, p. 407). In other words, supervisor effectiveness is associated with enhancing supervisee competence across the spectrum of assessment, intervention, and professional domains. Satisfaction with supervision can coincide and correlate with supervisor effectiveness. For instance, being a caring and supportive supervisor may contribute to high ratings of satisfaction and may also be associated with supervisor effectiveness. However, an effective supervisor is more than a person who offers support and encouragement. Because constructive challenge and even confrontation to promote insight and self-awareness may be warranted to foster growth into competence, high levels of effectiveness may not always be associated with high levels of satisfaction (O'Donovan & Kavanagh, 2014). Supervisory competence refers to the extent to which supervisors exhibit a range of specific competencies embedded within a framework prescribed by a regulatory authority, guided by expert consensus, or determined by a theoretical approach.

Supervision effectiveness can be evaluated at first-order, supervisee levels (e.g., extent to which supervisee competence is enhanced) and at second-order, flow-on effects on client outcomes (individual/family/organisation) deriving from both supervisor and supervisee competence. Thus, competency-based theories would posit that supervisory competence mediates effectiveness at both first-order levels on supervisee competence, and at second-order levels on client outcomes. At each order, effectiveness may be evaluated by supervisee, supervisor (who provided supervision), peer supervisors, or supervisor experts, and evaluations can be made either by subjective ratings (supervisee ratings of improved competence) or by objective evidence of improved competence (e.g., supervisee demonstrates the use of higher levels of competence in simulated situations). The primary implication for measurement is that any overall measure of supervision evaluation must necessarily go beyond assessing supervisee satisfaction, and incorporate elements of supervisory competence and/or supervision effectiveness.

Second, a quick scan of the measures indicates an overriding emphasis on the measurement of the working alliance and related constructs (e.g., the Brief Supervisory Alliance Scale, the SWAI, the SRQ, and the Supervisory Relationship Measure). Although the contribution of supervisory working alliance to supervision satisfaction and effectiveness is not discounted, "measures of alliance and relationship quality may not capture all of the elements necessary for effective supervision" (O'Donovan & Kavanagh, 2014, p. 464). For instance, competency-based theorists identify working alliance as one of seven other generic supervision competencies, which itself is only one of four broad categories of supervisory competency (Pilling & Roth, 2014). It is indeed possible that a pivotal concept such as working alliance may interact in a complex way with other competencies, contribute to and even serve to mediate their effects. However, the unstated assumption inherent in all competency approaches is that overall evaluations of supervision and indeed of supervisor effectiveness would best be predicted by a combination of key competencies rather than by a single or a small number of global processes. The onerous task of identifying, defining, and categorising lists of competencies

would make little sense if a single omnibus variable such as working alliance could better predict supervisor effectiveness. Although the competency stance in favour of a key list of competencies makes intuitive sense, and can be justified as being consistent with pedagogic principles, in actual fact, we are unaware of research that has subjected this assumption to empirical scrutiny. On the other hand, the supervision literature is replete with research that supports the supervisory working alliance as a key variable (see Beinart, 2014).

Third, there is limited research especially on the measurement of supervisory competence. This is understandable because competency-based models of supervision themselves are relatively new (Falender & Shafranske, 2004; Gonsalvez et al., 2002) and although there is preliminary agreement about broad themes within competence (Olds & Hawkins, 2014), there is still need for both empirical data and an enhanced understanding of the nature and core elements of supervisory competence. A recently published instrument, Supervision: Adherence and Guidance Evaluation (SAGE, Milne & Reiser, 2014), is designed to measure competence in cognitive-behaviour therapy (CBT) supervision based on direct observation, and is a valuable first step although additional work on the psychometric properties of the instrument needs to be undertaken. It is obvious that "the absence of a tool for assessing competence in supervision is a serious omission" (Milne & Reiser, 2014) and additional instruments specifically designed to measure the various components of supervisory competence both from the perspective of the supervisee, peers and expert supervisors is a need the current study intends to address.

Finally, a critical evaluation of existing instruments highlights two additional points. Although the demand articulated by earlier reviewers for the development of more measures to evaluate supervision process and outcomes has been heeded by the creation of a number of instruments in the recent past, there is a need for measures that can demonstrate adequate scientific and psychometric rigour (Milne & Watkins, 2014). For instance, more than half of the 49 measures received low scores when assessed by Wheeler and Barkham (2014). Further, past research has largely focussed on the supervisee's and/or the supervisor's subjective evaluation of supervision process or outcomes. Although these appraisals are of salience, peer and expert ratings based on behavioural data (e.g., through videotapes of supervisor–supervisee interactions) and flow-down effects of supervisor effectiveness in terms of enhanced therapist competence and these effects on client outcomes are lacking.

Rationale for the Study

The primary aim of the study was to design a new scale to assess the supervisee's evaluation of supervision and ratings of supervisory competence and to establish its psychometric credentials. Second, the study intended to test the prediction of competency-based theories, namely that a combination of key supervisor competencies would better predict overall supervision evaluation scores than would omnibus process-oriented scores such as supervisory working alliance or supervisory styles.

The study addresses several recommendations from recent literature reviews on supervision instruments. First of all, unlike most other instruments, choice of the content of the items was primarily driven by competency-based theories of supervision, in this instance, the model advocated by Gonsalvez et al. (2002). As a result, supervisor behaviours and competencies emerging from several competency domains (e.g., intervention, professionalism), and across the many supervisory roles (e.g., teacher, counsellor, consultant) and across competency types (knowledge, skills, attitudes, relationship) were included in the scale. In addition to the independent itemised competencies, an overall scale was compiled to produce a global evaluation supervision score that comprised the supervisee's overall satisfaction with supervision and a rating of the extent to which supervision actually improved the supervisee's overall competence levels.

