

Assessing Psychologists in Practice: Lessons From the Health Professions Using Multisource Feedback

Jac J. W. Andrews, Claudio Violato, Ahmed Al Ansari, Tyrone Donnon, and Gia Pugliese
The University of Calgary

The aim of the present study was to conduct a systematic literature review of multisource feedback (MSF) instruments and to summarize the evidence of feasibility, reliability, generalizability, validity, and other psychometric characteristics of the instruments. Accordingly, we conducted a systematic literature review for English-language studies published from 1975 to 2012 using the following databases: MEDLINE, EMBASE, CINAHL, PubMed, and PsycINFO. The following terms were used in the search: *multisource feedback*, *360-degree evaluation*, and *assessment of medical professionalism*. Forty-eight studies conducted in Canada, the United States, the United Kingdom, Netherlands, China, and elsewhere met the inclusion criteria. The results indicate that MSF has adequate evidence of validity, reliability, and feasibility for providing health practitioners with quality improvement data (both formative and summative assessment) as part of an overall strategy of maintaining competence and certification. Professional psychology has not adopted MSF as a systematic competence-based method for evaluating, maintaining, and assuring competent practice of psychology and instead relies on self-assessment as the primary quality assurance approach for its public accountability. We make recommendations to adopt an MSF system of competence-based assessment of practicing psychologists by regulatory and licensing authorities in Canada and the United States.

Keywords: multisource feedback, 360-degree evaluation, professional competence, psychologist assessment, formative assessment

A major professional issue for psychology is the monitoring and assessment of psychologists in professional practice. *Multisource feedback* (MSF), sometimes referred to as *360-degree assessment*, has become increasingly common in health care

(Dupee, Ernst, & Caslin, 2011) as a way to monitor and assess the performance of professionals in practice (London & Smith, 1995). The use of multiple sources of information to evaluate performance has also been considered useful in business (Sala & Dwight, 2002). For example, it has been widely used in industry as a way of providing feedback to employees in order to improve workplace performance and guide self-directed learning (Sala & Dwight, 2002) where supervisors, peers, and occasionally clients provide MSF. Often, MSF is used in settings in which the staff person works in a team and/or cannot be directly and easily supervised by managers (Church, 1997). Similarly, in psychology, multiple sources of input can be valuable when evaluating the same dimension such as in the case of a trainee's development as a therapist (Falender & Shafranske, 2004a). Cone (2001), for example, suggested the use of measurement of multiple aspects of a client's constructs to gauge their relative effectiveness. More generally, Falender and Shafranske (2004b) believe the use of MSF provides an interesting analysis of training performance within the context of clinical supervision and psychology training. Although MSF has not been used as a way to evaluate the performance of practicing psychologists, Andrews and Violato (2010) presented information about how MSF instruments could be constructed and used as a competence-monitoring system for practicing school psychologists and what these instruments could look like. In the present article, we propose the use of MSF for the assessment of psychologists in practice. We reviewed developments and research findings in health care and psychology to provide some considerations and directions for MSF in professional psychology.

JAC J. W. ANDREWS, PhD, is Professor, Chair, and Director of Training of the School and Applied Child Psychology Program in the Faculty of Education, University of Calgary. His research areas include child psychopathology, psychological assessment and evaluation, and psycho-educational intervention.

CLAUDIO VIOLATO, PhD, is Professor with the Medical Education and Research Unit, Department of Community Health Sciences, Faculty of Medicine, University of Calgary. His research areas include psychometrics, multisource feedback, and assessment and evaluation.

AHMED AL ANSARI, MBBCh, MRCSI, MHPE, PhD, is Director of Training and Development, Department of Medical Education, Bahrain Defense Force Hospital. His research areas include multisource feedback and teaching and learning in medical education.

TYRONE DONNON, PhD, is Associate Professor with the Medical Education and Research Unit, Department of Community Health Sciences, Faculty of Medicine, University of Calgary. His research areas include assessment and evaluation in medical education.

GIA PUGLIESE, MSc, is a PhD student in the School and Applied Child Psychology Program in the Faculty of Education at the University of Calgary. Her research areas include social-emotional processing and bullying.

CORRESPONDENCE CONCERNING THIS ARTICLE should be addressed to Jac J. W. Andrews, School and Applied Child Psychology, Faculty of Education, University of Calgary, 2500 University Drive N.W., Calgary, Alberta, Canada, T2N 1N4. E-mail: jandrews@ucalgary.ca

Professional Competency and Assessment of Psychologists: An Overview

Psychologists are generally thought to be competent if they have the knowledge, skills, attitudes, and professionalism for practice (e.g., assessment, intervention, consultation; Sharpless & Barber, 2009). According to Epstein and Hundert (2002), *competence* is “the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice for the benefit of the individual and community being served” and is dependent on “habits of mind, including attentiveness, critical curiosity, self-awareness, and presence” (p. 227). According to the Continuing Competence Program for Psychologists Practicing in Nova Scotia (http://www.nsbep.org/downloads/Continuing_Competence) in Canada, competence for psychologists involves the interaction of four major components: knowledge (of a range of professional issues), skill (ability to apply knowledge), judgment (when and where to apply skills), and diligence (consistent application of knowledge, skills, and judgment). Moreover, competence is the *potential* for appropriate professional practice distinct from *performance* in daily practice, which is situation specific, and should be observable, measurable, and developmental.

The conceptualization and measurement of performance and competence in the health professions have been evolving processes over the years. Performance and competence are fluid states whereby practice standards are developed throughout one’s professional education, training, and experience. Health professions (including psychology) have established foundations and criteria for determining professional competence including ethical principles and standards of practice (e.g., American Psychological Association, 2002; Canadian Psychological Association, 2000), registration and licensure requirements through professional credentialing bodies (e.g., Association of State and Provincial Psychology Boards, which is the alliance of state, provincial, and territorial agencies responsible for the licensure and certification as well as the maintenance of competence of psychologists throughout the United States and Canada; <http://www.asppb.net>), education and training (e.g., American Psychological Association- and Canadian Psychological Association-accredited programs), and specialty board certification processes and procedures (e.g., American Board of Clinical Psychology, American Board of Counseling Psychology, American Board of School Psychology).

Over the past several decades, psychology has moved toward considering, developing, and implementing valid, reliable, and feasible assessment of performance and competence in order for the profession to improve its public accountability and as a way to provide both formative and summative feedback to psychologists to improve, enrich, or confirm their performance consistent with expected outcomes (Carraccio, Wolfsthal, Englander, Ferentz, & Martin, 2002; Roberts, Borden, Christiansen, & Lopez, 2005). Professional competence in health care is expected by patients and clients, as well as policymakers and regulators (Hoge et al., 2005). Accordingly, the assessment of competence and performance protects the public and encourages professional development (Kaslow, 2004). Although there has been recent support for the systematic assessment of competence in professional psychology (Kaslow, Rubin, Bebeau, et al., 2007), there is still a lack of consensus in the field about how to define and measure the

competence of professional psychologists in practice (Fantuzzo, Sisemore, & Spradlin, 1983; Kaslow, 2004; Shaw & Dobson, 1988). Moreover, licensing and regulatory boards as well as graduate training programs also vary in their competence-based assessment approach (Roberts et al., 2005). Although licensing and regulatory boards in Canada and the United States have not adopted agreed-on principles and objectives to assess the practice of psychologists, there have been recommendations made in psychology for the assessment of competence.

Kaslow et al. (2009), for example, presented a competence assessment toolkit for professional psychology in which various assessment instruments along with information regarding their implementation, reliability, validity, and fidelity were reported. They reviewed several methods deemed appropriate for assessing the foundational competencies (e.g., professionalism) and functional competencies (e.g., assessment) of psychologists and for measuring competence for education, training, and professional development. The methods described and reviewed included annual/rotation performance reviews, case presentation reviews, client/patient process and outcome data, competence evaluation rating forms, consumer surveys, live or recorded performance ratings, objective structured clinical examinations, portfolios, record reviews, self-assessment, simulations/role plays, standardized client/patient interviews, structured oral examinations, written examinations, and 360-degree evaluations.

In addition to providing descriptions and reviews of these assessment methods, Kaslow et al. (2009) noted the following considerations for choosing and using one or more of the methods: (a) types of competencies to be assessed; (b) psychometric properties of the methods; (c) feasibility, fidelity, strengths, and challenges of the methods; (d) appropriateness of the coding, scoring, and interpretation of the data; (e) who would serve as the evaluator(s); and (f) the context of the assessment. Moreover, they underscored that each method has strengths and challenges and that no single method can adequately assess the breadth and depth of all competencies. In addition, methods need to be cost-effective and congruent for different stages of professional development as well as for the purpose of assessment (e.g., formative vs. summative evaluation).

MSF or 360-degree evaluation involves the creation of instruments to assess and compare observable behaviors (performance) of professionals. Typically, this involves a self-assessment and an assessment from others (e.g., colleagues, coworkers, clients). An important issue with the use of a questionnaire-based MSF is whether or not it can provide adequate reliability and evidence of validity, demonstrate utility or feasibility, and be applicable to professional psychology. In the following section, we review the use of MSF in the health field addressing many of the issues identified by Kaslow et al. (2009) for its use as an approach for competence-based assessment of psychologists in practice.

The aim of the present study was to conduct a systematic literature review of MSF instruments and to summarize the evidence of feasibility, reliability, generalizability, validity, and other psychometric characteristics of the instruments. Based on the results, we wished to make recommendations about implementing an MSF procedure for assessing psychologists in professional practice.

