Measuring Perceived Self-Efficacy in Occupational Therapy

Marie Gage, Samuel Noh, Helene J. Polatajko, Violet Kaspar

Key Words: assessment process, occupational therapy

Objectives. Occupational therapists sometimes observe an alarming discrepancy between the occupational performance skills developed in the clinical setting and the degree to which the client willingly puts these skills to use outside the clinical environment. The literature strongly suggests that perceived self-efficacy partially explains this discrepancy; however, an understanding of perceived self-efficacy has not yet been integrated into the practice of occupational therapy. Both a lack of awareness of the construct and a lack of ability to assess perceived self-efficacy within the occupational performance domain are responsible for this lack of integration into practice. This article presents the development, reliability, and validity testing of a measure of perceived self-efficacy developed particularly for use by occupational therapists, the Self-Efficacy Gauge.

Method. Reliability and validity testing was performed through use of a mailed survey (n = 126). Test-retest reliability, internal consistency, and alternate form reliability were examined. In addition, the subjects' scores on the Gauge were compared to actual performance and a hopelessness scale.

Results. The Self-Efficacy Gauge has a reasonable degree of reliability when used on a clinical sample whose perceived self-efficacy for occupational performance activities is unlikely to change. The results also provide preliminary support for the validity of the instrument.

Conclusion. It is recognized that without a gold standard for comparison, validity testing will be an ongoing need.

Perceived self-efficacy as a concept has been a focus of research in social psychology (Bandura, 1977, 1981, 1982, 1986; Bandura & Shunk, 1981; Manning & Wright, 1983), health and counseling psychology (Allen, Becker, & Swank, 1990; Dolce, Crocker, & Doleys, 1986; O'Leary, 1985; O'Leary, Shoor, Lorig, & Holman, 1988), and related health sciences (Ewart, 1989; Ewart, Taylor, Reese, & DeBusk, 1983; Shoor & Holman, 1984). Most recently, the pertinence and salience of this concept has also been introduced, to a limited extent, to the literature of occupational therapy (Conte & Conte, 1977; Crist & Stoffel, 1992; Gage & Polatajko, 1994). Conceived as a psychological property, perceived self-efficacy refers to a person's belief in his or her performance capabilities with respect to a specific task (Bandura, 1977). More precisely, perceived self-efficacy is "concerned with beliefs in one's capabilities to mobilize the motivation, cognitive resources, and courses of action needed to meet given situational demands" (Bandura, Cioffi, Taylor, & Brouillard, 1988, p. 479). As such, perceived self-efficacy is believed to be an important determinant of actual performance, and therefore has been used to improve the
understanding or prediction of actual performance or achievement levels. It has also been suggested that perceived self-efficacy may explain at least partially the inconsistency, observed frequently by occupational therapists, between clients’ acquired skills and the use of the skills outside the learning (clinical) environments (Gage & Polatajko, 1994). Unfortunately, there has been no attempt to develop an instrument that may be suitable for the assessment of perceived self-efficacy for occupational performance activities.

In this article, we present a newly developed scale of perceived self-efficacy, the Self-Efficacy Gauge (the Gauge), and some preliminary results of its psychometric properties. The Gauge was developed to assess the current level of perceived self-efficacy, as well as its change over time, among adults with physical disabilities that may limit competent occupational performance. Occupational performance is defined as the daily occupations in which people engage in order to look after themselves (self-care), to produce meaningful commodities (productivity), or to derive personal enjoyment or satisfaction (leisure) (Christiansen, 1991).

It should be noted that the Gauge was developed to measure perceived self-efficacy at the microanalytic (task specific level) as conceptualized by Bandura (1977). Other authors suggest that perceived self-efficacy can be measured at a generalized level, being defined as a generalized belief in one’s ability to succeed at whatever task one is faced with. However, generalized self-efficacy measures have been found to be less useful in predicting performance and less sensitive to a real change in performance (Tipton & Worthington, 1984; Wang & RieCharde, 1988). A substantial convergence in research findings suggests that when perceived self-efficacy is measured at the microanalytic level it is a strong predictor of performance of the specified tasks, or of tasks that require similar performance skills (Allen et al., 1990; Bandura, 1977; Ewart et al., 1983). Moreover, some have argued that generalized self-efficacy as an empirical construct has not been well defined, and the difference between generalized self-efficacy and self-esteem is a subject of debate in the literature (Gist & Mitchell, 1992). Thus, these findings support the development of a microanalytic measure of perceived self-efficacy for occupational performance.