Although the supervisee's appraisal of supervisory competence and effectiveness are essential aspects of supervision evaluation, they remain insufficient measures. Supervisees, especially during early developmental stages, may have simplistic and naive appraisals of and poor discrimination between, levels of supervisory competence. It is evident that evaluations by supervisees should necessarily be supplemented by assessments by professionals better qualified to pass judgment on supervisors, including informed peers and supervision experts. It is also worth noting that the current study uses the term "supervision effectiveness" in a narrow sense, representing first-order effects of supervision on trainee competence. Whilst this is consistent with use of the term in the supervision literature (e.g., Barnett, Erickson Cornish, Goodyear, & Lichtenberg, 2007; Edwards et al., 2005), an examination of flow-down effects on client outcomes was beyond the scope of the current project and was not investigated.

Method

Participants

The sample comprised a total of 142 supervisees who were relatively young (mean age = 30.2; *SD* = 8.2) and mostly women (83% female; 17% male) with age and sex distributions being representative of current demographic profiles of psychology trainees in Australia. All supervisees had completed an accredited, 4-year undergraduate degree in psychology (Bachelors with Honours, or equivalent) with the vast majority pursuing further studies in psychology either through a current enrolment in a Professional Masters (*n* = 99 or 70%) or Doctoral degree (*n* = 17 or 12%) in an accredited training institution in Australia. The remaining supervisees (*n* = 26 or 18%) were enrolled in a 2-year internship programme approved by the Psychology Board of Australia. The group was largely inexperienced with 61%, 17%, and 22% having respectively completed 500 or less, 501–1,000, and over 1,000 hr of placement experience before they commenced the current placement they evaluated. Of 80% of trainees (*n* = 114) who responded, 81% reported their preferred theoretical orientation as behavioural, cognitive or CBT, 6% as psychodynamic, and 13% as a mix of other orientations (eclectic, narrative, etc.).

All supervisees were undertaking a clinical placement as part of their training when they were invited to participate in the online survey through a group email, and evaluated the supervision they received and the competencies of the principal supervisor at their current or recently completed placement. All participants completed the Supervision Evaluation and Supervisory Competence (SE-SC) scale, a subgroup (*N* = 70) also completed the SWAI (SWAI-supervisee version), and a second subgroup (*N* = 51) completed the SSI. Further, to examine test–retest reliability, a small group (*N* = 20) were re-administered the SE-SC scale 4–6 weeks after initial completion.

Supervisors of clinical psychology trainees were clinical psychologists who had a recognised postgraduate qualification in clinical psychology and were deemed eligible to supervise by the concerned university. Trainees undertaking an internship approved by the Psychology Registration Board were supervised by board-accredited supervisors.

Materials

Data from the following three measures are relevant to the study, and each of the measures is described below.

SE-SC scale

The SE-SC scale is a 31-item Likert, scale that comprises an overall evaluation of supervision (6-items) subscale, and an additional 25 items that lists a range of supervisor competencies. The scale derived from an early competency-based model of supervision (Gonsalvez et al., 2002), with the individual items reflecting discrete competencies across supervisor roles (therapist, teacher, assessor, facilitator) and across competency types (knowledge, skills, relationship, and attitude-value competencies) and has been in use to evaluate supervision within a clinical training programme over a 12-year period. Although the scale was largely designed for an assessment of individual formative supervision, a few items also cover group supervision and summative assessment. Each item is rated on a 7-point Likert scale ranging from 1 (*not at all, strongly disagree*) to 7 (*very much so, strongly agree*). All items are worded positively (e.g., the supervision significantly enhanced my competencies as a therapist and professional) with higher scores indicating better supervision outcomes and higher competence levels of the supervisor. An earlier version of the scale, called the supervision evaluation scale was used in a previous study and yielded good internal consistency and high correlations with other measures (Livni, Crowe, & Gonsalvez, 2012). The SE-SC generates a mean "overall score" that is a measure of both supervision satisfaction (three items; e.g., overall the supervisor provided me with high quality supervision) and supervision effectiveness (three items; e.g., the supervisor significantly enhanced my competence as a therapist and professional), and separate scores for each of 25 independent items measuring a range of other supervisor competencies. Each item is scored on a scale from 1 to 7, with higher scores indicating better supervision outcomes and/or higher levels of supervisory competence.
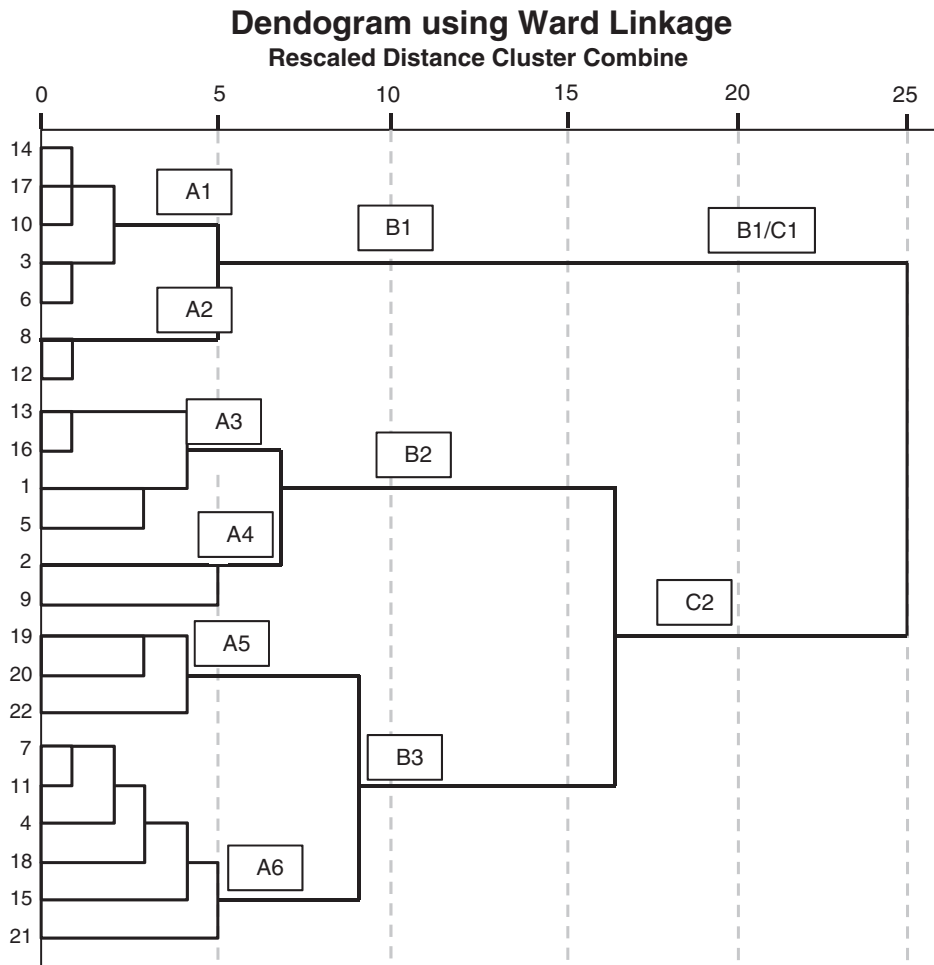
Dendogram using Ward Linkage
Rescaled Distance Cluster Combine