Method

Data Sources

A systematic literature review was conducted searching for English-language studies published from 1975 to 2012 using the following databases: MEDLINE, EMBASE, CINAHL, PubMed, and PsycINFO. The following terms were used in the search: *multisource feedback*, *360-degree evaluation*, and *assessment of medical professionalism*.

Studies were included if they (a) described the instrument design; (b) identified factors measured by the instrument; (c) were used with medical/health professionals; (d) included information about feasibility, reliability, generalizability, and validity of the MSF; and (e) were published in English. We excluded studies if (a) they were used with nonhealth/medical professionals; (b) the instruments were not adequately described; (c) they provided only general description and information about MSF; (d) they were published in languages other than English; (e) they provided inadequate information about sample size, study design, data analyses, and psychometric results; and (f) they provided only changes in raters' performance after feedback.

Data Extraction

Each article in this study was evaluated for inclusion by the third and fourth authors independently based on the title and abstract. Any disagreements were solved by retrieving the full article and having it reviewed by the second author. Based on discussions among the three coders, we achieved 100% agreement for all the included studies.

The initial search yielded 1,061 articles. Of these, 743 articles were excluded based on the title, a further 219 articles were excluded based on the abstract, and another 51 were eliminated after reading the full articles. Finally, we agreed on 48 articles to be included. The detailed characteristics of those studies are described in Table 1.

Implementation Process of Surveys and Rater Selection Across Studies

Candidates selected for MSF (e.g., family physicians, psychiatrists, occupational therapists) are typically provided self-report questionnaires, patient questionnaires (typically 25 questionnaires), coworker questionnaires (typically eight to 10 questionnaires), and colleague questionnaires (typically eight to 10 questionnaires) from the researcher (MSF office). Generally, the candidates or designates (e.g., receptionist) administer questionnaires to the patients, which are then collected and returned to the researcher or MSF office. The candidates distribute questionnaires to his or her coworkers and colleagues to complete, and these raters return the questionnaires to the researcher or MSF office in self-addressed, prestamped envelopes. The candidates select raters (e.g., colleagues) who are known to the candidate and who have similar practice (e.g., surgeons for surgeons, etc.).

Results

As shown in Table 1, of 1,061 articles, 48 met the inclusion criteria and 1,013 were excluded. Most studies ($n = 42, 87\%$) were

published between 2000 and 2012, another four (8%) were published between 1990 and 1999, and two (5%) were published before 1989. Sixteen studies (33%) were conducted in Canada, 14 (29%) in the United States, 13 (27%) in the United Kingdom, one (2%) in the Netherlands, one (2%) in Australia, one (2%) in China, one in Denmark (2%), and one (2%) in Taiwan. MSF has been used in almost all the medical specialties but primarily with family physicians and pediatricians. The specific medical and health care specialties were family physicians ($n = 8, 17\%$), obstetrics and gynecology ($n = 3, 6\%$), pediatrics ($n = 5, 10\%$), radiology ($n = 1, 2\%$), anesthesia ($n = 3, 6\%$), histopathology ($n = 1, 2\%$), various specialties ($n = 11, 23\%$), psychiatry ($n = 4, 8\%$), occupational therapists ($n = 1, 2\%$), medical radiation technologists ($n = 1, 2\%$), pathologists ($n = 1, 2\%$), emergency medicine ($n = 1, 2\%$), surgery ($n = 4, 8\%$), urology ($n = 1, 2\%$) and internal medicine ($n = 3, 6\%$).

In total, 32 of the studies (66%) used multiple questionnaires to assess candidates, whereas 16 studies (34%) used a single questionnaire. The domains assessed by MSF included professionalism (reported in 39 [81%] of total studies), clinical competence ($n = 37, 77\%$), communication ($n = 35, 72\%$), case management ($n = 20, 42\%$), interpersonal relations ($n = 28, 58\%$), and overall assessment ($n = 4, 8\%$).

Description of MSF Instruments

Information about different types of MSF instruments is provided in Table 1. Thirty-two (66%) studies used multiple surveys in MSF in assessing participants. Most of those studies were conducted in Canada and the United States (see Table 1). The remaining 16 (34%) studies used single survey in MSF (five used the Sheffield Peer Ratings Assessment Tool, which consists of 24 items). Those studies were mainly conducted in the United Kingdom. The other studies used a single survey with different number of items in each specific survey (see Table 1).

Feasibility

Thirty-nine studies (81%) addressed feasibility. Most of the studies focused on the response rates for the surveys (see Studies 3–7, 9–13, 15, 17–25, 27–29, 31–36, and 38–48 in Table 1). In general, feasibility was classified as good but the results varied across the studies. Davis (2002), DiMatteo and DiNicola (1981), Lockyer, Violato, Fidler, and Alakija (2009), and Violato, Lockyer, and Fidler (2006) reported response rates of 100% across some of the questionnaires (see Studies 25, 17, 10, 32, respectively, in Table 1). Other researchers reported response rates ranging from 56.2% to 95.1%. One of the studies reported time needed per doctor to complete the questionnaire as a measurement of responses (Lockyer, Violato, & Fidler, 2006a). The authors found that the time required to complete each questionnaire is 6 min, which is considered feasible. In another study, Wood et al. (2006) reported that more than 90% of the 360-degree forms were completed by all raters in less than 1 min, which further supports the feasibility of using such methods.

Internal Structure, Reliability, and Generalizability

Several reliability coefficients are reported in Table 1. Internal consistency reliability (Cronbach's alpha) was reported as $\geq .90$ in

Table 1
Design, Sample, Specialty, and Psychometric Characteristics of Multisource Feedback Studies

Study; country	Specialty (sample size)	Instrument	Raters	Domains assessed	Reliability coefficients	Feasibility	Generalizability	Highest validity evidence
1. Violato, Lockyer, & Fidler (2008b); Canada	Family physicians (250)	Medical colleagues, coworkers, self	25 patients, 8 coworkers, 8 medical colleagues	Professional, clinical competence, communication, manager	$\alpha = .96-.98$	Response rates = 92.75–96.9%	$Ep^2 = .78-.83$	Construct validity
2. Wood et al. (2006); United Kingdom	Obstetricians and gynecologists (180)	MSF: 4 items	8 raters (peers and staff mixed)	Manager, interpersonal relationship	8 raters; ICC = .80			Construct validity
3. Archer, McGraw, & Davies (2010); United Kingdom	Pediatrics (577)	MSF (SPRAT): 24 items	8–12 mixed raters	Professional, clinical competence	8 raters; ICC = .80	Response rate = 83%		Construct validity
4. Wood et al. (2004); United States	Radiology (7)	MSF: 10 items	3 mixed raters	Professional, communication	Residents; ICC = .85–.87	Response rate = 100%		Content validity
5. Lockyer, Violato, & Fidler (2007); Canada	Family physicians (250)	MSF: 31 items	1 rater (self-rating)	Professional, clinical competence, communication, manager	$\alpha = .74-.98$	Response rate = 100%		Construct validity
6. Meng, David, & Rita (2009); United States	Anesthesia (15)	MSF: 31 items	Mean = 22 mixed raters	Professional, communication	ICC = .82–.87	Response rate = 88%		Content validity
7. Davies et al. (2008); United Kingdom	Histopathology (92)	MSF (PAT-SPRAT): 24 items	10 raters	Professional, clinical competence, manager, interpersonal	With 8 raters, ICC > .70	Response rate = 92%		Construct validity, factor analysis
8. Brinkman et al. (2007); United States	Pediatric (36)	MSF: parents = 10 items; nurses = 14 items	Parents and nurses	Professional, clinical competence	Parent $\alpha = .95$; nurse $\alpha = .95$			Construct validity, between-groups differences
9. Archer, Norcini, Southgate, Heard, & Davies (2008); United Kingdom	Various specialties (553)	MSF (Mini-PAT): 16 items	6–8 raters	Professional	Mini-PAT $\alpha = .98$; ICC > .70	Response rate = 67%		Construct validity, factor analysis, between-groups differences
10. Davis (2002); United States	Obstetricians and gynecologists	MSF: 16 items	15 peers, 14 attending doctors, 6 nursing and self	clinical, interpersonal overall	ICC = .34–.84	Response rates = 92.9–100%		Criterion-related validity, Pearson's r
11. Violato, Marini, Toews, Lockyer, & Fidler (1997); Canada	Family physicians and specialists (28)	Peer, self, patients, coworker, consultant, referring physician	Patients, physician, coworkers, peer, self	Professional, clinical competence, communication, manager, interpersonal	$\alpha = .89-.95$	Response rates = 73.5–87.4%	$Ep^2 = .82-.84$	Construct validity, factor analysis, between-groups differences
12. Lockyer & Violato (2004); Canada	Psychiatrists, pediatricians, and internal medicine (304)	MSF: 36 items	8 peer raters	Professional, clinical competence, communication	Peer $\alpha = .98$	Peer response rate = 94.8%	$Ep^2 = .70-.82$	Content validity

Table 1 (continued)