Construction of the Self-Efficacy Gauge

A preliminary form of the Gauge was developed with the performance areas and activity items of the Sickness Impact Profile (SIP) (Bergner et al., 1976). The SIP was selected to assist in the development of items for the Gauge after an extensive review of literature and scales of health indexes. The SIP is a reliable and valid index of health and well-being that was developed on the basis of information from 1,100 respondents to a survey (Pollard, Bobbitt, Bergner, Martin, & Gilson, 1976). It was important that information was derived from a broad spectrum of professionals in health care as well as clients with a variety of health conditions. In addition, the SIP also adequately encompasses the three subdomains of occupational performance and hence seemed appropriate for generating items to tap perceived self-efficacy in a population manifesting occupational performance deficits. The items of the SIP subscale of Emotion were irrelevant to occupational performance and thus were discarded.

During the construction of the Gauge, consideration was also given to what Law and Letts called “a universal disability framework” (1989, p. 523). Thus, a serious attempt was made to address the needs of clients with a variety of diagnoses, and to cover a wide range of occupational performance activities that were judged (by experts and clients) to be generally applicable to clients of many diagnoses. This approach prevents a proliferation of similar measurement tools of unknown reliability or validity. Thus, issues relevant to specific diagnostic groups have been purposefully omitted.

A concern that the instrument be responsive to clinically significant change also guided the instrument development. Responsiveness refers to the ability of an instrument to detect the minimal, clinically important change in the construct being measured (Guyatt, Walter, & Norman, 1987). It is in essence “the power of the index to detect a change when one is present” (Kirshner & Guyatt, 1985, p. 34). The magnitude of a clinically significant change in perceived self-efficacy has not been determined and a standardized method of increasing or ensuring responsiveness is yet to be developed. However, to a point, an increase in the number of scale points (response categories) is likely to increase the responsiveness of the instrument (Kirshner & Guyatt, 1985). Kirshner and Guyatt recommend a minimum of 7 scale steps for an evaluative measurement tool. Thus, in an attempt to enhance responsiveness, each item of the Gauge is measured on a 10-point scale ranging from 1 to 10, with two anchor points stipulated; 1 for “not at all confident” and 10 for “completely confident.”

In addition to the large scale client input into the generation of items for the SIP, considerable effort was made to incorporate occupational therapy clients’ participation and input into the development and refinement of the Gauge. As noted by Law and Letts (1989), sampling the domain of occupational performance from the perspective of clients experiencing occupational performance dysfunctions is crucial in the design of occupational therapy assessment instruments. Too often, measurement instruments are developed solely on the basis of clinicians’ views, and the occupational performance activities that are most important to clients are left out. In the process of developing and refining the Gauge, clients’ participation was actively recruited. The structure as well as the clarity and appropriateness of the items on the initial version of the Gauge were rated by 12 therapists...
ductivity, and leisure. This review also indicated that the sample questions). It should be noted that the Gauge was designed to measure persons' beliefs in their ability to perform at the competence level, that is, performance at a level that is adequate to meet demands (Kielhofner, 1985). This level was chosen because of the investigator's belief that the most frequent goal of clinical intervention is to facilitate competent occupational performance, not superior competence. Thus, this instrument is likely to have a ceiling effect when used on a population that is not experiencing an occupational performance dysfunction. In operational terms, each item assessed the perceived confidence of doing the specified activity "without the help of another person" but with the understanding that they may "make use of something such as a cane or wheelchair."

Assessing Psychometric Properties of the Gauge

The most important psychometric property for a newly developed instrument is that it be reliable. Thus, this study was designed with reliability testing as a primary objective, but with a desire to begin the validity testing of the instrument. Three forms of reliability were central to the design of this study:

1. Given that one important goal of the Gauge was to assess the improvements or changes in perceived self-efficacy, the stability or test-retest reliability was considered to be a critical component of the evaluation.
2. The internal consistency of the Gauge was also examined to determine whether the 27 Gauge items could be combined to derive a composite score.
3. The alternate form reliability of the Gauge was assessed to determine whether subjects' responses would be consistent between the two forms.