Figure 1 Dendrogram Depicting the Hierarchical Clustering of Competencies Derived from the Supervisor Evaluation and Supervisory Competence (SE-SC) Scale.

Supervisory Working Alliance Inventory

The SWAI (Efstation et al., 1990) is a self-report measure that assesses the perceived strength of the supervisory relationship. The SWAI has two versions with separate questionnaires for each member within the supervisor–supervisee dyad. The supervisee form comprising 19 items and two subscales (client focus, rapport) was administered in this study. Items are rated along a 7-point Likert scale ranging from 1 (*almost never*) to 7 (*almost always*). Items include questions such as "I feel comfortable working with my supervisor". The SWAI scales have been demonstrated to have acceptable levels of inter-item consistency of $\alpha = .71$ and above, and adequate divergent and convergent validity (Efstation et al., 1990).

Supervisory Styles Inventory

The SSI (Friedlander & Ward, 1984) is a 33-item measure of supervision that assesses both supervisor and supervisee of supervisor styles. The measure contains three subscales—Task Orientation, Interpersonal Sensitivity, and Attractiveness. The SSI is presented as a list of 25 adjectives only (e.g., open, intuitive, flexible) and participants are asked to rate their view of

the supervisor on a 7-point Likert scale ranging from 1 (*not very*) to 7 (*very*). The initial reliability study of the SSI found alphas ranging from .76 to .93 for both versions of the measure (Friedlander & Ward, 1984).

Procedure

Supervisees were invited to complete the measures at the end of their clinical placement via email and were provided a link to an online survey that administered the three measures if they consented to participate. Supervisees were informed that their evaluations would remain anonymous, and were identified by a code. Participants were offered the opportunity to enter a "book prize" competition of a $200 book voucher. Ethics approval for the project was obtained from the university's Human Research Ethics Committee.

Analyses

The statistical package, SPSS version 22 (2015) was used for all analyses. Instead of imposing theoretical and a-priori classification of items into subscales, an innovative, empirical method to determine the clustering of items into

Table 1 The Clustering of Individual Supervisor Competencies into Clusters, Super-Clusters (B-Clusters) and Mega-Clusters (C-Clusters) Based on Hierarchical Clustering Analyses

Items	A-clusters (RD = 4 or 5)	B-clusters (RD = 9–16)	C-clusters (RD = 17–24)
In day-to-day dealings, I got along well with the supervisor	A1: Openness, Caring, and Support (5 items)	B1: Supervisor as Person and Therapist (7 items)	C1: Supervisor as Person and Therapist (7 items)
The supervisor was understanding and open to a sharing of ideas			
The supervisor was accepting of my mistakes and inadequacies			
The supervisor was caring and supportive			
The supervisor was approachable and interested in my personal and professional development			
The supervisor impressed me as a skilled therapist	A2: Supervisor's Knowledge and Expertise as Therapist (2 items)		
The supervisor was knowledgeable and could communicate theoretical concepts clearly			
The supervision plan appropriately reflected important clinical competencies	A3: Supervision Planning and Management (4 items)	B2: Competency-Based Supervision (6 items)	C2: Supervisory Competencies (15 items)
Supervision objectives were in accordance with my level of professional development			
The supervisor organised and managed supervision efficiently			
Supervision methods were varied to match supervision objectives			
Supervision objectives (goals) were negotiated and clearly articulated	A4: Goal-Directed Supervision (2 items)		
Supervision sessions were structured and supervision activities were goal driven			
I felt comfortable discussing my professional inadequacies in supervision.	A5: Restorative Competencies (3 items)	B3: Restorative, Relationship, and Reflective Practitioner Competencies (9 items)	
The supervisor was sensitive to my emotional and self-care needs			
Supervision facilitated emotional ventilation and support as appropriate			
The supervisor enhanced my abilities to reflect on my clinical work	A6: Insight into and Management of Therapist-Client Dynamics and Reflective Practitioner Competencies (6 items)		
The supervision sessions enhanced my self awareness as a person			
The supervision furthered my understanding of my own positive and negative interaction patterns with clients			
The supervisor helped me gain an understanding of my emotional reactions within therapy			
The supervisor helped inspire me to remain excited about my clinical work and professional responsibilities			