Study; country	Specialty (sample size)	Instrument	Raters	Domains assessed	Reliability coefficients	Feasibility	Generalizability	Highest validity evidence
13. Violato, Worstfold, & Polgar (2009); Canada	Occupational therapists (238)	MSF: coworkers, self, and clients	15 clients and 12 coworkers	Professional, clinical competence, communication, manager	$\alpha = .97-.93$	Response rates = 80.7–95.2%		Construct validity, factor analysis, between-groups differences
14. Lipner et al. (2002); United States	Various specialties (356)	MSF: patient, peer	25 patients, 10 peer, 2 self-ratings	Professional, communication	Between items ranging from .43 to .62		$Ep^2 = .61-.67$	Construct validity, factor analysis, between-groups differences
15. Violato & Saberton (2006); Canada	Medical radiation technologists (307)	MSF: patients, coworker, colleagues, self	25 patients, 6 colleagues, 6 coworker, and self-ratings	Professional, clinical competence, manager, interpersonal	$\alpha = .96-.97$	Response rates = 91–100%	$Ep^2 = .78-.81$	Construct validity, factor analysis, between-groups differences
16. Violato & Lockyer (2006); Canada	Psychiatrists, pediatricians, and internal medicine (304)	MSF: 38 items	Patients, coworkers, psychiatrist colleagues, and self	Professional, clinical competence, communication	$\alpha [\mu r] .90$		$Ep^2 = .78-.82$	Construct validity, factor analysis, between-groups differences
17. Lockyer et al. (2009); Canada	Pathologists/laboratory medicine (101)	MSF: patients, coworkers, colleagues, self	8 coworkers, 8 peers, and 8 referring physicians	Professional, clinical competence, communication, manager, interpersonal	$\alpha = .95-.98$	Response rates = 91–100%	$Ep^2 = .78-.81$	Construct validity, factor analysis, between-groups differences
18. Violato et al. (2008a); Canada	Psychiatrists (101)	MSF: patients, coworkers, colleagues, self	30 patients, 8 medical colleagues, 8 coworkers, and self-assessment	Professional, clinical competence, communication, manager, interpersonal	$\alpha = .96-.98$	Response rates = 92.1–100%	$Ep^2 = .78-.82$	Construct validity, factor analysis, between-groups differences
19. Lockyer et al. (2006a); Canada	Emergency medicine (187)	MSF: patients, coworkers, colleagues, self	25 patients, 8 medical colleagues, 8 coworkers, and self-assessment	Professional, clinical competence, communication, manager, interpersonal	$\alpha = .68-.97$	Response rates = 93–96%	$Ep^2 = .68-.85$	Construct validity, factor analysis, between-groups differences
20. Lockyer et al. (2006b); Canada	Anesthesiologists (197)	MSF: patients, coworkers, colleagues, self	30 patients, 8 medical colleagues, 8 coworkers, and self-assessment	Professional, clinical competence, communication, manager, interpersonal relationship	$\alpha = .93-.97$	Response rates = 56.2–95.1	$Ep^2 = .56-.69$	Construct validity, factor analysis, between-groups differences
21. Violato et al. (2003); Canada	Surgery (252)	MSF: patients, coworkers, colleagues, self	25 patients, 8 medical colleagues, 8 coworkers, and self-assessment	Professional, clinical competence, manager, interpersonal	$\alpha = .93-.97$	Response rates = 83.2–96.5%, medical	$Ep^2 > .70$ for 8 assessors and 25 patients	Construct validity, factor analysis, between-groups differences
22. Lockyer, Violato, & Fidler (2003); Canada	Surgery (153)	MSF: patients, coworkers, colleagues, self	Self-evaluation	Professional, communication	$\alpha = .78-.93$	Response rate = 76.5%		Construct validity, factor analysis (table continues)

Table 1 (continued)

Study; country	Specialty (sample size)	Instrument	Raters	Domains assessed	Reliability coefficients	Feasibility	Generalizability	Highest validity evidence
23. Allerup et al. (2007); Denmark	Internal medicine (42)	MSF: 15 items	1 secretary, 4 nurses, 5 senior doctors	Professional, clinical competence, communication, manager	$\alpha = .46-.89$	Response rate = 95.4%		Criterion-related, Pearson's r , between-groups differences
24. Hall et al. (1999); Canada	Family physicians (308)	Self; peer, patients, consultants, coworkers	25 patients, 8 colleagues, 6 coworkers, 6 peers, 6 referrals, self	Professional, clinical competence, manager	$\alpha = .93-.95$	Response rates = 79.7–88.67%	$Ep^2 = .82$	Construct validity, factor analysis, between-groups differences
25. Violato et al. (2006); Canada	Pediatricians (100)	MSF: patients, coworkers, colleagues, self	8 medical colleagues, 8 coworker, 25 patients, self	Professional, clinical competence, communication, manager, interpersonal	$\alpha = .95-.99$	Response rates = 93.6–100%	$Ep^2 = .78-.85$	Construct validity, factor analysis, between-groups differences
26. Archer & McAvoy (2011); United Kingdom	Various specialties (68)	SPRAT: 24 items	11 coworkers, 23 patients	Professional, clinical competence, communication				Construct validity, factor analysis, between-groups differences
27. Wennich, Carlene, Giles, & Ramsey (1993); United States	Internal medicine (318)	MSF: 13 items	12 nurses	Professional, clinical competence, communication, interpersonal overall		Response rate = 68.2%	Nurses $Ep^2 = .70$	Construct validity, factor analysis
28. Lelliott et al. (2008); United Kingdom	Consultant psychiatrists (347)	MSF: self, colleagues, patients	Self, 15 colleagues, 30 patients	Professional, communication, interpersonal	$\alpha = .97-.98$	Colleagues $M = 12.7$; patients $M = 19.2$	$Ep^2 = .75$	Construct validity, factor analysis, between-groups differences
29. Archer et al. (2005); United Kingdom	Pediatricians (112)	MSF (SPRAT): 24 items	10 mixed (doctors, nurses)	Professional, clinical competence, communication	Four raters ICC > .70	25 min per doctor		Construct validity, factor analysis
30. Campbell, Richards, Greco, Narayanan, & Brearley (2008); United Kingdom	Various specialties (291)	MSF: patients = 18 items, colleagues = 25 items	12 medical colleagues, 30 patients	Professional, clinical competence, communication, manager	$\alpha = .89-.92$		$Ep^2 = .75-.76$	Construct validity, factor analysis
31. Ramsey et al. (1993); United States	Physicians (313)	MSF: 11 items	11 (physicians and nurses)	Professional, clinical competence, overall		Response rate = 98.7%	7–11 physicians $Ep^2 = .70$	Construct validity, factor analysis
32. DiMatteo & DiNicola (1981); United States	Various specialties (141)	MSF: physician = 9 items, self = 8 items, patients = 3 items	6 physicians, 10 house staff, 16 patients	Clinical skills, interpersonal relationship	$\alpha = .40-.93$	Response rate = 100%		Construct validity, factor analysis
33. Noonan, Monagle, & Castanelli (2011); Australia	Consultant anesthetists (60)	Medical colleagues and coworkers = 49 items		Professional, clinical competence, communication, manager, interpersonal		Annual staff appraisal in Monash Medical Centre in Australia		Construct validity, factor analysis

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 1 (continued)

Study; country	Specialty (sample size)	Instrument	Raters	Domains assessed	Reliability coefficients	Feasibility	Generalizability	Highest validity evidence
34. Campbell, Narayanan, Burford, & Greco (2010); United Kingdom	General practice (179)	Colleague Feedback Evaluation Tool (CFET) and Doctor's Interpersonal Skills Questionnaire (DISQ) = CFEP360; colleagues = 18 items, patients = 12 items	1 attending doctor, 2 colleagues, 3 coworkers, 2 office staff, 7 patients, self	Professional, clinical competence, communication, manager, interpersonal	$\alpha = .84-.95$	Screen for underperformance	$Ep^2 = .80-.81$	Construct validity, factor analysis
35. Qu, Zhao, & Sun (2012); China	Various specialties (258)	Attending doctors = 21 items, colleagues = 21 items, coworkers = 21 items, office staff = 5 items, patients = 25 items, self = 21 items	1 attending doctor, 2 colleagues, 3 coworkers, 2 office staff, 7 patients, self	Professional, clinical communication, competence, manager, interpersonal	$\alpha = .90-.93$	MSF program by China Medical Board for resident physicians in China		Construct validity, factor analysis, between-groups differences
36. Warm, Schauer, Revis, & Boer (2010); United States	Internal medicine (22)	Colleagues, coworkers, attending doctors = 4 items, self = 45 items, patients = 10 items	1 attending doctor, 2 colleagues, 3 coworkers, 2 office staff, 7 patients, self	Professional, communication, clinical competence, manager, interpersonal	$\alpha = .84-.89$	MSF as mandatory for internal medicine residents		Construct validity, factor analysis, between-groups differences
37. Lockyer, Blackmore, et al. (2006); Canada	General practice (37)	PAR: self = 21 items, colleagues = 22 items, coworkers = 12 items, patients = 13 items	8 colleagues, 6 coworkers, 9 patients, 25 self	Professional, clinical competence	$\alpha = .88-.97$	Response rates = 88-100%	$Ep^2 = .67-.71$	Construct validity, factor analysis, between-groups differences
38. Pollock, Donnelly, Plymale, Stewart, & Vasconez (2007); United States	Plastic surgery (6)	Not applicable (NA); colleagues + coworkers = 60 items	12 colleagues, 28 coworkers	Professional, communication, clinical competence, manager, interpersonal		Evaluates resident performance using accessible colleagues and coworkers		Criterion-related: r between colleague and coworker
39. Yang et al. (2011); Taiwan	Various specialties (245)	NA; colleagues = 12 items	5 colleagues	Professional, clinical competence	Colleague $\alpha = .86$	Feasible to evaluate Year 1 residents	5 colleague $Ep^2 = .80$	Construct validity, factor analysis & between-groups differences
40. Overeem et al. (2012); Netherlands	Various specialties (146)	NA; colleagues = 33 items, coworkers = 22 items, patients = 18 items	7 colleagues, 7 coworkers, 15 patients	Professional, communication, clinical competence, manager, interpersonal	Coworker $\alpha = .94-.95$	MSF system in 23 hospitals		Construct validity, factor analysis, between-groups differences

