Quantitating Qualitative Issues in Residency Training:

Development and Testing of a Scaled Program Evaluation Questionnaire

CHARLES B. SEELIG, MD, MS

Objectives: To develop and test a scaled program evaluation questionnaire focusing on resident satisfaction with workload, learning environment, and stress.

Design/participants: Phase 1: A cross-section of 92 residents from five programs completed questionnaires for factor analysis and descriptive statistics. Phase 2: A three-year prospective study of questionnaire responses in a single program.

Questionnaire development: After extensive literature review, 33 Likert-scaled statements were written, dealing with situational, personal, and professional issues. After pilot testing, the 92 questionnaires obtained in Phase 1 were factor analyzed, resulting in three distinct scales: workload, faculty/learning environment, and stress.

Interventions: Program changes in Phase 2 included the introduction of a night float between 1989 and 1990, and an increase in senior resident call between 1990 and 1991.

Results: Phase 1: The first-year residents reported significantly less stress and generally greater satisfaction with workload and learning environment than did the second- or third-year residents. Marked differences between programs were also present in the mean scores on all three scales. Phase 2: The introduction of a night float did not significantly affect response to the questionnaire, but the increase in nights on call significantly increased stress and dissatisfaction with the faculty/learning environment.

Conclusions: This study demonstrates the techniques needed to develop and use scaled program evaluation questionnaires. It is hoped that widespread use and validation of such instruments may result in greater responsiveness to the needs of trainees and more facilitative environments in which to acquire medical knowledge.

Key words: residency; program evaluation; questionnaires; night float; education; stress; satisfaction.

J GEN INTERN MED 1993;8:610–613.

RECENTLY, changes in career expectations of medical school graduates, as well as pressures in society, have caused the reevaluation of internal medicine as a career, and graduate medical education in general.[1-6] In particular, the process of educating physicians has received greater attention. Stress in residency has been the subject of many reports,[7-35] and some recommendations for how to reduce this stress have been made.[22-35] Studies of residency stress have utilized a variety previously described psychological testing instruments[7, 8, 16-18, 20, 21] and locally developed questionnaires.[9, 13] This article describes the development and testing of a scaled residency program evaluation questionnaire focusing on resident satisfaction with workload, faculty/learning environment, and overall stress. It presents the descriptive statistics and discusses the factor analysis used to derive distinct attitudinal scales from the questionnaires completed by a diverse group of 92 medical residents. Last, the results obtained in three years of using this instrument in a single university-affiliated community hospital internal medicine residency training program are analyzed.

Received from the University of North Carolina, Chapel Hill, and the Internal Medicine Residency Training Program, New Hanover Regional Medical Center, Wilmington, North Carolina.

Address correspondence and reprint requests to Dr. Seelig: Internal Medicine Training Program, New Hanover Regional Medical Center, P.O. Box 9025, Wilmington, NC 28402-9025.

METHODS

Phase 1 Participants

Ninety-two internal medicine residents were enrolled in the study, 45 from three community hospitals, and 47 from two university training programs. Forty-eight were postgraduate year 1 (PGY-1) residents, 16 were PGY-2, 25 were PGY-3, and there were three whose year of training was unknown.

Questionnaire Development, Factor Analysis, and Scale Development

Issues felt to be important to residents were gleaned from prior reports and the *Directory for Graduate Medical Education Programs.*[13, 15,18-28, 36] Thirty-three specific questions (Appendix A) dealing with these situational, personal, and professional issues were written in statement form for use in a five-point Likert-scaled questionnaire (1 = strongly disagree to 5 = strongly agree). An optional page of five open-ended questions followed. The questionnaire was then judged for face validity by the program faculty at New Hanover Regional Medical Center (NHRMC), a university affiliated, community hospital internal medicine residency training program, and by an outside consultant. Next, the questionnaire was pilot tested at NHRMC. The residents who participated felt that the questionnaire results accurately reflected their attitudes. We therefore proceeded to administer the Phase 1 questionnaire without further modification.