Design issues related to validity testing of the Gauge centered on the fact that there was no gold standard to which the Gauge could be compared. Before 1970, construct validation was a process of testing the 3 Cs: content, criterion, and construct validity (Streiner & Norman, 1989). However, because more instruments have been designed to measure constructs that do not have a gold standard for comparison there has, of necessity, been a proliferation of new techniques for validity testing, such as trait, convergent, and discriminant validity. Construct validation is now considered to be an ongoing process whereby one continually tests hypotheses about the validity of the instrument; the more tests (of null hypothesis) an instrument sustains, the more confidence one gets about the validity of the instrument (Streiner & Norman, 1989). Because the Gauge was designed to measure subjective perception, widely accepted gold standards are not available. In this study, preliminary results on convergent validity are presented. Two constructs were used to examine convergent validity—the sense of hopelessness and actual performance of the activities included in the scale.

Hopelessness refers to the degree to which one holds negative expectancies concerning oneself and one's future life (Beck, Weissman, Lester, & Trexler, 1974). An assumption consistent with self-efficacy theory is that clients with greater confidence in their ability to regain their usual level of competence for daily activities are less likely to develop hopeless feelings. This assumption seems plausible in that the occupational activities listed in either the SIP or the Gauge are believed to be important and desirable from the client's perspective. Thus, a significant negative correlation between the scores of the Gauge and the Hopelessness Scale is anticipated.

As discussed above, perceived self-efficacy for a specific task has been found to be related to the level of actual performance (Bandura, 1977); therefore, a measure of performance is an important source of validation. The empirical data on the correlation between these two constructs appear to vary depending on many factors, such as the subject's recent experience with the task (or activity) and the complexity of the task. The most important factor, however, is the similarity between the task items used in a self-efficacy measurement tool and in the performance instrument (Gist & Mitchell, 1992). Substantial high correlations, ranging from .62 to .72, were reported between specific single parameter behaviors and the performance of the same behaviors (Bandura & Adams, 1977; Bandura et al., 1988). Correlations between instruments using items of the same domain of activity, but not with exactly the same tasks or activities, were considerably compromised, ranging from a low of .44 (Allen et al., 1990) to a high of .5 (Ewart et al., 1983).

Method

Subjects

The Gauge was developed as an occupational therapy assessment tool for clients experiencing occupational performance dysfunctions. As such, the clients who re-
ently acquired major disabilities and those who are actively participating in therapeutic programs constitute the most appropriate target for assessment. However, in these clients, the sense of efficacy is likely to change rapidly (Lorig, Chastain, Ung, Shoor, & Holman, 1989), making it impossible to determine the instruments' reliability. Given that one important purpose of developing the Gauge was for use in evaluative research, the stability over time of the instrument, not of persons, was considered to be a key question. Thus, subjects with limited but stabilized occupational performance dysfunctions were thought to be the most suitable sample.

Adults older than 21 years of age with chronic health problems such as stroke, heart disease, arthritis, and chronic pain were recruited from regional agencies and physicians dealing with such chronic medical conditions. Subjects who were judged by the agency staff members or physicians as unlikely to be currently receiving therapies were selected from the lists. Pretest questionnaires were mailed to a random sample of 260 subjects who satisfied these conditions. Of these, 26 were returned address unknown, thus decreasing the sample to 254. Subjects were asked to complete the questionnaire only if they could complete it without assistance. If they required assistance, they were asked to tick a box on the front and mail the package to the researcher. This exclusion was necessary because a person helping a subject complete the questionnaire might influence the ratings selected. (A total of 24 responses were received with this box ticked.) A reminder post card was mailed if the pretest questionnaire was not received within 2 weeks. An additional 126 pretest packages were returned completed, a response rate of 59% (including those who replied that they were unable to complete the questionnaire without assistance). Of the 126 subjects, as shown in Table 1, more than 20% had arthritis or chronic pain syndrome or both. Nearly one third indicated that they had had either a stroke or a heart attack. Many subjects had more than one of the conditions listed, thus 214 conditions were reported by 126 subjects.