(table continues)

Table 1 (continued)

Study; country	Specialty (sample size)	Instrument	Raters	Domains assessed	Reliability coefficients	Feasibility	Generalizability	Highest validity evidence
41. Chandler et al. (2010); United States	Pediatrics (66)	NA; colleagues = 10 items, coworkers = 10 items, patients = 10 items, self = 10 items	3 colleagues, 7 coworkers, 11 patients, 1 self	Professional, communication		Ratings of resident's professionalism and interpersonal skills		Construct validity, factor analysis, between-groups differences
42. Massagli & Carline (2007); United States	Physician and medicine rehabilitation (56)	NA; nurses, medical students, rehab staff = 12 items	4 nurses, 3 medical students, 10 rehab staff	Professional, clinical competence	Nurses, medical students, rehab staff $\alpha = .89$	Feasibility for physical medicine and rehabilitation residents	3 nurses, 2 rehab staff, 13 medical students $Ep^2 = .70$	Construct validity, factor analysis, between-groups differences
43. Risucci et al. (1989); United States	Surgical residents (32)	NA; colleagues + self = 18 items	23 colleagues, 1 supervisor, 27 self	Professional, clinical interpersonal		Differentiates attending surgeons from surgical residents		Construct validity, factor analysis, between-groups differences
44. Sinclair et al. (2009); United Kingdom	Urologist consultants (10)	SHEFFPAT: patients = 13 items	23 patients	Professional, clinical interpersonal	23 raters ICC > .70	Patients completed instruments on urologists	23 patients $Ep^2 = .70$	Construct validity, factor analysis, between-groups differences
45. Crossley et al. (2008); United Kingdom	Various specialties (137)	SPRAT + SHEFFPAT: colleagues = 24 items, patients = 13 items	9 colleagues + coworkers, 15 patients	Clinical competence, interpersonal	ICC = .78	High response rates for instruments		Construct validity, factor analysis, between-groups differences
46. Joshi, Ling, & Jaeger (2004); United States	Obstetricians and gynecologists (10)	Colleagues = 7 items, coworkers = 53 items, patients = 10 items, student = 12 items	16 Colleagues, 25 coworkers, 10 patients, 12 students	Clinical competence, interpersonal, communication	ICC = .54-.82	Various informants	9 colleagues + coworkers $Ep^2 = .80$	Construct validity, factor analysis, between-groups differences
47. Whitehouse, Hassell, Bullock, Wood, & Wall (2007); United Kingdom	Various specialties (171)	TAB: colleague + coworker = 4 items	3 colleagues, 7 coworkers (74% for colleague + coworker)	Interpersonal, communication		Assessment of interpersonal problems in doctors in training		Construct validity, factor analysis, between-groups differences
48. Wall, Singh, Whitehouse, Hassell, & Howes (2012); United Kingdom	Various specialties (834)	TAB: self = 4 items, colleagues + coworkers = 4 items	834 self, 10 colleagues + coworkers	Interpersonal, communication		Used self-assessment, compared raters scores on TAB		Construct validity, factor analysis, between-groups differences

Note. ICC = intraclass correlation; MSF = multisource feedback; PAT = Peer Assessment Tool; PAR = Physician Achievement Review; SPRAT = Sheffield Peer Review Assessment Tool; SHEFFPAT = Sheffield Patient Assessment Tool; TAB = Team Assessment of Behaviours.

most of the studies. Intraclass correlations (ICCs; correlation between items and overall ratings) were reported in six (12%) studies, with values typically in the high range ($>.70$) but with the occasional low outlier ($ICC = .34$; Davis, 2002). Generalizability/reproducibility coefficients (Ep^2) were reported for 23 (48%) studies and were adequate with eight or more raters (i.e., $Ep^2 \geq .70$).

Validity

Of the 48 studies that were included in the review, 35 (73%) had results for construct validity, several studies reported evidence of criterion-related validity, and several others reported evidence of content validity.

Content validity. The content validity studies generally focused on determining whether the content that the instrument contained is an adequate sample of the domain it is supposed to represent. Enhancing content validity of instruments can be achieved (sampling of appropriate content and skills) by using a table of specifications based on the list of core competency areas and methods to assess them and having experts systematically review items to ensure that each competency is adequately assessed. Applying this procedure, W. Hall et al. (1999) and Violato, Lockyer, and Fidler (2003), for example, constructed instruments to assess family physicians and surgeons, respectively, in communication skills, interpersonal skills, collegiality, professionalism, and ability to continuously improve. These researchers developed a committee of experts (i.e., physicians, surgeons, psychometric experts) to construct questionnaires of 34 items for medical colleagues, 19 items for coworkers, 33 items for self-assessment, and 39 items for patients. The questionnaires were subsequently sent to physicians and surgeons to provide systematic feedback (a modified Delphi procedure). Questionnaires were edited following the feedback to enhance content validity of the instruments. Overeem et al. (2012) employed a similar procedure to address content validity of MSF instruments adapted for use in the Netherlands.

Criterion-related validity. Several studies reported criterion related-validity evidence by comparing the results of MSF with the results obtained using another assessment method. Criterion-related validity refers to the relationship between scores obtained using the MSF instruments and scores obtained using one or more other instruments or measures. Risucci, Tortolonia, and Ward (1989) examined the predictive validity by comparing MSF with the American Board of Surgery in Training Examination (ABSITE). They found a significant correlation between MSF and the ABSITE ($r = .58, p < .01$). This relationship suggests that as surgeons received higher ratings in MSF, they also received higher rating scores in the ABSITE.

In their MSF study of 356 physicians, Lipner, Blank, Leas, and Fortana (2002) found that the health of the patient was significantly correlated with overall rating ($r = .11, p < .001$); those in better health tended to rate their doctors higher. Patients who had spent more time under the doctor's care tended to rate the doctor higher ($r = .08, p < .001$), and female doctors received higher ratings than did male doctors ($r = .17, p < .001$). Internal medicine program directors' ratings were positively correlated with patient ratings of participants' humanistic qualities ($r = .20, p < .05$) but not with the program directors' overall clinical competence ratings.

Crossley et al. (2008) compared the MSF assessment in the form of Non-Technical Skills for Surgeons (NOTSS) with the Procedures Based Assessment (PBA) global summary and Objective Structured Assessment of Technical Skills (OSATS). They found that the NOTSS scores were positively correlated with PBA global summary scores ($r = .48, p < .001$). Also MSF in the form of NOTSS was positively correlated with the generic part of the OSATS score ($r = .51, p < .001$). Similarly, Yang et al. (2011), in their work in Taiwan with residents, found significant correlations between medical colleague assessments and scores on objective performance measures ($r = .37$ and $.72$, respectively, $p < .01$). Evidence for criterion-related validity was adduced in several other studies as well (see Table 1).

Construct validity. Evidence for construct validity, which refers to the nature of the psychological construct or characteristic being measured by the instrument, was reported in the majority of studies. Most of the work for construct validity were factor analyses to determine the number of factors in the various instruments, the factors' variance-accounting properties, their theoretical meaningfulness, and their coherence. Another common analysis investigating construct validity was between-groups differences typically using analyses of variance.

Violato et al. (2003), for example, conducted principal component factor analysis to derive a five-factor solution for the medical colleague questionnaire accounting for 69% of the variance, three factors for the coworker questionnaire accounting for 70.9%, five factors for the patient questionnaire accounting for 73.5%, and four factors for the self-assessment questionnaire accounting for 65.1%. In addition, the mean score was calculated between self-assessment and medical colleague assessment. Surgeons rated themselves lower than did the medical colleagues, with self-assessment $M = 4.07$ ($SD = 0.73$) and medical colleague $M = 4.50$ ($SD = 0.64$). In a British study of surgical trainees, Crossley et al. (2008) derived four factors with principal component factor analyses in their MSF instruments with six surgical specialties.

Overeem et al. (2012) investigated construct validity of MSF instruments adapted for use in the Netherlands with factor analyses. The peer, coworker, and patient instruments, respectively, had six factors, three factors, and one factor with high internal consistencies (Cronbach's alphas = .95–.96), accounting for 67%, 70%, and 60% of the variance, respectively. They found that peer ratings were positively associated with the patient ratings ($r = .214, p < .01$) and with coworker ratings ($r = .352, p < .01$). Coworker ratings were positively associated with patient ratings ($r = .220, p < .01$).

Although additional work is required to further investigate evidence for the validity of MSF (especially criterion-related, construct, and consequential validity), the research in the present study points to the reliability, feasibility, and tentative validity of MSF and supports its use with health professionals, including its potential use with practicing psychologists.