Following the principles of scale development as outlined by DeVellis[37] and Comrey,[38] prior to factor analysis, the questions were face-assigned to six scales: General Learning Environment, Faculty Issues, Clerical/Administrative Issues, Caseload Stress, Personal Stress, and Time Demands. Next, several factor analytic methods were applied. Each resultant factor pattern matrix was evaluated for the distribution of the variance explained by the factors. The best solution was obtained

with an iterated principal components factor analysis, retaining five factors in the final analysis, followed by a procrustean rotation of the correlation axes. This resulted in three main factors, fairly evenly sharing the variance explained in the analysis, and two other, less important, factors. All questions that loaded (correlated) strongly on these five factors were checked for face validity, consistency with the six original face-assigned scales, and internal correlations with repeated alpha factor analyses. These procedures resulted in five distinct scales. Three consisted of five related questions each, and two consisted of two related questions each. The latter were too specific to be useful scales, leaving three final, five-question scales, which were then tested by Cronbach's alpha analysis.

Phase 2 Participants

One hundred percent of the residents at NHRMC completed the questionnaire in October 1989, 1990, and 1991. This resulted in nine paired responses to compare between 1989 and 1990 and ten paired responses to compare between 1990 and 1991.

Phase 2 Program Interventions

Between 1989 and 1990, the only major program change at NHRMC was the institution of a night-float system. Between 1990 and 1991, several interrelated program changes were made, involving the intensive care and elective rotations, which resulted in an increase in the number of nights on call for second- and third-year residents by as much as 50%. Resident input was not obtained prior to the institution of these changes.

TABLE 1

Final Scale Questions*

Workload Scale (Cronbach's alpha = 0.82)
- Hospital support services are sufficient to help me care for my inpatients.
- The caseload on the wards is about right.
- The average number of work-ups on call days is reasonable.
- There is enough clerical and administrative support provided by the program.
- The workload is generally excessive on the wards.

Faculty/Learning Environment Scale (Cronbach's alpha = 0.80)
- I get timely and appropriate feedback from faculty.
- I have received sufficient counseling from faculty to help with career planning.
- Full-time faculty members contribute to a great extent to the teaching I've received.
- I have enough personal support from faculty.
- I receive enough instruction on what is expected of me in each level of my training.

Stress Scale (Cronbach's alpha = 0.87)
- I have far too little leisure time.
- There are too many difficult patient management problems on the wards (AIDS, alcoholism, nursing home placement, etc.).
- I often feel "stressed out" or depressed.
- I often feel tired and overworked.
- I rarely have time to read.

*Negatively worded questions are scored in reverse.

TABLE 2

Mean Phase 1 Scale Scores—Mean (SD)—Analyzed by Postgraduate Year (PGY) of Training and by Program

	n	Workload	Faculty/Learning Environment	Stress
PGY-1	47	3.42 (0.95)	3.46 (0.69)	2.85* (0.84)
PGY-2	16	2.89 (1.10)	3.16 (0.76)	3.45* (1.04)
PGY-3	26	3.24 (0.80)	3.11 (0.90)	3.35* (0.65)
Program 1	31	3.95† (0.56)	3.70‡ (0.72)	2.69† (0.69)
Program 3	10	3.50† (0.56)	3.18‡ (0.75)	2.94† (0.68)
Program 4	35	2.55† (0.93)	3.13‡ (0.71)	3.43† (0.94)
Program 5	12	3.65† (0.44)	3.15‡ (0.92)	3.29† (1.29)

*$p = 0.02$ (ANOVA).
†$p = 0.0001$ (ANOVA).
‡$p = 0.03$ (ANOVA).

Analysis

Factor analysis, alpha factor analysis, and Cronbach's alpha analysis were carried out using SAS.[39] Differences in Phase 1 responses, by resident year of training and by program, were compared by analysis of variance (ANOVA). Differences in Phase 2 responses, by year, were compared by paired-sample t-tests.