Note. RD, rescaled distance.

subscales, the hierarchical clustering statistical technique was employed to the new measure. This technique has been used effectively in many studies, and can be used to either cluster cases as is done in conventional cluster analyses, or to cluster variables to determine the cluster structure of items within a scale or inventory, as is done in the current study. This clustering technique has several advantages including the capability to plot (the plot is called a dendrogram) how the items form clusters and super clusters as the level of kinship or proximity is systematically manipulated from close to distant kinship. The index of proximity is the rescaled distance (RD) unit that ranges from 1 to 25 with tight associations indicated by low numbers. The maximum RD of 25 represents a point when all items merged into a single all-encompassing cluster.

Internal consistency measures for the overall and subscales/clusters of the instruments were determined by the Cronbach's alpha statistic. Concurrent validity was assessed by computing correlations between the subscales of the SE-SC and subscales of the SWAI and SSI. Finally, three independent multiple linear regression analyses were conducted to determine the extent to which the subscales of each of the three instruments (the SE-SC, the SWAI, and the SSI) could predict the overall supervision evaluation scores.

Table 2 Means For and Correlations Between Subscales of the Supervision Evaluation and Supervisory Competence (SE-SC), Supervisory Working Alliance Inventory (SWAI), and Supervisor Styles Inventory (SSI)

Cluster	*M* (*SD*)	SWAI-rapport	SWAI client-focus	SSI attractive	SSI IPS	SSI-task oriented
Cluster A1: Openness, Caring, and Support	6.04 (1.03)	.82**	.50**	.65**	.87**	.36*
Cluster A2: Supervisor's Knowledge and Expertise as Therapist	6.07 (1.00)	.61**	.59**	.63**	.77**	.35
Cluster A3: Supervision Planning and Management	5.29 (1.07)	.68**	.71**	.49**	.68**	.54**
Cluster A4: Goal-Directed Supervision	4.96 (1.37)	.52**	.67**	.39*	.50**	.68**
Cluster A5: Restorative Competencies	5.49 (1.32)	.80**	.59**	.64**	.84**	.27
Cluster A6: Insight into and Management of Therapist-Client Dynamics and Reflective Practitioner Competencies	5.44 (1.09)	.42**	.67**	.52**	.68**	.15
Cluster B1: Supervisor: Person and Therapist	6.05 (0.95)	.81**	.55**	.68**	.88**	.37*
Cluster B2: Competency-Based Planning and Implementation	5.17 (1.10)	.64**	.72**	.47*	.64**	.61**
Cluster B3: Restorative, Relationship, and Reflective Competencies	5.46 (1.09)	.62**	.70**	.60**	.79**	.21
Cluster C1: Supervisor: Person and Therapist	6.05 (0.95)	.81**	.55**	.68**	.88**	.37*
Cluster C2: Supervisory Competencies	5.37 (1.00)	.68**	.77**	.60**	.79**	.41*
Overall supervision satisfaction and effectiveness	5.69 (1.07)	.67**	.59**	.59**	.75**	.35

Note. Att, attractive, IPS, interpersonal sensitivity; *N* for SE-SC scales = 140, for SWAI scales = 70, and SSI scales = 51; **p* < .01; ***p* < .001.

Results

Cluster Analyses

Four items that related to summative assessment and group supervision were endorsed as "not-applicable" by a third of the sample of supervisees, and were omitted from further analyses. The results from the cluster analyses of the remaining 22 items are presented visually in the traditional form of a dendrogram (see Figure 1). Because the overall items constituted an independent scale against which trends were to be compared, it was inappropriate to include these items within the cluster analyses. As depicted in Figure 1, when a strict measure of proximity or kinship (RD units between 4 or 5) was adopted, the

22 supervisor competencies congregated into six clusters that were named: Openness, Caring, and Support (see Table 1 and Figure 1; A1, five items), Supervisor's Knowledge and Expertise as Therapist (A2, two items), Supervision Planning and Management (A3, four items), Goal-Directed Supervision (A4, two items), Restorative Competencies (A5, three items), and Insight into and Management of Therapist-Client Dynamics and Reflective Practitioner Competencies (A6, five items).