Domains Assessed by MSF

Professionalism, clinical competence, communication, case management, interpersonal relations, and overall assessment were reported in 81%, 77%, 72%, 42%, 58%, and 8% of the studies, respectively. The term *professionalism*, however, encompassed several subdomains (psychosocial skills, psychosocial manage-

ment, humanistic qualities, compassion, attitude, professional development, teaching, professional responsibilities, and professional management), and *clinical competence* included several subdomains (clinical care, good medical practice, patient care, safe practice, clinical performance, knowledge, critical thinking, diagnosis, and management of complex problem). *Communication* encompassed subdomains as well (communication with staff and interpersonal communication skills), as did *case management* (reporting, self-management, administrative skills, office personal, access to doctor, practice process, physical office, and physical space). *Interpersonal relationships* encompassed relationships with patients, colleagues, and family members; collegiality; collaborator; patient education; information provision; and patient interaction. The last factor was overall assessment.

Discussion

A summary of the MSF empirical review indicates some opportunities for professional psychology. First, MSF has been effectively used with many medical and health professionals (e.g., family physicians, nurses, psychiatrists, anesthesiologists, obstetricians, gynecologists, radiologists, occupational therapists) primarily in Canada, the United States, and the United Kingdom to assess their performance on professionalism, interpersonal relations, clinical competence, communication, and case management. Moreover, the surveys that have been developed and used for these evaluations have proven to be appropriate for all of the informants (self, colleague, coworker, and patient) across all the reviewed studies. Hence, it appears that MSF could similarly be appropriately constructed for use with practicing psychologists as well as confidentially and effectively used throughout Canada, the United Kingdom, and the United States to assess the performance and competence of psychologists on professionalism, interpersonal skills, clinical competence, communication, and case management. As MSF has been used effectively to assess the competence and performance of medical health professionals in a number of countries, it can be used for the same purposes for psychologists as well.

Second, the accumulated evidence from reported studies from 1980 to 2012 indicates that MSF has adequate to good reliability, validity, and feasibility. The overall internal consistency reliability (Cronbach's alpha) of MSF instruments is generally greater than .90 for self and other raters such as patients, coworkers, and colleagues. Mean ratings for all MSF instruments are typically from 4 to 5 (when using 5-point Likert scales and an *unable to assess* category). Generalizability coefficients for the assessors across persons are approximately .80.

There is substantial evidence of content and criterion-related validity and some evidence of construct validity of the MSF instruments applied in the medical professions. Factor analytic studies, for example, have shown that factors tend to be consistent with the intent of the questionnaires, theoretically meaningful, and coherent. Hence, MSF surveys could be constructed for and used by practicing psychologists with similar accumulated evidence of validity, reliability, and feasibility. Although there are issues that need to be further addressed for the use of MSF (e.g., variable reliability across domains), the overall evidence is that the system has very good response rates (i.e., >70%) and is generally considered to be cost-effective. In this regard, the empirical review

indicated that MSF can be carried out by a relatively low number of raters (eight to 10 colleagues, coworkers, 25 patients/clients), and that it generally takes these raters about 6 min or less to respond to the items on the questionnaires. Hence, it would seem logical to assume that the use of MSF in the profession of psychology could be reliably, validly, and feasibly done.

Third, respondents report that the MSF system benefits their practice for personal/professional development (e.g., helps them to focus learning activities to legitimate needs) and for multidisciplinary teamwork (Violato & Lockyer, 2006). Based on our review, MSF appears to be useful for both formative (Allerup et al., 2007; Musick, McDowell, Clark, & Salcido, 2003) and summative evaluations (Lockyer & Clyman, 2008) as well as helpful with quality assurance processes and procedures for training programs (Archer, Norcini, & Davies, 2005). MSF could be similarly beneficial for practicing psychologists (i.e., for personal/professional development, formative and summative evaluation, professional quality assurance). It appears that negative feedback from MSF can evoke negative feelings and interfere with its acceptance in some situations. To overcome this possible distress, it may be helpful to provide interventions for professionals that help them focus their feedback on performance tasks and that facilitate their reflection on the feedback. From our review, the research indicates that feedback might be better received if the respondents are familiar with whom they are rating and are able to observe their professional practice (Sargeant et al., 2003). In addition, this feedback needs to be specific, credible, and useful (Sargeant, Mann, Sinclair, van der Vleuten, & Metsemakers, 2008). Researchers contend that MSF can be a positive approach for practice improvement provided that, for example, skilled facilitators are available to encourage reflection, concrete goals are set, and follow-up interviews are carried out (Oveerem et al., 2009).

Lastly, it is important to note that a typical barrier in using MSF is the occasional difficulty in recruiting enough coworkers and colleagues to do the ratings in some practice situations. In this regard, a unique barrier for the use of MSF with practicing psychologists as was found with psychiatrists (Violato, Lockyer, & Fidler, 2008a) might be that some patients/clients could be too cognitively and/or emotionally incapacitated (e.g., severely depressed, psychotic, intellectually disabled, etc.) to be able to provide patient/client ratings. Hence, selection and recruitment of patients/clients for some practicing psychologists may be more difficult than for others. Moreover, it is also the case that patient satisfaction ratings within the medical field must be contextualized and interpreted relative to particular patient illnesses and issues; in a similar fashion, it will be important to contextualize the type of patients/clients selected and recruited for the evaluation of practicing psychologists with respect to their clinical diagnoses and associated issues (e.g., attention-deficit/hyperactivity disorder, mood disorder, personality disorder). However, notwithstanding these types of barriers and challenges, MSF is quite flexible and can be constructed and used across many different types of practicing psychologists (e.g., clinical psychologists, counseling psychologists, school psychologists) and with respect to their patients/clients.

In addition, MSF can be adapted for unique practice situations (as it has been for radiology and laboratory medicine in the medical field) in which patients are not directly involved with the professional being assessed. In such cases, including situations like

this in psychology, other instruments (e.g., surveys for referring psychologists) can be developed when patient instruments cannot be employed. In addition, there are particular areas of competence unique to specialties of psychology (e.g., counseling psychology, school psychology) that need to be considered and addressed in the development and use of MSF questionnaires with particular practicing psychologists that have not been addressed in the fields of medicine and business (e.g., the evaluation of a psychologists' ability to create and maintain a therapeutic alliance with his or her patient/client or a working alliance with allied professionals within the various settings in which he or she works).

Recommendations for Assessing Psychologists in Practice

In 2002, a conference titled "Future Directions in Education and Credentialing in Professional Psychology" provided a forum for interorganizational discussion for competence needs in the profession that included members from many associations including the American Psychological Association and the Association of Psychology Postdoctoral and Internship Centers (Kaslow, 2004). The conference resulted in a model of understanding the developmental nature of competence in professional psychology across the professional life span: A cube model of *foundational domains* (e.g., professionalism, reflective practice/self-assessment, scientific knowledge and methods, relationships, ethical and legal practice, individual and cultural diversity, interdisciplinary systems), *functional domains* (e.g., assessment, diagnosis, and conceptualization, intervention, consultation, research, evaluation, supervision, training management, administration) of competence, and the stages of professional development of these domains (doctoral education, internship/residency, postdoctoral residency, and continuing competency; Rodolfa et al., 2005). Based on our review of the use of MSF for the assessment of competence of practicing health professionals and the developmental nature of competence in professional psychology, MSF seems well suited for assessing the competence of psychologists in practice.

In 2006, the American Psychological Association Task Force on Assessment of Competency in Professional Psychology provided a report that noted four systems used by health professions to produce competent professionals: entry-level education and training, licensure and practice regulation, continuing professional education, and specialty certification. Four models of assessment were proposed: (a) measures of knowledge (e.g., by way of the Examination for Professional Practice in Psychology), (b) measures of professional decision making, (c) measures of practice performance including professional attributes, and (d) integrated assessments of practice-based skills and tasks (e.g., by way of assessment within education, training, and supervision experiences; American Psychological Association, 2006). Based on our review, it appears that an MSF system can be employed as a measure of performance and competence of practicing psychologists. Licensure and practice regulation of psychologists can potentially involve the assessment of their performance and competence (including professional attitudes) by way of a valid, reliable, and feasible MSF assessment during education, training, and supervision experiences as well. However, this suggestion widens the scope beyond the postlicensure stage of psychologists, and more groundwork will have to be done to review and propose the use of

MSF relative to its use with respect to education, training, and supervision. In this regard, this present study and our recommendations have focused on (and have laid the groundwork for) the potential use of MSF with practicing psychologists. Our suggestion would be that MSF profiles be reviewed by members of a psychologists performance committee (PPC), a multimember (e.g., $n = 6$) regulatory body appointed by the group responsible for administering the program. Should the MSF surveys flag a potential problem, the PPC can work with the psychologist from a quality-improvement perspective. Peer office reviews or other competency assessment tools may be used to assist these psychologists. Confidentiality of information gathered by the questionnaires should be guaranteed under the jurisdiction's privacy laws. MSF information should be for educational (formative) purposes and should not be used in legal or disciplinary proceedings.