Retest questionnaires were mailed approximately 2 weeks after the date the pretest questionnaire was completed. The retest questionnaire contained the Gauge and a few questions regarding "recent life events and changes" that might have occurred since the completion of the pretest questionnaire. Subjects were asked to complete the retest within 3 days. Of the 126 retest questionnaires mailed, 107 (85%) were received. The age of the respondents (n = 107) ranged from 25 to 93 (M = 58.5) years. There were 61 (49%) male and 63 (51%) female respondents (two had not specified their gender). Time since onset of the major medical conditions ranged from 4 months to 76 years (1 subject was 93 years old and had developed arthritis at a young age).

### Measurement

The pretest questionnaire included a number of measures. As discussed previously, a few were measures of constructs used for the evaluation of convergent validity, whereas other measurements obtained data required for reliability tests.

### Occupational Performance

Few would contend that objectively assessed actual performance is a useful criterion for the validation of a measure of perceived self-efficacy. However, the selection of a survey research design, chosen with a view to maximizing the efficacy of the reliability testing by increasing the possible sample size, made it impossible to observe the subjects' actual performance. Additionally, some of the items would be impossible to observe due to ethical considerations (e.g., when subjects were having sex) and logistic considerations (e.g., when subjects were enjoying themselves). Therefore, as a preliminary investigation, the present study employed a self-reported instrument to measure performance level.

When existing self-administered performance measures were examined, it became apparent that self-administered measures of performance typically ask subjects if they believe that they are able to perform each activity, not whether they have performed the activity. Thus, existing self-administered measures of performance may be tapping perceived self-efficacy rather than performance. It was, therefore, necessary to create a self-report measure of performance for this study called the Performance Checklist.

The Performance Checklist is a 32-item scale developed to tap the performance of activities relevant to the Gauge items. Persons are asked to indicate when they last actually performed each of the items (within the last day, week, month, year). Although the items of the Performance Checklist were derived from the domain of occupational performance, they do not exactly replicate the list of activities of the Gauge. For example, the Gauge question asks the subjects how certain they are that

#### Table 1

<table>
<thead>
<tr>
<th>Medical Condition</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke</td>
<td>28</td>
<td>13.1</td>
</tr>
<tr>
<td>Heart attack</td>
<td>40</td>
<td>18.7</td>
</tr>
<tr>
<td>Heart surgery</td>
<td>17</td>
<td>7.9</td>
</tr>
<tr>
<td>Arthritis</td>
<td>55</td>
<td>22.7</td>
</tr>
<tr>
<td>Pain</td>
<td>48</td>
<td>22.4</td>
</tr>
<tr>
<td>Other</td>
<td>26</td>
<td>12.2</td>
</tr>
</tbody>
</table>

*Note that many respondents reported more than one condition, hence 214 conditions are shown for 126 subjects.

*Other medical conditions reported were back problems (n = 5), neurological disorders (n = 5), heart conditions (n = 5), asthma (n = 5), surgery (n = 4), and miscellaneous (n = 4).
they can “do the things they like to do” was broken into several tasks such as “gardening,” “gone out to do fun things,” “gone shopping,” and so on. However, 11 items are identical on the Performance Checklist and the Gauge. The Performance Checklist was subjected to the same pretesting procedure as the Gauge, and the suggestions from experts and clients were incorporated.

The format and example items of the Performance Checklist are presented in Table 2. As can be seen, the Checklist asks clients to document the last time they actually performed a task or activity. It is important to note that the Checklist did not require subjects to indicate how well they could perform the tasks. Performance scores were derived by allotting one point for each activity that was performed by subjects within the last month. The internal consistency of the scale, as estimated by Cronbach’s alpha, was .86. Scores obtained from the sample (n = 126) ranged from 1 to 32, with a mean of 23.64.

Perceived Ability

As a part of the Performance Checklist shown in Table 2, the subjects were also asked whether they believed that they were capable of performing the activity if required to do so within the next week. The purpose of including this question in the Checklist was to obtain information that may be used as an alternate form of the Gauge. This scale was named Perceived Ability, and a score was derived by counting the number of activities or tasks that the subjects believed that they could perform if required. Scores range from 1 to 32 with a mean of 23.89 and a standard deviation of 7.6.

