

Establishing the Reliability and Validity of a Competency Based Evaluation Tool to Assess M.S.W. Students' Performance in Field Practicum

Authors: Bogo, M., Regehr, C., Hughes, J., Power, R., and Globerman, J.

Editor of Fact Sheet: Lambert, L.

Funding from the Social Sciences and Humanities Research Council of Canada

Research Question

How reliable and valid is the tool currently used to evaluate the performance of M.S.W. students in practicum? In other words, is the Competency Based Evaluation Tool (CBE), used in one social work program, effective in identifying variations in student ability?

Description of the Competency Based Evaluation Tool Used to Evaluate M.S.W. Student Performance in Field Practicum

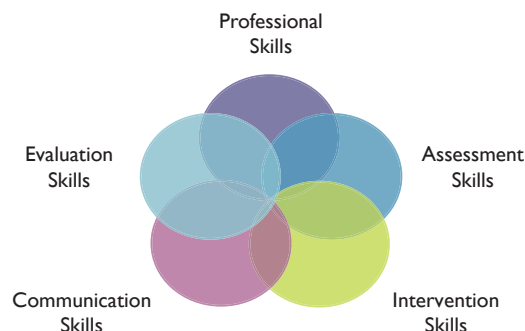
Since 1979, the University of Toronto, Faculty of Social Work has used a 5-point Likert scale to rate social work students' performance or competence on multiple dimensions of social work practice.

The scale identifies stages in students' skill acquisition in five key areas with a total of eighty items. Revisions have been made to the items over time to reflect changes in the field and findings from research on practice.

What Dimensions of Social Work Practice Does the Tool Seek to Measure?

- Professional skills
- Assessment skills
- Intervention skills
- Evaluation skills
- Communication skills

Although knowledge of the organization and of relevant policies, programs, and community resources are part of the tool, these dimensions were not used in this study as they rate student comprehension rather than behavior.



How are Students Evaluated in Practicum?

Students are evaluated over a period of approximately 20 weeks in a number of ways including weekly audio or videotapes of their one-on-one sessions with clients, weekly process and summary records with one to two clients, and their participation in team meetings.

Sample

480 files of students who entered the two-year M.S.W. program between 1992 and 1998 were included for analysis in the study. Inclusion criteria were that a student must have participated in at least one micro-level practicum, and the student file could not contain missing data. After removing those that did not meet the criteria, the sample was reduced to 227 for Year 1 Evaluation Forms and 253 for Year 2 Evaluation Forms.

Data Collection

Students' age, gender, GPA entering and exiting the program, year entering the program and ratings on Year 1 and Year 2 final practicum evaluation forms were collected. An extra variable was added that identified whether a student had difficulty in his/her practicum and could have been at risk of failing. The Director of Field Education identified these students.

* This fact sheet is a summary of the following article:

Bogo, M., Regehr, C., Hughes, J., Power, R., & Globerman, J. (2002). Evaluating a measure of student field performance in direct service: Testing reliability and validity of explicit criteria. *Journal of Social Work Education*, 38(3), 385-401.

STUDY FINDINGS

Reliability of the Measure

Internal Consistency

Internal consistency is a measure of the reliability of various items on a tool that are intended to measure the same characteristic (statistics.com). Statistical analyses of variance among the first-year and second-year students revealed that there is very high internal consistency of the items within each subscale on the tool.

Comparing the Evaluations of the Same Students in Year 1 and Year 2

The correlation between Year 1 and Year 2 scores, although significant was low. The low level of correlation may be due to some students improving their skills more than others, variability in how field instructors use the scale or stated expectations in the field instructor manual that stipulate that Year 2 students should score a 4 or 5 on most items.

Validity Criteria

Content Validity

Content validity is a qualitative measure of the relevance of the items on an instrument (statistics.com). The evaluation tool used to assess student performance in practicum was developed based on consultations with faculty members in social work and experienced field practice educators, competency based education models, and existing practice outcome research.

Construct Validity

Construct validity is a measure of how well the instrument performs in practice (statistics.com). A comparison of the factor analysis of this scale on the same group of students in both their first and second year practicum resulted in some support for the construct validity of the measure.

Criterion Validity

Criterion validity is a measure of the level of agreement between the results from the instrument under question, and other comparable instruments that are well established in the field (statistics.com). The analysis produced the following findings:

- **AGE** - Age was not significantly associated with Year 1 or Year 2 evaluation scores (contrary to some of the literature).

- **GRADE POINT AVERAGE (GPA)** - Students with higher GPAs entering and exiting the M.S.W. program tended to have higher total scores on the evaluation forms in Year 1 and Year 2 (consistent with other findings in the literature).
- **GENDER** - There were no significant differences in male and female students' scores on the Year 1 or Year 2 evaluations (contrary to some of the literature).
- **STUDENTS IDENTIFIED AS HAVING DIFFICULTY** - There was a significant difference in Year 1 total scores between those students identified as having difficulty in practicum and those who were not identified. This difference was not significant in Year 2 (this variable was not discussed in the literature).

Limitations

1. Field instructors are provided with expected levels of student performance for Year 1 and Year 2 in the field manual. These expectations may explain why the scale was unable to differentiate between students identified as having difficulty in Year 2 of the program and the reason for the low test-retest reliability.
2. Items on the evaluation tool are grouped by type of skills. Grouping items in this way may create a "halo effect" whereby the perception of a factor is influenced by the perception of other factors in that category.
3. Field instructors and students collaboratively evaluate the student's performance. If field instructors feel uncomfortable confronting students about areas of concern and assigning a low rating, they may negotiate with students thereby inflating the ratings.
4. Field instructors may be reluctant to assign low ratings to students that are below expected levels of performance because in such cases, the field instructor has to work with the faculty field liaison to evaluate the student's work and devise an appropriate response which can involve lengthy, time-consuming negotiations.

Conclusions

Measuring competency in student practicum continues to represent a challenge for social work educators, but these outcomes are critical because social work schools are responsible for producing competent professional practitioners.

- The findings from this study indicate that there is some theoretical and practice coherence in the factors and associated skills that comprise the Competency Based Evaluation Tool (CBE).
- It is recommended that all social work programs test and report the reliability and validity of the tools they use to evaluate students' performance in Field Practicum.

Further Research

Studies comparing this tool with other tools and methods of evaluation are reported in:

Bogo, M., Regehr, C., Power, R., & Regehr, G. (2006). Improving measures of social work practice performance in the field. Council on Social Work Education, APM, Chicago.

Regehr, C., Bogo, M., Power, R., & Regehr, G. (2006). An innovative approach to field assessments of student competence. Council on Social Work Education, APM, Chicago.

Regehr, G., Bogo, M., Regehr, C., & Power, R. (2007). Can we build a better mousetrap? Improving measures of social work practice performance in the field. *Journal of Social Work Education*, 43(2), 327-343.

Studies exploring the dynamics in the evaluation process which influence the effectiveness of any evaluation tool are reported in:

Bogo, M., Regehr, C., Power, R., & Regehr, G. (2007). When values collide: Providing feedback and evaluating competence in social work. *The Clinical Supervisor* 26(1/2), 99-117.

RIEBSW information sheets are produced and distributed by the Research Institute for Evidence-Based Social Work in the Faculty of Social Work, University of Toronto to provide timely access to evidence-based research.

This information sheet can be downloaded from: www.socialwork.utoronto.ca/competency