Currently, the primary method used by practicing psychologists to evaluate their performance and competence as well as assess their learning needs is by way of self-assessment (Belar et al., 2001; Pope, Sonne, & Greene, 2006). Self-assessment involves self-reflection and evaluation of one's professional strengths and areas of improvement in foundational and functional domains as well as an evaluation of one's limitations and decisions about how to address their developmental needs (Caverzagie, Shea, & Kogan, 2008; Kaslow, Rubin, Forrest, et al., 2007). A major problem with self-assessment as an approach for evaluating competence is that very few self-assessment measures have established adequate psychometric properties, and they tend not to correlate well with ratings by peers and/or supervisors and with measures of performance (Dunning, Heath, & Suls, 2004; Eva, Cunningham, Reiter, Keane, & Norman, 2004; Fletcher & Baldry, 2000; Mattheos, Nattestad, Falk-Nilsson, & Attstrom, 2004; Swick, Hall, & Beresin, 2006).

Systematic reviews in medicine (e.g., Davis et al., 2006) as well as empirical investigations in social psychology (Krueger & Mueller, 2002) have suggested that physicians can be limited in their ability accurately self-assess when their assessments are compared with objectively observed and measured competencies. In a meta-analysis of quantitative self-assessment studies in higher education, Falchikov and Boud (1989) reported the results of 44 studies in a variety of subject areas (e.g., medicine, law, engineering, psychology) that showed, on average, that self-assessors were poor to moderate judges of their performance (correlations between self-assessed and external measures of performance ranged from $-.05$ to $.82$, with a mean correlation of $.39$). Hence, it is possible that practicing psychologists may over- or underestimate their performance, abilities, and skills in their self-assessments. Moreover, as suggested by Wise (2010), most psychology registration and licensing boards are unable to ensure that psychologists are maintaining their professional competence, adequately enhancing their professional skills, or remediating their skill deficits by way of psychologist self-assessment.

Although educational training, professional development, ethical guidelines, professional standards, and self-reflection/assessment are necessary aspects of quality assurance for the practice of psychology, they are not sufficient. The primary, or sole, use of self-assessment by practicing psychologists to ensure the public and psychology registration and licensing boards that they are providing quality care is insufficient. Accordingly, psychology registration and licensing boards should mandate that the perfor-

mance and competencies of practicing psychologists be reviewed and evaluated by way of an MSF system on some temporal cycle (e.g., every 5 years).

Rodolfa, Schaffer, and Webb (2010) remind us that our field has “operated on the assumption that once a person is evaluated to be competent to practice, that person remains competent throughout his or her entire career” (p. 296). J. Hall and Boucher (2003), however, have cautioned that psychologists with the highest incidence of disciplinary action have been in practice for 11 to 25 years. Professional psychology lags behind other health professionals for monitoring of continuing competence after licensure in the public’s perception (Nutt, 2010). There is a public assumption that health professionals undergo periodic evaluation and assessment of their professional skills (AARP, 2007). Once a psychologist is registered and licensed in a state (in the United States) or province (in Canada), however, no further competence or performance assessment is typically conducted; instead, psychologists may be directed to self-reflect and self-assess their professional development and attend continuing education events to maintain their competence. Moreover, there is no independent evidence that provides support for these efforts in maintaining or developing one’s competence (Rodolfa et al., 2010). Many medical health professionals have taken steps toward assuring the public of their continued competence by developing a framework for maintenance of licensure that includes an ongoing process of self-assessment, self-evaluation, and professional development in deficit areas identified through these assessments as well as their demonstrated competence for patient care, professionalism, and communication skills by way of peer assessment, patient reviews, satisfaction surveys, and MSF (Special Committee on Maintenance of Licensure, 2008). Although medical health professionals appear to be far ahead of psychologists in attending to quality assurance, there have been attempts to move the profession of psychology toward using approaches that more comprehensively and objectively monitor and evaluate the practice of psychology beyond educational and training environments (e.g., Kaslow et al., 2009). To date, psychology registration and licensure boards in Canada and the United States have been not moved beyond self-assessment as a means of competence evaluation of practicing psychologists. Therefore, provincial and state licensing boards for psychology in Canada and the United States, respectively, must undertake greater responsibility beyond their efforts to date (e.g., recommending/mandating periodic self-reflection and self-assessment among licensed psychologists) for ensuring the public that psychologists are practicing in a competent manner because they are the only entity with legal authority over psychologists’ practice (Swankin, LeBuhn, & Morrison, 2006).

Conclusion

Based on empirical evidence from health care and our own professional experience, we propose that an MSF assessment system (that is reliable, valid, and feasible) should be developed to monitor the core competencies of psychologists (for an example of what MSF might look like for school psychologists, see Andrews & Violato, 2010). Moreover, our contention is that such a system cannot only identify strengths and weaknesses of the core competencies of psychologists (i.e., summative assessment), but also provide useful information and feedback for professional develop-

ment and enrichment (i.e., formative assessment). In addition, an MSF competence-based assessment system for psychologists can provide feedback to psychologists about their performance and improve their practice. An MSF assessment system that incorporates self-assessment along with peer, coworker, and client/patient assessment could provide information that is not only personally/professionally useful for practicing psychologists but useful for national, provincial, and statewide psychology associations as well as provincial and statewide psychology regulatory boards for their oversight and governance of professional psychology.

One step toward incorporating MSF in psychology more broadly is to have one provincial or statewide psychology association and/or regulatory board or association of state and provincial psychology boards support the formation of an MSF advisory group, develop a table of specifications for MSF instruments, and create items from which a pilot study involving practicing psychologists within that jurisdiction could be conducted to provide evidence of the reliability, validity, and feasibility of MSF with practicing psychologists.

References

- AARP. (2007). *Strategies to improve health care quality in Virginia: Survey of residents age 50+*. Washington, DC: Author.
- Allerup, P., Aspegren, E., Ejlersen, G., Jorgensen, A., Malchow-Moller, A., Moller, K., & Sorensen, B. (2007). Use of 360-degree assessment of residents in internal medicine in a Danish setting: A feasibility study. *Medical Teacher, 29*, 166–170. doi:10.1080/01421590701299256
- American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*, 1060–1073. doi:10.1037/0003-066X.57.12.1060
- American Psychological Association. (2006). *American Psychological Association Task Force on the Assessment of Competence in Professional Psychology: Final report*. Washington, DC: Author.
- Andrews, J., & Violato, C. (2010). The assessment of school psychologists in practice through multisource feedback. *Canadian Journal of School Psychology, 25*, 328–346. doi:10.1177/0829573510373585
- Archer, J., & McAvoy, P. (2011). Factors that might undermine the validity of patient and multi-source feedback. *Medical Education, 45*, 886–893. doi:10.1111/j.1365-2923.2011.04023.x
- Archer, J., McGraw, M., & Davies, H. (2010). Assuring validity of multisource feedback in a national programme. *Postgraduate Medical Journal, 86*, 526–531. doi:10.1136/pgmj.2008.146209rep
- Archer, J. C., Norcini, J., & Davies, H. A. (2005). Use of SPRAT for peer review of pediatricians in training. *British Medical Journal, 330*, 1251–1253.
- Archer, J., Norcini, J., Southgate, L., Heard, S., & Davies, H. (2008). Mini-PAT (Peer Assessment Tool): A valid component of a national assessment programme in the UK? *Advance in Health Sciences Education: Theory and Practice, 13*, 181–192.
- Belar, C. D., Brown, R. A., Hersch, L. E., Hornyak, L. M., Rozensky, R. H., Sheridan, E. P., . . . Reed, G. W. (2001). Self-assessment in clinical health psychology: A model for ethical expansion of practice. *Professional Psychology: Research and Practice, 32*, 135–141. doi:10.1037/0735-7028.32.2.135
- Brinkman, W. B., Geraghty, S. R., Lanpher, B. P., Khoury, J. C., Gonzalez del Rey, J. A., Dewitt, T. G., & Britto, M. T. (2007). Effect of multi-source feedback on resident communication skills and professionalism. *Archives of Pediatrics & Adolescent Medicine, 161*, 44–49. doi:10.1001/archpedi.161.1.44
- Campbell, J., Narayanan, A., Burford, B., & Greco, M. (2010). Validation of a multi-source feedback tool for use in general practice. *Education for Primary Care, 21*, 165–179.