Hopelessness

In this study, hopelessness was measured by the Beck’s Hopelessness Scale (Beck et al., 1974). The Beck Hopelessness Scale consists of 20 true–false items derived from two sources: a test of attitudes about the future and statements made by clients who were judged as hopeless by clinical experts (Beck et al., 1974). The Hopelessness scores for the present study sample ranged from 21 to 37 (M = 25.17). Coefficient alpha was .89.

Recent Life Events and Changes

There is a substantial accumulation of evidence to support the assumption that an exacerbation of health problems and profound changes in life environment or lifestyle influence one’s confidence in the ability to perform daily activities (e.g., Holmes & Rahe, 1967; Kransnoff, Askenasy, & Dohrenwend, 1982; Turner & Avison, 1992). This evidence suggested that it was important to determine the extent of profound changes that had occurred recently in the lives of our subjects. If a large proportion of the study sample had recently experienced a sudden change in health status or availability of support, it might have resulted in instability of test scores between pretest and retest, because of the subjects’ uncertainty about their performance abilities not due to poor instrument reliability. Thus, it was considered to be important to determine the stability of the subjects’ health status and support systems.

To this end, two measures were put into the survey in an attempt to determine whether the study sample had been stable both before the administration of the first test and during the interval between test and retest. First, subjects were asked to indicate the extent to which they believed their ability to perform daily life activities had changed during the 3 months before the completion of the first questionnaire. Responses were rated on a 5-point Likert scale, ranging from 1, for significant improvement in ability to perform their daily life activities, to 5, for significant reduction of ability. The response “no change” was coded as 3. This question was repeated in the follow-up (retest) questionnaire, asking them to consider only experiences that occurred in the interval between the completions of the two questionnaires (pretest and retest).

Second, a brief list of life events was developed for this study. Events that were indicative of changes in health, disability, or lifestyle were selected from several standardized life events scales (Holmes & Rahe, 1967; Kransnoff et al., 1982; Turner & Avison, 1992). Because existing life events scales tend to contain exhaustive lists of events covering a broad spectrum of life experiences, a scale that dealt with only critical health and life-style items relevant for this study could not be found. At the time of the initial test administration, subjects were asked to check the number of events that had occurred in the previous 3 months. At retest, the subjects were asked to check the number of life events that had occurred in the interval since completion of the first test.

Table 2
Examples of Questions From the Performance Checklist

<table>
<thead>
<tr>
<th>Activity</th>
<th>Perceived Ability</th>
<th>Required? (Circle One)</th>
<th>Activity Done Most Recently (Circle One)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walk from one room to another</td>
<td>Yes</td>
<td>No</td>
<td>day week month year longer</td>
</tr>
<tr>
<td>Climb one flight of stairs</td>
<td>Yes</td>
<td>No</td>
<td>day week month year longer</td>
</tr>
<tr>
<td>Get dressed</td>
<td>Yes</td>
<td>No</td>
<td>day week month year longer</td>
</tr>
<tr>
<td>Do gardening</td>
<td>Yes</td>
<td>No</td>
<td>day week month year longer</td>
</tr>
<tr>
<td>Go out to eat</td>
<td>Yes</td>
<td>No</td>
<td>day week month year longer</td>
</tr>
<tr>
<td>Play cards</td>
<td>Yes</td>
<td>No</td>
<td>day week month year longer</td>
</tr>
</tbody>
</table>
Results

Descriptive Results

The score for the Gauge was derived by calculating a total score for the first 27 items of the scale. Of the 107 subjects who provided a complete set of data (pretest and retest), 5 had missing values on 6 or more of the 27 Gauge items. These cases were discarded from the analyses. Another 30 had one or two items missing. For each subject in this group, the person's average of validly responded items were substituted for the missing items; for some the means of 26 items were used for 1 missing item, and for some others, the mean of 25 item for 2 missing items. Yet another 6 cases had missing responses on more than 2 but fewer than 5 missing items, for these the same mean substitution procedure was used. Given the minimal number of subjects involved in this modification, and relatively small proportion of missing items that required mean substitution (maximum of 5 out of 27, or 18%), this procedure is considered preferable to dropping cases from the analysis (De Vaus, 1990).