When the proximity unit was relaxed (RD = 7), the six A-clusters (A1–A7) re-grouped into three larger, B-clusters, namely, Supervisor as Person and Therapist (B1-cluster, seven items, comprising both A1 and A2 items), Competency-based Supervision: Competency Development Planning and Implementation (B2-cluster, six items, comprising A3, A4, and A5 items), and Restorative,

Table 3 Internal Consistency Measures (α) and Multiple Regression Results From the Subscales of Each of the Three Supervision Instruments: the Supervisory Styles Inventory (SSI), Supervisory Working Alliance Inventory (SWAI), and the Supervision Evaluation and Supervisory Competence (SE-SC) Scale

Supervision instruments	*N*	α	*R²*	β	*t*	*p*
SSI subscales	51		.58			
Attractive		.98		.23	1.99	.05
Interpersonal		.92		.62	5.03	<.001
Task-oriented		.87		−.01	−.14	.89
SWAI subscales	70		.57			
Rapport		.95		.42	4.45	<.001
Client focus		.88		.45	4.75	<.001
SE-SC subscales	137		.84			
Cluster A1: Openness, Caring, and Support		.92		.27	4.00	<.001
Cluster A2: Supervisor's Knowledge and Therapist Expertise		.80		.26	4.87	<.001
Cluster A3: Supervision planning and management		.86		.06	5.07	<.001
Cluster A4: Goal-Directed Supervision		.75		−.02	−.24	.81
Cluster A5: Restorative Competencies		.84		−.11	−1.80	.07
Cluster A6: Insight into and Management of Therapist-Client Dynamics and Reflective Practitioner Competencies		.90		.36	6.28	<.001

Relationship, and Reflective Practice Competencies (B3-cluster, nine items, comprising A6 and A7 items). The B3-clusters were robust and stable, with these clusters retaining their kinship within their groups over an extended scale distance (RD from 7 to 16). At a RD of 17, the three clusters merged into two mega clusters: The B1-cluster remained unchanged and maintained its identity, Supervisor as Person and Therapist (B1/C1, seven items), whereas B2- and B3-clusters merged into a larger cluster (C2, 15 items: Supervisory Competence).

Reliability and concurrent validity

Internal consistency measures were uniformly high for the SE-SC instrument, and this applied to both the overall scale and the A1-clusters (ranging from 0.75 to 0.92; see Table 3). The results for test–retest reliability were appropriately high: for each of the three B-clusters, reliability correlations (Pearson's *r*, $N = 20$) were, respectively, 0.93, 0.81, and 0.87. The results indicated good agreement between measures of the SE-SC on the one hand and the other two instruments, the SWAI and the SES. Correlations between the SE-SC overall score and the SWAI were high, $r(69) = 0.75$, $p < .001$; as were correlations between SE-SC overall score and the SES, $r(49) = 0.86$, $p < .001$. Further, higher correlations were observed between subscales of concurrent instruments that were conceptually similar to each other. For instance, higher correlations between conceptually similar subscales: (a) Supervisor Openness, Caring and Support (Cluster A1) and the Rapport Scale of the SWAI *r* (70) = 0.82, $p < .001$; (b) Supervisor Openness, Caring and Support (Cluster A1) and the Interpersonal Sensitivity Subscale of the SSI, $r(50) = 0.87$, $p < .001$; (c) Goal-Directed Supervision (Cluster 4) and SSI-Task Orientation, $r(70) = 0.68$, $p < .001$; Restorative Competencies (A5) with SWAI-Rapport Subscale, *r* (70) = 0.80, $p < .001$; Restorative Competencies (A5) with SSI-Interpersonal Sensitivity, $r(51) = 0.84$, $p < .001$; Conversely, lower correlations were observed between conceptually dissimilar subscales, Reflective Practitioner Competencies (Cluster A6) and the SSI-Task Orientation, $r(51) = 0.15$, $p > .05$, and Restorative Competencies (Cluster A5), and the SSI-Task Orientation, $r(51) = 0.27$, $p > .05$ (Table 2).

Predictive validity

Importantly, the hypothesis that the SE-SC clusters would better predict overall SE-SC was supported. The results of the multiple regression analyses are presented in Table 3. Each of the three regressions was significant: for SE-SC, $F(6,137) = 116.48$, $p < .001$; for SWAI, $F(2,69) = 27.98$, $p < .001$; for SES, $F(3,50) = 23.39$, $p < .001$. However, the clusters from the SE-SC predicted a larger proportion of the variance (85%) when regressed on the overall SE-SC score, as compared to the subscales of the SSI (58%) and those of the SWAI (57%). Of the six clusters, four clusters including Openness, Caring and Support, Knowledge and Therapist Expertise, Supervision Planning and Management, and Reflective Practitioner Competencies were significant independent predictors of overall evaluations of supervision satisfaction and effectiveness ($p < .001$ in each instance). A fifth, Restorative Competencies approached statistical significance, $t(137) = 1.80$, $p < .10$. Finally, because our

overall score combined supervision satisfaction and the trainee's ratings of supervisor effectiveness, we repeated the regression analyses separately for each of the two subscales. The results confirmed the previous analyses with the clusters contributing to much larger proportions of the variance (77% to supervision satisfaction and 79% to supervisor effectiveness) than did the SSI (51% to supervision satisfaction and 50% to supervisor effectiveness) and the SWAI (48% to supervision satisfaction and 46% to supervisor effectiveness). Interestingly, some clusters predicted both satisfaction and effectiveness (Clusters 2, 3, and 6 from the SE-SC; client focus from the SWAI; and interpersonal sensitivity from the SSI), whereas supervisor caring and support (from the SE-SC) and rapport (from the SWAI) were strong predictors of supervision satisfaction but not of effectiveness.

