Evaluating Career Development in School-Based Programs: Performance Assessments

Nancy L. Hutchinson
Queen's University

Abstract

This paper makes a case for using performance assessments in career development interventions conducted in the classroom. Examples are drawn from *Pathways*, a recently developed Canadian program for career development. Limitations of performance assessment are discussed, followed by recommendations for practice and research.

Résumé

Cet article insiste sur l'utilisation d'évaluation des performances lors des interventions sur le développement de carrière conduites dans les classes. Des exemples sont extraits du programme *Pathways*, un programme canadien qui a été développé récemment pour le développement de carrière. Les limitations de l'évaluation de la performance sont discutées, suivi par des recommandations pour la pratique et la recherche.

Recent research by the author in career development has focused on the development and evaluation of an instructional program, called *Pathways* (Hutchinson & Freeman, 1994), used in the classroom by teachers and counsellors. *Pathways* is designed for students with learning disabilities and others at risk for dropping out of school, as well as regular secondary students. Consistent with the general movement toward detracking (Oakes, 1985), and the destreaming of grade 9 in Ontario (Ontario Ministry of Education and Training, 1993), *Pathways* is suitable for mixed ability groups. *Pathways* also responds to recent calls for instruction of authentic tasks based on complex views of learning (Marshall, 1992), and evaluation approaches capable of assessing such learning (Wolf, Bixby, Glenn & Gardner, 1991).

Purpose

The purpose of this paper is to explore recent developments in classroom learning, teaching, and assessment, and to apply these to the assessment of career development in classroom interventions, using examples from *Pathways*. Neither instruction nor assessment can be understood in isolation, as the two are interdependent parts of a system whose purpose is to promote student learning (Camp, 1992). A case is made for using performance assessments (Messick, 1992) in classroom interventions, the limitations of performance assessments are discussed, and recommendations for practice and research are made.

Recent Views of Learning and Teaching

The perspective on teaching and learning that dominated much of this century is teaching as transmission and learning as passive acquisition, derived from the behavioural tradition. Evaluation that has accompanied such teaching and learning has focused on the production and recognition of correct answers, comparisons to norms, and often multiple-choice formats. Evaluation of career development in secondary schools has frequently been conducted with norm-referenced tests of maturity (e.g., *The Career Maturity Inventory*, Crites, 1978) and interests (e.g., Holland, 1985). Career interventions applied to groups of secondary students sometimes consist of transmitting the results of such instruments, "test-and-tell" (Crites, 1981). Peterson, Sampson and Reardon (1991) argued that cognitive psychology offers concepts that overcome limitations in current career counselling.

Recently, much has been written about constructivism wherein learning consists of learners building on what they bring to the situation and restructuring or recasting initial knowledge as increasingly complex understanding (Marshall, 1992). Whereas cognitive constructivists emphasize that learners construct knowledge, social constructivists place greater emphasis on the role of social interaction in meaning construction. Whether students come to see themselves as "sense makers" and problem solvers or "rememberers and forgetters" (Lampert, 1986) is a function of what counts as learning in the particular setting. In constructivist perspectives, teaching means providing challenging tasks that require active involvement by students, options and guidance, and opportunities for shared responsibility. Meaningful tasks are emphasized that are related to real-world knowledge and skills (e.g., Seely Brown, Collins & Duguid, 1989).

Pathways consists of five instructional modules focusing on: (a) awareness of self and careers, (b) employment writing, (c) interview skills, (d) problem solving on the job, and (e) anger management. In each area, students work in pairs and small groups, think aloud providing explanations for their choices and actions, and learn through role-playing and authentic tasks. For example in employment writing, students learn to think about what employers are likely to demand in application forms and learn about the reasons for these demands by role-playing employers who must judge completed application forms and make hiring decisions. They also practise completing application forms and judge their own performance.