RESULTS

Phase 1

Table 1 contains the three distinct, five-question scales, derived by factor analysis. The high Cronbach's alphas reflect their internal consistency.

Table 2 shows the mean scale scores obtained, analyzed by postgraduate year of training and by program. The stress scale responses of the PGY-1 residents were significantly lower than those of the PGY-2 or PGY-3 residents, indicating a lower level of stress among the PGY-1 respondents. Significant differences were also found between programs. Program 2 returned only three questionnaires and was left out of this analysis. Resident satisfaction with workload and the faculty environment was greatest, and stress lowest, among the respondents from Program 1 (a small community hospital program). Satisfaction was lowest, and stress highest, in Program 4 (a large university program). Results were intermediate in Program 3 (a small university program) and Program 5 (a large community hospital program).

Phase 2

Table 3 shows the results of the questionnaire study of a single community hospital (NHRMC) residents' responses, obtained in October 1989, 1990, and 1991.

TABLE 3

Mean Phase 2 Scale Scores—Mean (SD)—from New Hanover Regional Medical Center, Analyzed by Year

Year	Workload	Faculty/Learning Environment	Stress
1989	4.14 (0.41)	4.18 (0.55)	2.62 (0.72)
1990	4.24 (0.41)	3.81* (0.87)	2.45† (0.81)
1991	4.08 (0.63)	2.85* (1.08)	3.14† (1.08)

*p = 0.001 (paired sample t-test, $n = 10$).
†p = 0.05 (paired sample t-test, $n = 10$).

From 1989 to 1990, after the introduction of a night-float system, none of the changes in mean scale results achieved statistical significance. However, from 1990 to 1991, after the increase in nights on call, the stress scale showed a significant increase in stress, and satisfaction with the faculty/learning environment showed a highly significant decrease.

DISCUSSION

This report contributes to the study of residency stress and program evaluation. It demonstrates techniques that can be used to develop objective measures of subjective issues in residency training. Despite careful preparation of questionnaires, not all questions are perceived as anticipated. This is particularly true for questions dealing with broad issues, such as situational and personal stress. Some questions may be found, after factor analysis, to be too specific or too vague. Others, such as the caseload and administrative questions, in this study, load on the same factor, resulting in a single, combined scale.

The descriptive statistics obtained with the scaled questionnaire responses provide some interesting new insights into residency stress. In the past, studies have focused on stress in the PGY-1 year.[7-10, 18, 23] It has been suggested that a nadir of satisfaction and mood occurs in the first two to three months, with improvement occurring toward the beginning of the PGY-2 year.[7-9] Our study, done in October, should have caught this nadir, yet the PGY-1 respondents were more satisfied than their higher-level peers. This suggests that efforts to lessen the stresses in the PGY-1 year have been successful. Perhaps, in the past, the grueling workload in the PGY-1 year overshadowed any increased stress due to the greater supervisory and patient care responsibilities in the later years of training. The results of the between-program differences are interesting, but variability in the proportion of PGY-1, PGY-2, and PGY-3 respondents makes any conclusions suspect.

The results of the "real life" experiment at NHRMC illustrate the usefulness of the questionnaire in reflecting changes in attitudes among residents over time. It clearly demonstrated its ability to separate issues of workload (which didn't change) from those of stress and satisfaction with the faculty/learning environment. It is possible that there were pressures other than the program interventions described that might have influenced the results, but the changes in resident attitudes reflected in the questionnaire responses also were subjectively apparent to the program faculty.

Although the current scaled questionnaire could now be used generally, the author hopes to continue to improve it through cooperation with other investigators. This study demonstrates the feasibility of producing useful tools to help program directors quantitate some of the qualitative issues in residency training, resulting, it is hoped, in better, more facilitative learning environments in which to acquire medical knowledge.