- Campbell, J., Richards, S., Greco, M., Narayanan, A., & Brearley, S. (2008). Assessing the professional performance of UK doctors: An evaluation of the utility of the General Medical Council patient and colleague questionnaires. *Quality & Safety in Health Care, 17*, 187–193. doi:10.1136/qshc.2007.024679
- Canadian Psychological Association. (2000). *Canadian code of ethics for psychologists* (3rd ed.). Ottawa, ON: Author.
- Carraccio, C., Wolfsthal, S. D., Englander, R., Ferentz, K., & Martin, C. (2002). Shifting paradigms: From Flexner to competencies. *Academic Medicine, 77*, 361–367. doi:10.1097/00001888-200205000-00003
- Caverzagie, K. J., Shea, J. A., & Kogan, J. R. (2008). Resident identification of learning objectives after performing self-assessment based upon the ACGME core competencies. *Journal of General Internal Medicine, 23*, 1024–1027. doi:10.1007/s11606-008-0571-7
- Chandler, N., Henderson, G., Park, B., Byerley, J., Brown, W. D., & Steiner, M. J. (2010). Use of a 360-degree evaluation in the outpatient settings: The usefulness of nurse, faculty, patient/family, and resident self-evaluation. *Graduate Medical Education, 2*, 430–434.
- Church, A. H. (1997). Do you see what I see? An exploration of congruence in ratings from multiple perspectives. *Journal of Applied Social Psychology, 27*, 983–1020. doi:10.1111/j.1559-1816.1997.tb00283.x
- Cone, J. J. (2001). *Evaluating outcomes: Empirical tools for effective practice*. Washington, DC: American Psychological Association. doi:10.1037/10384-000
- Crossley, J., McDonnell, J., Cooper, C., McAvoy, P., Archer, J., & Davies, H. (2008). Can a district hospital assess its doctors for re-licensure? *Medical Education, 42*, 359–363. doi:10.1111/j.1365-2923.2007.02977.x
- Davies, H., Archer, J., Bateman, A., Deears, S., Crossley, J., & Southgate, L. (2008). Specialty-specific multisource feedback: Assuring validity, informing training. *Medical Education, 42*, 1014–1020. doi:10.1111/j.1365-2923.2008.03162.x
- Davis, D. A., Mazmanian, P. E., Fordis, M., van Harrison, R., Thorpe, K. E., & Perrier, L. (2006). Accuracy of physician self-assessment compared with observed measures of competence: A systematic review. *JAMA, 296*, 1137–1139.
- Davis, J. D. (2002). Comparison of faculty, peer, self, and nurse assessment of obstetrics and gynecology residents. *Obstetrics and Gynecology, 99*, 647–651. doi:10.1016/S0029-7844(02)01658-7
- DiMatteo, M., & DiNicola, D. (1981). Sources of assessment of physician performance: A study of comparative reliability and patterns of intercorrelation. *Medical Care, 8*, 829–842.
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*, 69–106. doi:10.1111/j.1529-1006.2004.00018.x
- Dupee, J. M., Ernst, N. P., & Caslin, E. K. (2011). Does multisource feedback influence performance appraisal satisfaction? *Nursing Management, 42*, 12–16. doi:10.1097/01.NUMA.0000394060.41121.ce
- Epstein, R. M., & Hundert, E. M. (2002). Defining and assessing professional competence. *Journal of the American Medical Association, 287*, 226–235. doi:10.1001/jama.287.2.226
- Eva, K. W., Cunnington, J. P. W., Reiter, H. I., Keane, D. R., & Norman, G. R. (2004). How can I know what I don't know? Poor self-assessment in a well-defined domain. *Advances in Health Sciences Education, 9*, 211–224. doi:10.1023/B:AHSE.0000038209.65714.d4
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 59*, 395–430. doi:10.3102/00346543059004395
- Falender, C. A., & Shafranske, E. P. (2004a). *360-degree evaluation applied to psychology training*. Manuscript in preparation.
- Falender, C. A., & Shafranske, E. P. (2004b). *Clinical supervision: A competency-based approach*. Washington, DC: American Psychological Association. doi:10.1037/10806-000
- Fantuzzo, J. W., Sisemore, T. A., & Spradlin, W. H. (1983). A competency-based model for teaching skills in the administration of intelligence tests. *Professional Psychology: Research and Practice, 14*, 224–231. doi:10.1037/0735-7028.14.2.224
- Fletcher, C. E., & Baldry, C. (2000). A study of individual differences and self-awareness in the context of multi-source feedback. *Journal of Occupational and Organizational Psychology, 73*, 303–319. doi:10.1348/096317900167047
- Hall, J., & Boucher, A. (2003). Professional mobility for psychologists: Multiple choices, multiple opportunities. *Professional Psychology: Research and Practice, 34*, 463–467. doi:10.1037/0735-7028.34.5.463
- Hall, W., Violato, C., Lewkonja, R., Lockyer, J., Fidler, H., Toews, J., . . . Moores, D. (1999). Assessment of physician performance in Alberta: The Physician Achievement Review. *Canadian Medical Association Journal, 161*, 52–57.
- Hoge, M. A., Morris, J. A., Daniels, A. S., Huey, L. Y., Stuart, G. W., Adams, N., . . . Dodge, J. M. (2005). Report of recommendations: The Annapolis Coalition on Behavioral Health Work Force Competencies. *Administration and Policy in Mental Health, 32*, 651–663. doi:10.1007/s10488-005-3267-x
- Joshi, R., Ling, F. W., & Jaeger, J. (2004). Assessment of a 360-degree instrument to evaluate residents' competency in interpersonal and communication skills. *Academic Medicine, 79*, 458–463. doi:10.1097/00001888-200405000-00017
- Kaslow, N. J. (2004). Competencies in professional psychology. *American Psychologist, 59*, 774–781. doi:10.1037/0003-066X.59.8.774
- Kaslow, N. J., Grus, C. L., Campbell, L. F., Fouad, N. A., Hatcher, R. L., & Rodolfa, E. R. (2009). Competency assessment toolkit for professional psychology. *Training and Education in Professional Psychology, 3*, S27–S45. doi:10.1037/a0015833
- Kaslow, N. J., Rubin, N. J., Bebeau, M., Leigh, I. W., Lichtenberg, J., Nelson, P. D., . . . Smith, I. L. (2007). Guiding principles and recommendations for the assessment of competence. *Professional Psychology: Research and Practice, 38*, 441–451. doi:10.1037/0735-7028.38.5.441
- Kaslow, N. J., Rubin, N. J., Forrest, L., Elman, N. S., Van Horne, B. A., Jacobs, S. C., . . . Thorn, B. E. (2007). Recognizing, assessing, and intervening with problems of professional competence. *Professional Psychology: Research and Practice, 38*, 479–492. doi:10.1037/0735-7028.38.5.479
- Krueger, J., & Mueller, R. A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology, 82*, 180–188. doi:10.1037/0022-3514.82.2.180
- Lelliott, O., Williams, R., Mears, A., Andiappan, M., Owen, H., Reading, P., . . . Hunter, S. (2008). Questionnaires for 360-degree assessment of consultant psychiatrists: Development and psychometric properties. *The British Journal of Psychiatry, 193*, 156–160.
- Lipner, R. S., Blank, L. L., Leas, B. F., & Fortana, G. S. (2002). The value of patient and peer ratings in recertification. *Academic Medicine, 77*, S64–S66. doi:10.1097/00001888-200210001-00021
- Lockyer, J., Blackmore, D., Fidler, H., Crutcher, R., Salte, B., Shaw, K., . . . Wolfish, N. (2006). A study of a multi-source feedback system for international medical graduates holding defined licenses. *Medical Education, 40*, 340–347. doi:10.1111/j.1365-2929.2006.02410.x
- Lockyer, J., & Clyman, G. S. (2008). Multisource feedback (360-degree evaluation). In E. S. Holmboe & R. E. Hawkins (Eds.), *Practical guide to the evaluation of clinical competence* (pp. 75–85). Philadelphia, PA: Mosby.
- Lockyer, J., & Violato, C. (2004). An examination of the appropriateness of using a common peer assessment instrument to assess physician skills across specialties. *Academic Medicine, 79*, S5–S8. doi:10.1097/00001888-200410001-00002