The mean scores at pretest \( M_1 = 228.6, SD_1 = 45.6 \) and retest \( M_2 = 233.2, SD_2 = 40.2 \) were not highly discrepant. The distribution of the Gauge scores was slightly negatively skewed. An attempt was made to reduce the skewness by taking the square root or using a log transformation. In both cases the skewness increased. Although this may warrant some caution in the interpretation of results, skewed data are not unusual in psychological and behavioral research, and investigation into the effect of skewed data has determined that one need not assume that the distributions of the two variables are normal when using correlational analysis (Ferguson & Takane, 1989). It is suggested that the selection of a stable sample resulted in higher functioning subjects, and thus some degree of a ceiling effect was experienced.

Convergent Validity of the Gauge

The scores of the Gauge correlated significantly \( r = .44, p < .001 \) with total Performance Checklist scores. Once again, considering only the 11 items that were exactly the same on the Gauge and the Performance Checklist, the correlation increased in the theoretically predicted manner to \( r = .76 \). The Gauge also correlated in the theoretically expected direction with hopelessness \( r = -.47 \).

Although the life events and change scores were used to ensure that the population was not experiencing a real change in perceived self-efficacy between pretest and retest, the results of correlating the Gauge (pretest) scores with the number of life events and the subjects' perception of change in the 3 months before completion of the pretest are interesting. As anticipated, both life events and change scores correlated negatively with the Gauge, \( r = -.24 (p < .05) \) and \( r = -.19 (p < .05) \), respectively. These findings suggest that people who experienced more stressful life events and reported greater deterioration in their perceived ability showed significantly lower levels of perceived self-efficacy. This finding is consistent with the self-efficacy theory.

Reliability of the Gauge

Several forms of reliability coefficients were estimated. First, the stability of the Gauge was estimated with an intraclass correlation between pretest and retest. The intraclass correlation coefficient was .90. The intraclass correlation was considered to be the proper statistic to test the pretest–retest reliability of the Gauge because it detects systematic differences in scores, unlike the Pearson correlation coefficient, which only accounts for random differences in scores (Guyatt et al., 1987).

Internal consistency was assessed by Cronbach's alpha. The estimated alpha coefficient was .94. Statistics also indicated that removal of any of the items would reduce the internal consistency of the scale, thus supporting the retention of all 27 items as one scale rather than creating subscales.

Alternate form reliability was also explored. (In the Performance Checklist (see Table 2) subjects were asked about their Perceived Ability.) Perceived Ability and the Gauge are strongly, positively correlated \( r = .72, df = 111, p < .001 \), with Perceived Ability explaining 51% of the variance in the Gauge score (or vice versa). When only those 11 items of the Gauge and Perceived Ability that were identical were considered in the calculation, the coefficient escalated to .87, sharing more than 75% of variances. The high correlation of the identical items is suggestive of an alternate form of reliability.

It was thought that the occurrence of life events might be affecting the stability of the Gauge in that 54 of the subjects reported the occurrence of at least 1 life event during the 2 weeks between pretest and retest. To examine this possibility, the test–retest reliability was reestimated. The intraclass correlation among those who experienced no events between pretest and retest \( (N = 48) \) was .95, and .86 among those who experienced at least one event during the 2 weeks between the two tests \( (N = 54) \). The discrepancy between the two intraclass correlations is relatively small, and both represent an acceptable reliability (Nunnally & Durham, 1975). The finding, however, suggests that the occurrence of life events may indeed affect the stability of the test score. The finding also suggests that the instrument is sensitive to a real change in perceived self-efficacy. However, one cannot draw conclusions about sensitivity from this study design.

Discussion

The purpose of this study was to develop and test the reliability and validity of a self-efficacy measurement tool that would be useful in occupational therapy clinical and
research settings. The Gauge was designed to measure the confidence with which persons who manifest occupational performance deficits believed they could perform self-care, productivity, and leisure activities. The reliability results (internal consistency, test–retest, and alternate form) collectively suggest that client responses may be consistent over time. However, these results must be interpreted with caution because the sample consisted of subjects with chronic medical conditions. Therefore, the reliability results may not be generalizable to clients with conditions of recent onset. However, test–retest reliability would have been impossible to assess in a more acute population due to the dynamic nature of perceived self-efficacy. In the acute stage of treatment, clients are relearning occupational performance skills, thus their perceived self-efficacy for these tasks may be changing rapidly.