REFERENCES

1. Swanson AG. Medical students' specialty choices. In: Harris DL. Reform in Medical Education and Medical Education in the Ambulatory Setting. Washington, DC: US Dept of Health and Human Services, 1991;35-9.
2. Schwartz MD, Linzer M, Babbott D, et al. Medical Student Interest in Internal Medicine. Initial Report of the Society of General Internal Medicine Interest Group Survey on Factors Influencing Career Choice in Internal Medicine. Ann Intern Med. 1991; 114:6-15.
3. Babbott D, Levey GS, Weaver SO, Killian CD. Medical student attitudes about internal medicine: a study of U.S. medical school seniors in 1988. Ann Intern Med. 1991;114:16-22.
4. Reitemeier RJ. The leadership crisis in internal medicine: what can be done? Ann Intern Med. 1991;114:69-75.
5. Lewis CE, Prout DM, Chalmers EP, Leake B: How satisfying is the practice of internal medicine? A national survey. Ann Intern Med. 1991;114:1-5.
6. Lipkin M, Levinson W, Barker R, et al. Primary care internal medicine: a challenging career choice for the 1990's. Ann Intern Med. 1990;112:371-8.
7. Uliana RL, Hubbell FA, Wyle FA, Gordon GH. Mood changes during the internship. J Med Educ. 1984;59:118-23.
8. Gordon GH, Hubbell FA, Wyle FA, Charter RA. Stress during internship: a prospective study of mood states. J Gen Intern Med. 1986;1:228-31.
9. Girard DE, Elliot DL, Hickam DH, et al. The internship—a prospective investigation of emotions and attitudes. West J Med. 1986;144:93-8.
10. Girard DE, Sack RL, Reuler JB, Chang MK, Nardone DA. Survival of the medical internship. Forum Med. 1980;3:460-3.
11. Zeigler JL, Strull WM, Larsen RC, et al. Stress and medical training. West J Med. 1985;142:814-9.
12. Rudner HL. Stress and coping mechanisms in a group of family practice residents. J Med Educ. 1985;60:564-6.
13. Alexander D, Monk JS, Jonas AP. Occupational stress, personal strain, and coping among residents and faculty members. J Med Educ. 1985;60:830-9.
14. Smith JW, Denny WF, Witzke DB. Emotional impairment in internal medicine house staff. JAMA. 1096;255:1155-8.
15. Reuben DB. Psychologic effects of residency. South Med J. 1983;76:380-3.
16. Reuben DB. Depressive symptoms in medical house officers. Arch Intern Med. 1985;145:286-8.
17. Samuel SE, Lawrence JS, Schwartz HJ, Weiss JC, Setzer JL. Investigating stress levels of residents: a pilot study. Med Teacher. 1991;13:89-92.
18. Badger LW, Chesebro MJ, Hartman JA. First year residency stress: sources and mediators. Fam Pract Res J. 1987;7:104-13.
19. Walerstein SJ, Rosner F, Wallace EZ. House staff stress. N Y State J Med. 1989;89:454-7.
20. Schwartz AJ, Black ER, Goldstein MG, Jozefowicz RF, Emmings FG. Levels and causes of stress among residents. J Med Educ. 1987;62:744-53.