Discussion

The current study aimed to design and validate the SE-SC scale, and to test a key prediction of competency approaches. An innovative cluster-analytic approach yielded six key competencies represented by six subscales (A-clusters). The overall score and the subscale scores derived from the instrument demonstrated excellent internal reliability, test–retest reliability, and concurrent validity. Importantly, the subscales of the SE-SC better predicted overall measures of SE-SC than subscales from the two other instruments: the SWAI and the SSI.

The current study makes a substantive contribution to the extant literature in several ways. First, the study documents, as far as we are aware, the first empirical data to support the basic tenet of competency-based approaches, namely, that a range of key competencies would better predict overall evaluations of supervision than generic processes. The development and validation of the new instrument is also of salience and has several advantages. (a) It provides a more comprehensive evaluation of supervision and of supervisory competence than most other published measures. Specifically, the overall score measures supervision satisfaction *and* effectiveness, and the 22 other items measure six key supervisor competencies, albeit as evaluated by the supervisee. (b) The instrument has excellent psychometric properties. (c) The instrument also provides a good index of supervisory alliance (Cluster A1, Openness, Caring and Support Correlates with SWAI-Rapport, $r(70) = 0.82$), but goes beyond the scope of the construct to measure supervisory competencies that span the normative (e.g., Supervisor Knowledge and Therapist Expertise, and Goal-Directed Supervision), formative (e.g., Reflective Practitioner Competencies), and restorative competencies (Restorative Competencies). (d) The instrument is better aligned with new theoretical developments in training and supervision, including competency-based approaches (Falender & Shafranske, 2007; Gonsalvez et al., 2002; Pilling & Roth, 2014), with the results supporting the importance of systematic formulation of a competency-development plan for supervision (e.g., Falender & Shafranske, 2012; Gonsalvez, 2014), and collaborative goal setting (e.g., O'Donovan, Halford, & Walters, 2011). The emphasis on tailoring supervision methods to align with skills training and competencies is also vindicated (e.g., Cluster A3) because this cluster contributed uniquely to overall measures of supervision

effectiveness independent of process variables, a position supported by other reports (Gonsalvez & McLeod, 2008; Kavanagh et al., 2003; Reiser, 2014).

Supervisory Rapport measured by the SWAI was also related to two relatively independent clusters, "Open, Caring and Supportive Supervisor" and "Restorative Competencies" (r = 0.82 and 0.80, respectively). A similar pattern was observed for interpersonal sensitivity (r = 0.87 and 0.84) and suggests that it would be useful to conceptualise them, as disparate competencies. In fact, the cluster analysis helps identify three independent competencies: a generic personable supervisor characteristic of being open, caring and supportive within the supervisor–supervisee dyad, a more specific restorative competency characterised by being sensitive to the supervisee's affective reactions including the supervisee's self-care needs, and a third competency that captures reflective practice competencies, including an incisive analyses of the therapist-client dynamics. It is also of interest that although the two latter competencies are somewhat related to each other, the generic relationship competency remains independent even at later stages of the cluster analyses. Put bluntly, supervisors who are open, caring and supportive as persons, may have neither the restorative skills to remain attuned to the supervisee's self-care needs and create a safe place in supervision, nor the competencies to enhance reflective practitioner skills among trainees and the set of skills necessary to foster insight into and effectively manage the interpersonal therapist-client dynamics (e.g., transference and counter transference). The results also supported the distinction between supervision satisfaction and effectiveness, with supervisor caring and support (SE-SC) and rapport (SWAI) predicting satisfaction but not effectiveness, underlining the assertion that an effective supervisor is more than a "nice caring person" (O'Donovan & Kavanagh, 2014).

In general, the competency clusters map well onto the preliminary thematic analyses conducted by Olds and Hawkins (2014). Specifically, Cluster A1, "Openness, Caring and Support" shares features with the "supervisory alliance" of these authors; Cluster A2, "Supervisor's Knowledge and Therapist Skills" maps well onto "Knowledge of the Profession;" Cluster A3, "Supervision Planning and Management" is akin to "Structuring Supervision;" and Cluster A6, "Insight into and Management of Therapist-Client Dynamics and Reflective Practitioner Competencies" encompasses Olds and Hawkin's "Reflective Practice." Similarly, several clusters are consistent with the results of the factor structure derived from the SRQ (Palomo et al., 2010), with Cluster 1 resembling "safe base;" Cluster 2 resembling "role model;" Cluster 3 resembling "structure;" and Cluster 6 resembling "reflective education."

Limitations and Future Directions

As stated earlier, supervisee evaluation of supervisory competence is an essential aspect of supervisory competence assessment, but remains, by itself, an insufficient measure. Ideally, evaluations by supervisees should be supplemented by assessments by informed peers and supervision experts. Further, subjective report of supervisees of what transpired in supervision is an incomplete if not poor report of supervisory competence. Specifically, pedagogically sound and ecologically valid assessment of supervisory competence should include a demonstration of key competencies by the supervisor (e.g., on a video tape) and reliable assessment of these competencies by credible assessors in the field (e.g., informed peers and supervision experts). Further, there is legitimacy in the argument that a comprehensive evaluation of supervisory competence should include evidence that this competence has flow-on effects that actually improve trainee competence (not just perceived improvements) and in demonstrated client outcomes. Consequently, although the new scale is of value to the evaluation of competence, comprehensive assessment should necessarily include additional assessments through objective methods, an issue that requires systematic attention in future supervision training and research. Despite the lack sophistication and other potential limitations, supervisee evaluation of supervisory competence constitutes a legitimate aspect because these perceptions inform attitudes, techniques, assessment, and process in supervision.