Recent Views of Assessment

The dissatisfaction with traditional assessment and the calls for new forms are based, in part, on the concern that traditional assessments do

not yield information critical to learning. They provide highly constrained simulations of the performances in which we want students to engage in the classroom and beyond (Camp, 1992). Standardized, multiple-choice, and short answer evaluations rarely elicit evidence of ability to engage in complex performances, of the processes and strategies used in such performances, or of the mental representations students use to solve complex problems (Snow & Lohman, 1988).

In a study of classroom evaluations in two Canadian provinces, Wilson (1990) found widespread use of multiple-choice and short answer evaluations. He reported a general trend toward using performance measures at the primary level, and completion and multiple-choice measures at the junior/intermediate grades. At the secondary level, subject content and evaluation pragmatics (e.g., appearance on government examinations) influenced selection of assessments. Wilson summarized the use of constrained assessments:

> . . . there is a clear indication that short answers requiring single word or sentence responses characterize much of the evaluation at both levels [elementary and secondary]. Objective performance items (checklists and rating scales) were also commonly used. . . . [Matching items] is the most common selection type of item in widespread use among teachers in this sample, although multiple-choice was the third-ranked item in use by elementary teachers in Ontario. (1990, p. 11)

Beyond the primary grades, teachers used evaluations mainly to generate marks for reporting. In interviews teachers stated these reports usually reaffirmed their judgments about learning, but had little influence on instruction.

Proponents of new approaches argue that assessments should exhibit the characteristics of worthwhile educational experiences (Snow, 1989). The implication is that assessments should be based in meaningful tasks—tasks that are complex and challenging, consistent with goals for learning, closely related to real-world skills, and allow students to use processes and strategies relevant to genuine performance. This may require extended or flexible time frames, open-ended formats, and collaboration with peers (Wolf, et al., 1991). These approaches have been called performance assessments (Messick, 1992).

Because *Pathways* is a classroom intervention to enhance career-related behaviours, cognitions, and metacognitions, the authors (Hutchinson & Freeman, 1994) attempted to develop authentic learning tasks and performance assessments involving these three aspects of performance. They were guided by concerns that teachers, counsellors, and students understand the assessments and use them to guide teaching and learning. They emphasized self-referenced evaluation, and demonstration of student growth, recognizing the controversy and issues that surround performance assessments (Messick, 1992).

Attempting Performance Assessments in Pathways

Each of the five modules in *Pathways* contains performance assessments that may be used as pre- and post-test indicators to show growth. The module "Knowing about Yourself, Knowing about Careers" includes an assessment measure that provides descriptions of individuals seeking employment and "want ads" containing descriptions of jobs. Students decide whether or not applicants are well matched to advertised positions on a number of criteria, and explain their decisions.

The module teaching employment writing ("Succeeding with the Resume and the Application") contains a realistic one-page application form (with scoring criteria) as one assessment measure. For the interview module ("Succeeding with the Interview"), one assessment is a simulated employment interview. Two sets of scoring criteria provide for holistic scoring and a behaviour checklist.

The module "Solving Problems on the Job" contains an assessment measure consisting of five realistic scenarios for which students must provide a number of solutions and evaluations of these solutions. Each scenario describes an incident that students, who participated in the developmental phases, rated as likely to occur and as posing a dilemma. The module "Anger Management on the Job" also contains an assessment measure with scenarios. Students demonstrate that they recognize the signs and feelings associated with anger and that they can apply a strategy for growing calmer and controlling reactions in anger-provoking situations.

Additionally, in the learning activities within the modules, students generate performances with peers and independently, including role playing that can be video-taped, résumés and personal data sheets, solved problem scenarios, and self-awareness activities. These performances can be accumulated in a portfolio. Suggestions for extension involve employers and others from the community in responding to student questions, evaluating applications and videotaped interviews, thus enhancing authenticity. Students practise, evaluate their performance, and receive feedback from videotape, peers, and adults.

Using instruments similar to those contained in the instructional modules, a series of control-group studies has shown that each module is effective alone (e.g., Freeman, Hutchinson & Porter, 1991; Hutchinson, Freeman, Downey & Kilbreath, 1992; Taves, Hutchinson & Freeman, 1992), and the five together make a viable course of instruction with learning and maintenance at as much as five months after the last module (Hutchinson, Freeman & Fisher, 1993). However, questions remain about the performance assessments and how confidence in them can be increased.