21. Whitley TW, Gallery ME, Allison EJ Jr, Revicki DA. Factors associated with stress among emergency medicine residents. Ann Emerg Med. 198;18:1157-61.
22. Ditts PV Jr, Dilts SL. Stress in residency: proposals for solution. Am J Obstet Gynecol. 1987;157:1093-6.
23. McCue JD. The distress of internship, causes and prevention. N Engl J Med. 1985;312:449-52.
24. Matheson KH. Stress and stress counselling. Postgrad Med J. 1990;66:738-42.
25. Levin R. Beyond "The Men of Steel": the origins and significance of house staff training stress. Gen Hosp Psychiatry. 1988; 10:114-21.
26. Resident Services Committee, Association of Program Directors in Internal Medicine. Stress and impairment during residency training: strategies for reduction, identification, and management. Ann Intern Med. 1988;109:154-61.
27. Colford JM Jr, McPhee SJ: The ravelled sleeve of care: managing the stresses of residency training. JAMA. 1989;261:889-93.
28. Butterfield PS. The stress of residency: a review of the literature. Arch Intern Med. 1988;148:1428-35.
29. Martin AR: Stress in residency: a challenge to personal growth. J Gen Intern Med. 1986;1:252-7.
30. McCue JD, Sachs CL. A stress management workshop improves residents' coping skills. Arch Intern Med. 1991;151:2273-7.
31. McCall TB. The impact of long working hours on resident physicians. N Engl J Med. 1988;318:775-8.
32. Alpert JS, Coles R. Residency reform: an urgent necessity. Arch Intern Med. 1988;148:1507-8.
33. Giardino AP, Giardino ER. Medical education and resident work hours. JAMA. 1990;263:620.
34. Resident Physician Section: American Medical Association Board acts to limit resident work hours. JAMA. 1989;262:2970.
35. Trontell MC, Carson JL, Taragin MI, Duff A. The impact of the night float system on internal medicine residency programs. J Gen Intern Med. 1991;6:445-9.
36. Accreditation Council for Graduate Medical Education. Directory of Graduate Medical Education Programs. Chicago, IL: American Medical Association, 199;9-47.
37. DeVellis RF. Guidelines in scale development. In: Scale Development: Theory and Applications. Newbury Park, CA: Sage Publications, 1991;51-90.
38. Comrey AL. Common methodological problems in factor analytic studies. J Consult Clin Psychol 1978;46:648-59.
39. SAS 6.04. Cary, NC: SAS Institute, Inc., 1987.

APPENDIX A

Residency Program Evaluation Questionnaire

Using the scale provided below, please fill out the following questionnaire to help us improve the program and your experience at:_____

Strongly Disagree				Strongly Agree
1	2	3	4	5

1. I get timely and appropriate feedback from faculty.
2. Time demands (call schedules, rounds, conferences) are reasonable and allow me to get my work done.
3. Salary and benefits packages are adequate.
4. The scheduled conferences are generally a valuable learning experience.
5. I have far too little leisure time.
6. I am satisfied that the inpatient medical record as currently structured is an effective tool for storage and retrieval of information.
7. I have sufficient personal support from peers.
8. There are too many difficult patient management problems on the wards (AIDS, alcoholism, nursing home placement, etc.).
9. Hospital support services are sufficient to help me care for my patients.
10. I generally have enough time to recover between call days.
11. Junior level residents have too much primary responsibility on the inpatient ward rotations.
12. The caseload in clinic is about right.
13. The training program satisfies my professional needs for training in internal medicine.
14. Clinic support is sufficient to meet my needs in caring for my clinic patients.
15. I have received sufficient counseling from faculty to help with career planning.
16. I often feel "stressed out" or depressed.
17. The caseload on the wards is about right.
18. I often need to perform administrative functions more appropriate to clerical staff.
19. Ward medicine is generally a valuable learning experience.
20. I often feel tired and overworked.
21. I have enough personal support from family and friends.
22. Rotations in the intensive care unit(s) are generally a valuable learning experience.
23. I am satisfied that the outpatient medical record as currently structured is an effective tool for storage and retrieval of information.
24. I rarely have time to read.
25. Full-time faculty members contribute to a great extent to the teaching I've received.
26. The average number of work-ups on call days is reasonable.
27. Junior-level residents have too much primary responsibility for patient care in clinics.
28. I have enough personal support from faculty.
29. Appropriate levels of responsibility for patient care are present at each year in the training program.
30. There is enough clerical and administrative support provided by the program.
31. I receive enough instruction on what is expected of me in each level of my training.
32. The workload is generally excessive on the wards.
33. The patient mix ("interesting" vs "routine," "good teaching case" vs "difficult, or hard to deal with") on Ward Medicine is appropriate.