The 22 items of the current scale subjected to psychometric scrutiny, comprise most but not all important competencies. Notable omissions include supervisory competence in dealing with ethical, multicultural, and legal aspects, and also of adhering to best-practice guidelines in relation to providing accurate summative assessments in an interpersonally sensitive manner. The current version of the scale includes these important competencies, but there were limited data on these competencies to meaningfully contribute to the cluster analyses undertaken within the study. The promising findings augur well for future research initiatives: to determine if current results are replicated in samples across developmental stages and theoretical persuasions, to investigate whether the superior results favouring competency clusters over supervisory alliance would hold if compared against newer conceptualisations of the supervisory relationship (e.g., as measured by the SRQ), to extend these findings by including a more comprehensive array of competencies, and to compare ratings of supervisory competence derived from different groups including supervisees, peers and experts.

Finally, the study's sample whilst being representative of clinical psychology training programmes within Australia comprised a large percentage of supervisees who were novices (61% had less than 500 hr of practicum experience) and who preferred a CBT theoretical orientation (81%). Consequently, an examination of trends among supervisees across multiple theoretical orientations and developmental stages is warranted.

References

American Psychological Association. (2014). *Guidelines for clinical supervision for health service psychologists*. Retrieved from http://www.apa.org/about/policy/guidelines-supervision.pdf

Association of State and Provincial Psychology Boards. (2015). *Supervision guidelines for education and training leading to licensure as a health service provider*. Retrieved from http://www.asppb.org/

Barnett, J. E., Erickson Cornish, J. A., Goodyear, R. K., & Lichtenberg, J. W. (2007). Commentaries on the ethical and effective practices of clinical supervision. *Professional Psychology: Research & Practice*, 38, 268–275. doi:10.1037/0735-7028.38.3.268

Beinart, H. (2014). Building and sustaining the supervisory relationship. In C. E. Watkins Jr. & D. L. Milne (Eds.), *The Wiley international handbook of clinical supervision* (pp. 257–281). West Sussex, England: John Wiley & Sons Ltd.. doi:10.1002/9781118846360

Bernard, J. M., & Goodyear, R. K. (2014). Supervision models. In J. M. Bernard & R. K. Goodyear (Eds.), *Fundamentals of clinical supervision* (5th ed., pp. 21–62). Boston, MA: Pearson.

Edwards, D., Cooper, L., Burnard, P., Hanningan, B., Adams, J., Fothergill, A., & Coyle, D. (2005). Factors influencing the effectiveness of clinical supervision. *Journal of Psychiatric and Mental Health Nursing*, 12, 405–414.

Efstation, J. F., Patton, M. J., & Kardash, C. M. (1990). Measuring the working alliance in counselor supervision. *Journal of Counseling Psychology*, 37(3), 322–329. doi:10.1037/0022-0167.37.3.322

Falender, C. A., & Shafranske, E. P. (2004). *Clinical supervision: A competency-based approach*. Washington, DC: American Psychological Association.

Falender, C. A., & Shafranske, E. P. (2007). Competence in competency-based supervision practice: Construct and application. *Professional Psychology: Research and Practice*, 38(3), 232–240. doi:10.1037/0735-7028.38.3.232

Falender, C. A., & Shafranske, E. P. (2012). The importance of competency-based supervision and training in the 21st century: Why bother? *Journal of Contemporary Psychotherapy*, 42, 5–13.

Fouad, N. A., Grus, C. L., Hatcher, R. L., Kaslow, N. J., Hutchings, P. S., Madson, M. B., … Crossman, R. E. (2009). Competency benchmarks: A model for understanding and measuring competence in professional psychology across training levels. *Training and Education in Professional Psychology*, 3(4 Suppl), S5–S26. doi:10.1037/a0015832

Friedlander, M. L., & Ward, L. G. (1984). Development and validation of the Supervisory Styles Inventory. *Journal of Counseling Psychology*, 31(4), 541–557. doi:10.1037/0022-0167.31.4.541

Gonsalvez, C. J. (2014). Establishing supervision goals and formalizing a supervision agreement. In C. E. Watkins Jr. & D. L. Milne (Eds.), *The Wiley international handbook of clinical supervision* (pp. 282–307). West Sussex, England: John Wiley & Sons Ltd.. doi:10.1002/9781118846360.ch12

Gonsalvez, C. J., & Calvert, F. L. (2014). Competency-based models of supervision: Principles and applications, promises and challenges. *Australian Psychologist*, 49(4), 200–208. doi:10.1111/ap.12055

Gonsalvez, C. J., Deane, F. P., Blackman, R., Matthias, M., Knight, R., Nasstasia, Y., … Bliokas, V. (2015). The hierarchical clustering of clinical psychology practicum competencies: A multisite study of supervisor ratings. *Clinical Psychology: Science and Practice*, 22, 390–403. doi:10.1111/cpsp.12123

Gonsalvez, C. J., & McLeod, H. J. (2008). Toward the science-informed practice of clinical supervision: The Australian Context. *Australian Psychologist*, 43, 79–87.

Gonsalvez, C. J., Oades, L., & Freestone, J. (2002). The objectives approach to clinical supervision: Towards integration and empirical evaluation. *Australian Psychologist*, 37(1), 68–77. doi:10.1080/00050060210001706706

Holloway, E. L. (2012). Professional competence in supervision. In J. N. Fuertes, A. Spokane & E. L. Holloway (Eds.), *Specialty competencies in counselling psychology* (pp. 165–181). New York, NY, US: Oxford University Press.