Controversy and Issues in Performance Assessments

Messick (1992) analyses the interplay of evidence and consequences in the validation of performance assessments, although he acknowledges that technical standards can be relaxed somewhat in instructional assessment compared to high stakes assessment (p. 2). Messick argues that the proponents of performance assessments must gather evidence to discount the two major threats to construct validity—construct underrepresentation and construct-irrelevant variance—for if they can do this they will have a much stronger case than if they developed specialized validity criteria.

He criticizes performance assessments because they focus on tasks more than on constructs, and educators are rarely concerned with just the particular performance, but rather with the knowledge and skill that enable a whole range of performances. This knowledge and skill becomes the construct at the centre of measurement. Messick assumes that constructs are central.

> A construct-centred approach would begin by asking what complex of knowledge skills, or other attributes should be assessed, presumably because they are tied to explicit or implicit objectives of instruction or are otherwise valued by society. Next, what behaviours or performances should reveal those constructs, and what tasks or situations should elicit those behaviors? Thus the nature of the construct guides the selection or construction of relevant tasks as well as the rational development of construct-based scoring criteria and rubrics. (Messick, 1992, p. 17)

Messick's recommendations show the assumed primacy of the evaluation construct, which is "presumably" related to teaching or valued by society. If we accept that the construct is the most important factor in assessment, then it follows that we want to get full representation of the construct and avoid variance from factors irrelevant to the construct. However, if we begin with the question of what knowledge and skills should be learned, then it seems to follow that we want to get full representation of what was taught and learned. "We must first decide what are the actual performances that we want students to be good at. We must design those performances first and worry about a fair and thorough method of grading them later" (Wiggins, 1989, p. 705). This makes fidelity of assessments to valued performances more important than fidelity of assessments to constructs. This means that when we say we are assessing employment writing, we use the term to mean a small number of simulated tasks, instances of writing applied to securing employment, rather than to mean we are obtaining a representative sample of the generalizable construct of employment writing. Thus, instructional assessment proceeds on the basis of "relaxed" technical standards at the expense of generalizations (Messick, 1992)—through simulations we assess what was valued and taught, rather than general

skills. Messick points out these simulations must detect relevant differences and changes in performance (p. 20). He acknowledges, "Such simulated tasks are authentic in that they replicate the challenges and standards of real-world performances and are representative of the ways in which knowledge and skills are used in real-world contexts, even though they do not simulate all of the complexity of real-world functioning" (p. 21). Beginning with the primacy of learning, performance assessments in instructional contexts should not underrepresent the authentic tasks or be characterized by learning-irrelevant variance, and should detect relevant differences and changes in performance.

Establishing Validity-Directions

What is needed is evidence, systematically accumulated, for the validity of performance assessments—evidence that begins with valued performances (knowledge, skills, and behaviours) rather than constructs, and works toward constructs. Representativeness and freedom from irrelevant variance are empirical questions about which data can and should be collected. The consequences of performance assessments have been defended through arguments based on the similarity between the learning and assessment tasks. However, vigilance for unintended effects will be just as important for performance assessments as for construct-driven assessments. The writings of Messick and other critics of performance assessments provide direction for validation. Proponents agree:

> If we want to pursue these new modes of assessment, we cannot do so on the mere conviction that they are better. We cannot use the notion of developmental accomplishments or holistic scores to excuse us from developing rigorous standards and thoughtful rules of evidence that will offer candid pictures of what students are learning. (Wolf et al., 1991, p. 62)

Closing Comments

Wolf and colleagues argue that our conceptions of learning are changing rapidly, and we must develop a new psychometrics to answer the changed questions about learning. They raise issues of multiple paths to excellence, multiple appraisers, and the need to break with single summary statistics for describing performance. "Unless we develop these kinds of differentiated portraits of student performance within a domain, it is difficult to envision student assessment ever informing, rather than merely measuring, the educational process" (Wolf et al., 1991, p. 63).