- Lockyer, J., Violato, C., & Fidler, H. (2003). Likelihood of change: A study assessing surgeon use of multi-source feedback data. *Teaching and Learning in Medicine, 15*, 168–174. doi:10.1207/S15328015TLM1503_04
- Lockyer, J., Violato, C., & Fidler, H. (2006a). The assessment of emergency physicians by a regulatory authority. *Academic Emergency Medicine, 13*, 1296–1303. doi:10.1111/j.1553-2712.2006.tb00293.x
- Lockyer, J., Violato, C., & Fidler, H. (2006b). A multi-source feedback program for anesthesiologists. *Canadian Journal of Anesthesia, 53*, 33–39. doi:10.1007/BF03021525
- Lockyer, J., Violato, C., & Fidler, H. (2007). What multisource feedback factors influence physicians' self-assessments? A five-year longitudinal study. *Academic Medicine, 82*, S77–S80. doi:10.1097/ACM.0b013e3181403b5e
- Lockyer, J., Violato, C., Fidler, H., & Alakija, P. (2009). The assessment of pathologists/laboratory medicine physicians through a multisource feedback tool. *Archives of Pathology and Laboratory Medicine, 133*, 1301–1308.
- London, M., & Smither, J. W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance related outcomes? Theory-based applications and directions for research. *Personnel Psychology, 48*, 803–839. doi:10.1111/j.1744-6570.1995.tb01782.x
- Massagli, T. L., & Carline, J. D. (2007). Reliability of a 360-degree evaluation to assess resident competence. *American Journal of Physical Medicine and Rehabilitation, 86*, 845–852. doi:10.1097/PHM.0b013e318151ff5a
- Matheos, N., Nattestad, A., Falk-Nilsson, E., & Attstrom, R. (2004). The interactive examination: Assessing students' self-assessment ability. *Medical Education, 38*, 378–389. doi:10.1046/j.1365-2923.2004.01788.x
- Meng, L., David, G., & Rita, M. (2009). Evaluating professionalism and interpersonal and communication skills: Implementing a 360-degree evaluation instrument in an anesthesiology residency program. *Journal of Graduate Medical Education, 1*, 216–220. doi:10.4300/JGME-D-09-00014.1
- Musick, D. W., McDowell, S. M., Clark, N., & Salcido, R. (2003). Pilot study of a 360-degree assessment instrument for physical medicine and rehabilitation residency programs. *American Journal of Physical Medicine and Rehabilitation, 82*, 394–402. doi:10.1097/01.PHM.0000064737.97937.45
- Noonan, C. L., Monagle, J., & Castanelli, D. (2011). Development of a multi-source feedback tool for consultant anaesthetist performance. *Australian Health Review, 35*, 141–145.
- Nutt, R. L. (2010). Are we meeting public expectations for competence? *Professional Psychology: Research and Practice, 41*, 294–295.
- Overeem, K., Wollersheim, H., Arah, O., Crujlsberg, J. K., Grol, P. T., & Lombarts, K. (2012). Evaluation of physicians' professional performance: An iterative development and validation study of multisource feedback instruments. *BMC Health Services Research, 12*, 80–91. doi:10.1186/1472-6963-12-80
- Overeem, K., Wollersheim, H., Driessen, E., Lombarts, K., van de Ven, G., Grol, R., & Onyebuchi, A. (2009). Doctors' perceptions of why 360-degree feedback does (not) work: A qualitative study. *Medical Education, 43*, 874–882. doi:10.1111/j.1365-2923.2009.03439.x
- Pollock, R. A., Donnelly, M. B., Plymale, M. A., Stewart, D. H., & Vasconez, H. C. (2007). 360-degree evaluations of plastic surgery resident accreditation council for graduate medical education competencies: Experiences using a short form. *Plastic and Reconstructive Surgery, 122*, 639–649.
- Pope, K. S., Sonne, J. L., & Greene, B. (2006). *What therapists don't talk about and why: Understanding taboos that hurt us and our clients*. Washington, DC: American Psychological Association. doi:10.1037/11413-000
- Qu, B., Zhao, Y. H., & Sun, B. Z. (2012). Assessment of resident physicians in professionalism, interpersonal and communication skills: A multisource feedback. *International Journal of Medical Sciences, 9*, 228–236. doi:10.7150/ijms.3353
- Ramsey, P. G., Wenrich, M. D., Carline, J. D., Inui, T. S., Larson, E. B., & LoGerfo, J. P. (1993). Use of peer ratings to evaluate physician performance. *Journal of the American Medical Association, 269*, 1655–1660. doi:10.1001/jama.1993.03500130069034
- Risucci, D. A., Tortolonia, A. J., & Ward, R. J. (1989). Ratings of surgical residents by self, supervisors and peers. *Surgery, Gynecology & Obstetrics, 169*, 519–526.
- Roberts, M., Borden, K., Christiansen, M., & Lopez, S. (2005). Fostering a culture shift: Assessment of competence in the education and careers of professional psychologists. *Professional Psychology: Research and Practice, 36*, 355–361. doi:10.1037/0735-7028.36.4.355
- Rodolfa, E. R., Bent, R. J., Eisman, E., Nelson, P. D., Rehm, L., & Ritchie, P. (2005). A cube model for competency development: Implications for psychology educators and regulators. *Professional Psychology: Research and Practice, 36*, 347–354. doi:10.1037/0735-7028.36.4.347
- Rodolfa, E., Schaffer, J. B., & Webb, C. (2010). Continuing education: The path to life-long competence? *Professional Psychology: Research and Practice, 41*, 295–297.
- Sala, F., & Dwight, S. A. (2002). Predicting executive performance within multirater surveys: Whom you ask makes a difference. *Consulting Psychology Journal: Practice and Research, 54*, 166–172. doi:10.1037/1061-4087.54.3.166
- Sargeant, J. M., Mann, K. V., Ferrier, S. N., Langille, D. B., Muirhead, P. D., & Sinclair, D. E. (2003). Responses of rural family physicians and their colleagues and co-worker raters to a multi-source feedback process: A pilot study. *Academic Medicine, 78*, S42–S44. doi:10.1097/00001888-200310001-00014
- Sargeant, J., Mann, K., Sinclair, D., van der Vleuten, C., & Metsemakers, J. (2008). Understanding the influence of emotions and reflection upon multi-source feedback acceptance and use. *Advances in Health Sciences Education, 13*, 275–288. doi:10.1007/s10459-006-9039-x
- Sharpless, B. A., & Barber, J. P. (2009). A conceptual and empirical review of the meaning, measurement, development, and teaching of intervention competence in clinical psychology. *Clinical Psychology Review, 29*, 47–56. doi:10.1016/j.cpr.2008.09.008
- Shaw, B. F., & Dobson, K. S. (1988). Competency judgments in the training and evaluation of psychotherapists. *Journal of Consulting and Clinical Psychology, 56*, 666–672. doi:10.1037/0022-006X.56.5.666
- Sinclair, A. M., Gunendran, T., Archer, J., Bridgewater, B., O'Flynn, K. J., & Pearce, U. (2009). Re-certification for urologists: Is the SHEFFPAT questionnaire valid for assessing clinicians' "relationships with patients"? *British Journal of Medical and Surgical Urology, 2*, 100–104. doi:10.1016/j.bjmsu.2008.12.011
- Special Committee on Maintenance of Licensure. (2008). *Draft report on maintenance of licensure*. Washington, DC: Federation of State Medical Boards.
- Swankin, D., LeBuhn, R. A., & Morrison, R. (2006). *Implementing continuing competency requirements for health care practitioners*. Washington, DC: AARP Public Policy Institute.
- Swick, S., Hall, S., & Beresin, E. (2006). Assessing the ACGME competencies in psychiatry training programs. *Academic Psychiatry, 30*, 330–351. doi:10.1176/appi.ap.30.4.330
- Violato, C., & Lockyer, J. (2006). Self and peer assessment of pediatricians, psychiatrists and medicine specialists: Implications for self-directed learning. *Advances in Health Sciences Education, 11*, 235–244. doi:10.1007/s10459-005-5639-0
- Violato, C., Lockyer, J., & Fidler, H. (2003). Multisource feedback: A method of assessing surgical practice. *British Medical Journal, 326*, 546–548. doi:10.1136/bmj.326.7388.546

- Violato, C., Lockyer, J., & Fidler, H. (2006). Assessment of pediatricians by a regulatory authority. *Pediatrics*, *117*, 796–802. doi:10.1542/peds.2005-1403
- Violato, C., Lockyer, J., & Fidler, H. (2008a). Assessment of psychiatrists with multisource feedback. *The Canadian Journal of Psychiatry*, *53*, 525–533.
- Violato, C., Lockyer, J., & Fidler, H. (2008b). Changes in performance: A 5-year longitudinal study of participants in a multi-source feedback programme. *Medical Education*, *42*, 1007–1013. doi:10.1111/j.1365-2923.2008.03127.x
- Violato, C., Marini, A., Toews, J., Lockyer, J., & Fidler, H. (1997). Feasibility and psychometric properties of using peers, consulting physicians, co-workers, and patients to assess physicians. *Academic Medicine*, *72*, S82–S84. doi:10.1097/00001888-199710001-00028
- Violato, C., & Saberton, S. (2006). Assessing medical radiation technologists in practice: A multi-source feedback system for quality assurance. *Canadian Journal of Medical Radiation Technology*, *37*, 10–17. doi:10.1016/S0820-5930(09)60131-6
- Violato, C., Worsfold, L., & Polgar, J. M. (2009). Multisource feedback systems for quality improvement in the health professions: Assessing occupational therapists in practice. *Journal of Continuing Education in the Health Professions*, *29*, 111–118. doi:10.1002/chp.20020
- Wall, D., Singh, D., Whitehouse, A., Hassell, A., & Howes, J. (2012). Self-assessment by trainees using self-TAB as part of the team assessment of behavior multisource feedback tool. *Medical Teacher*, *34*, 165–167. doi:10.3109/0142159X.2012.644840
- Warm, E. J., Schauer, D., Revis, B., & Boer, J. R. (2010). Multisource feedback in the ambulatory setting. *Journal of Graduate Medical Education*, *2*, 269–277.
- Wenrich, M. D., Carline, J. D., Giles, L. M., & Ramsey, P. G. (1993). Ratings of the performances of practicing internists by hospital-based registered nurses. *Academic Medicine*, *68*, 680–687. doi:10.1097/00001888-199309000-00014
- Whitehouse, A., Hassell, A., Bullock, A., Wood, L., & Wall, D. (2007). 360-degree assessment (multisource feedback) of UK trainee doctors: Field testing of team assessment of behaviours (TAB). *Medical Teacher*, *29*, 171–176. doi:10.1080/01421590701302951
- Wise, E. (2010). Maintaining and enhancing competence in professional psychology: Obsolescence, life-long learning, and continuing education. *Professional Psychology: Research and Practice*, *41*, 289–292.
- Wood, J., Collins, J., Burnside, E., Albanese, M. A., Propeck, P. A., Kelcz, F., . . . Schmaltz, M. (2004). Patient, faculty, and self-assessment of radiology resident performance: A 360-degree method of measuring professionalism and interpersonal/communication skills. *Academic Radiology*, *11*, 931–939.
- Wood, L., Wall, D., Bullock, A., Hassell, A., Whitehouse, A., & Campbell, I. (2006). Team observation: A six-year study of the development and use of multi-source feedback (360-degree assessment) in obstetrics and gynecology training in the UK. *Medical Teacher*, *28*, e177–e184. doi:10.1080/01421590600834260
- Yang, Y. Y., Lee, F. Y., Hsu, H. C., Huang, C. C., Chen, J. W., Cheng, H. M., . . . Huang, C. C. (2011). Assessment of first-year post-graduate residents: Usefulness of multiple tools. *Journal of the Chinese Medical Association*, *74*, 531–538.

Received October 26, 2012

Revision received March 12, 2013

Accepted March 14, 2013 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://notify.apa.org/> and you will be notified by e-mail when issues of interest to you become available!