Kavanagh, D. J., O'Halloran, P., Manicavasagar, V., Clark, D., Piatkowska, O., Tennant, C., & Rosen, A. (1997). The Family Attitude Scale: Reliability and validity of a new scale for measuring the emotional climate of families. *Psychiatry Research*, 70, 185–195. doi:10.1016/s0165-1781(97)00033-4

Kavanagh, D. J., Spence, S. H., Strong, J., Wilson, J., Sturk, H., & Crow, N. (2003). Supervision practices in allied mental health: Relationships of supervision characteristics to perceived impact and job satisfaction. *Mental Health Services Research*, 5, 187–195.

Lehrman-Waterman, D., & Ladany, N. (2001). Development and validation of the Evaluation Process Within Supervision Inventory. *Journal of Counseling Psychology*, 48(2), 168–177. doi:10.1037/0022-0167.48.2.168

Livni, D., Crowe, T. P., & Gonsalvez, C. J. (2012). Effects of supervision modality and intensity on alliance and outcomes for the supervisee. *Rehabilitation Psychology*, 57(2), 178. doi:10.1037/a0027452

Llewelyn, S. P. (1988). Psychological therapy as viewed by clients and therapists. *British Journal of Clinical Psychology*, 27(3), 223–237. doi:10.1111/j.2044-8260.1988.tb00779.x

McNeill, B. W., Stoltenberg, C. D., & Pierce, R. A. (1985). Supervisees' perceptions of their development: a test of the counselor complexity model. *Journal of Counseling Psychology*, 32, 630–633.

Milne, D. L., & Reiser, R. (2014). SAGE, a scale for rating competence in CBT supervision. In C. E. Watkins Jr. & D. L. Milne (Eds.), *The Wiley international handbook of clinical supervision* (pp. 403–415). West Sussex, England: John Wiley & Sons Ltd.. doi:10.1002/9781118846360.ch18

Milne, D. L., & Watkins, C. E. (2014). Defining and understanding clinical supervision. In C. E. Watkins Jr. & D. L. Milne (Eds.), *The Wiley international handbook of clinical supervision* (pp. 1–19). West Sussex, England: John Wiley & Sons Ltd.. doi:10.1002/9781118846360.ch1

O'Donovan, A., & Kavanagh, D. J. (2014). Measuring competence in supervisees and supervisors. In C. E. Watkins Jr. & D. L. Milne (Eds.), *The Wiley international handbook of clinical supervision* (pp. 458–467). West Sussex, England: John Wiley & Sons Ltd.. doi:10.1002/9781118846360.ch22

O'Donovan, A., Halford, W. K., & Walters, B. (2011). Toward best practice supervision of clinical psychology trainees. *Australian Psychologist*, 46, 101–112.

Olds, K., & Hawkins, R. (2014). Precursors to measuring outcomes in clinical supervision: A thematic analysis. *Training and Education in Professional Psychology*, 8(3), 158–164.

Olk, M. E., & Friedlander, M. L. (1992). Trainees' experiences of role conflict and role ambiguity in supervisory relationships. *Journal of Counseling Psychology*, 39(3), 389. doi:10.1037/0022-0167.39.3.389

Orlinsky, D. E., & Rønnestad, M. H. (2005). *How psychotherapists develop: A study of therapeutic work and professional growth*. Washington, DC: American Psychological Association.

Palomo, M., Beinart, H., & Cooper, M. J. (2010). Development and validation of the Supervisory Relationship Questionnaire (SRQ) in UK trainee clinical psychologists. *British Journal of Clinical Psychology*, 49(2), 131–149. doi:10.1348/014466509x441033

Pearce, N., Beinart, H., Clohessy, S., & Cooper, M. (2013). Development and validation of the Supervisory Relationship Measure: A self-report questionnaire for use with supervisors. *The British Journal of Clinical Psychology*, 52(3), 249–268. doi:10.1111/bjc.12012

Pilling, S., & Roth, A. D. (2014). The competent clinical supervisor. In C. E. Watkins Jr. & D. L. Milne (Eds.), *The Wiley international handbook of clinical supervision* (pp. 20–37). West Sussex, England: John Wiley & Sons Ltd.. doi:10.1002/9781118846360.ch2

Psychology Board of Australia. (2013). *Guidelines for supervisors and supervisión training providers*. Retrieved from Psychology Board of Australia website: http://www.psychologyboard.gov.au/Standards-and-Guidelines/Codes-Guidelines-Policies.aspx

Reiser, R. P. (2014). Supervising cognitive and behavioral therapies. In C. E. Watkins Jr. & D. L. Milne (Eds.), *The Wiley international handbook of clinical supervision* (pp. 493–517). West Sussex, England: John Wiley & Sons Ltd.. doi:10.1002/9781118846360

Rodolfa, E., Greenberg, S., Hunsley, J., Smith-Zoeller, M., Cox, D., Sammons, M., … Spivak, H. (2013). A competency model for the practice of psychology. *Training and Education in Professional Psychology*, 7, 71–83. doi:10.1037/a0032415

Rønnestad, H., & Lundquist, K. (2009). *Brief supervisory alliance questionnaire*. Oslo, Norway: Department of Psychology, University of Oslo.

Wheeler, S., & Barkham, M. (2014). A core evaluation battery for supervision. In C. E. Watkins Jr. & D. L. Milne (Eds.), *The Wiley international handbook of clinical supervision* (pp. 365–385). West Sussex, England: John Wiley & Sons Ltd.. doi:10.1002/9781118846360