These first steps, attempted in *Pathways*, remind us of the importance and difficulty of performance assessments for classroom interventions in career development, and of the need for validation data. Instructional studies must be complemented by classroom-based evaluation research on performance assessments.

References

Camp, R. (1992). Assessment in the context of schools and school change. In H. H. Marshall (Ed.), *Redefining student learning: Roots of educational changes* (pp. 241-63). Norwood, NJ: Ablex.

Crites, J. O. (1978). *The Career Maturity Inventory.* Monterey, CA: CTB/McGraw-Hill.

———. (1981). *Career counseling: Models, methods, and materials.* New York: McGraw-Hill.

Freeman, J. G., Hutchinson, N. L. & Porter, B. (1991). Improving problem solving on the job. *Exceptionality Education Canada, 1,* 45-65.

Holland, J. L. (1985). *Making vocational choices* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hutchinson, N. L. & Freeman, J. G. (1994). *Pathways.* Toronto: Nelson Canada.

Hutchinson, N. L., Freeman, J. G., Downey, K. H. & Kilbreath, L. (1992). Development and evaluation of an instructional module to promote career maturity for youth with learning disabilities. *Canadian Journal of Counselling, 26,* 290-99.

Hutchinson, N. L., Freeman, J. G. & Fisher, C. (1993). *A two-year cohort study: Career development for youth with learning disabilities.* Paper presented at the annual general meeting of the American Educational Research Association, Atlanta, GA.

Lampert, M. (1986). Knowing, doing, and teaching mathematics. *Cognition and Instruction, 3,* 305-42.

Marshall, H. H. (1992). Seeing, redefining, and supporting student learning. In H. H. Marshall (Ed.), *Redefining student learning: Roots of educational change* (pp. 1-32). Norwood, NJ: Ablex.

Messick, S. (April, 1992). *The interplay of evidence and consequences in the validation of performance assessments.* Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.

Oakes, J. (1985). *Keeping track.* New Haven, CT: Yale.

Ontario Ministry of Education and Teaching. (1993). *The common curriculum.* Toronto: Author.

Peterson, G. W., Sampson, J. P. & Reardon, R. C. (1991). *Career development and services: A cognitive approach.* Pacific Grove, CA: Brooks/Cole.

Seely Brown, J., Collins, A. & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Snow, R. E. (1989). Toward assessment of cognitive and conative structures in learning. *Educational Researcher, 18,* 8-14.

Snow, R. E. & Lohman, D. F. (1988). Implications of cognitive psychology for educational measurement. In R. Linn (Ed.), *Educational measurement* (3rd ed.), pp. 263-331. Washington, DC: American Council on Measurement.

Taves, R. A. H., Hutchinson, N. L. & Freeman, J. G. (1992). The effect of cognitive instruction in the development of employment interview skills in adolescents with learning disabilities. *Canadian Journal of Counselling, 26,* 87-95.

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan,* 79, 703-13.

Wilson, R. J. (1990). Classroom processes in evaluating student achievement. *The Alberta Journal of Educational Research, 36,* 4-17.

Wolf, D., Bixby, J., Glenn, J. & Gardner, H. (1991). To use their minds well: Investigating new forms of student assessment. *Review of Research in Education, 17,* 31-74.

About the Author

Nancy L. Hutchinson is an associate professor in the Faculty of Education at Queen's University. She conducts instructional research in career development and mathematics, focusing on students with learning disabilities and others at risk for dropping out of school. She received her doctorate from Simon Fraser University in 1987.

Address correspondence to: Dr. Nancy L. Hutchinson, Faculty of Education, Queen's University, Kingston, Ontario, K7L 3N6.

A version of this manuscript was presented at the symposium on Issues and Solutions for Evaluating Career Development Programs and Services, Mount Saint Vincent University, Halifax, Nova Scotia, March, 1994.