The strong internal consistency among the items of the Gauge (.94) was an unexpected finding. Wang and RiCharde (1988) had found that microanalytic measures of six different activities were not related, in any consistent manner, to one another. In reviewing the Gauge’s items, it became apparent that this commonality may be an artifact of the strong relationship between success with one scale task and success with other scale tasks. That is, the scale tasks could all be considered to be instrumental to occupational performance. For instance, dressing is a specific self-care task, but is also instrumental to being able to go out into the community to work or have fun. Another possible explanation is that physical disability affects one’s ability to perform self-care, productivity, and leisure activities consistently.

Correlational analyses provide preliminary support for the validity of the Gauge. However, further studies examining the relationship of the Gauge to other related constructs are needed. Because the Gauge is designed for use in evaluative research, the next step is to design studies to test the responsiveness of the Gauge to clinical change. For example, subjects could be tested at the beginning and end of treatment. The scores on the Gauge could be compared to changes in actual performance and client’s overall perception of change in perceived self-efficacy. The fact that the test–retest reliability was slightly lower for the group that did not experience a life event between test and retest is suggestive of sensitivity but certainly cannot be considered to support the responsiveness of the Gauge.

Perceived self-efficacy is believed to be a major determinant of behavior (Allen et al., 1990; Bandura, 1977, 1981, 1982, 1986; Conte & Conte, 1977; Dolce et al., 1986; Ewart, 1989; Ewart et al., 1983; Gage, 1992; Gage & Polatajko, 1994; Manning & Wright, 1983; Ol’Reary, 1985; O’Leary et al., 1988; Shoer & Holman, 1984) and as such it is important for clinicians to know whether their clients believe they are able to perform tasks within the occupational performance domain. If therapists are to enhance the performance of their clients they must develop a better understanding of, and response to, the factors that influence performance once clients are discharged from the protected clinical environment (Gage & Polatajko, 1994). Perceived self-efficacy theory is consistent with the root philosophical beliefs of the profession and justifies the task-specific nature of occupational therapy intervention (Gage & Polatajko, 1994). Therapists must understand that acquisition of a skill, in the absence of a belief in one’s ability to perform the skill without the support of the clinician, is not sufficient to improve occupational performance.

Although the Gauge clearly requires further development and more rigorous testing to reach its full potential in predicting performance, it has utility in a clinical setting at this stage of its development. For clients experiencing occupational performance deficits, occupational therapists are encouraged to administer the Gauge and to examine the clients’ responses and compare these responses to the therapists’ assessment of the clients’ abilities. Identified areas of discrepancy should be addressed through processes that are believed to enhance perceived self-efficacy. For a more complete understanding of these concepts, see Gage (1994) and Bandura (1977). The results of this study suggest that the Gauge shows promise as a measure of perceived self-efficacy within the occupational performance domain.

Acknowledgments

This study was supported by a research grant awarded to authors Marie Gage and Samuel Noh by the Health Services Research Fund, Victoria Hospital Foundation, London, Ontario, Canada.

This study was conducted in partial fulfillment of the requirements of Marie Gage’s master of science degree. She thanks R. C. Gardner for intellectual guidance during this process.

Appendix

Sample Questions From the Self-Efficacy Gauge

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>I’d like to know whether you can do everyday activities without the help of another person. It is OK if you carry out an activity with the use of something such as a cane or wheelchair.</td>
</tr>
</tbody>
</table>
| How Confident (Sure) Am I That I Can \_
| Write? \_
| Feed myself? \_
| Look after my family? \_
| Remember the things that I need to remember? \_
| Concentrate on something difficult? \_
| Do the things I like to do? \_
| Enjoy myself? \_
| Make my needs known to others? \_
| Get into a car? \_
| Walk 1 mile? \_

*Note. Complete copies of the Self-Efficacy Gauge are available from the first author.*
References


