Measuring Teachers’ Self-Efficacy
A validation study of a self-efficacy instrument used in the Prospective Analysis of Teachers’ Health

Elin Frögéli
Measuring teachers’ self-efficacy

A validation study of a self-efficacy instrument used in the Prospective Analysis of Teachers’ Health

Sammanfattning


Nyckelord: Psykometri, self-efficacy, socialkognitiv teori, validering

Abstract

The aim of this study was to evaluate the validity of the interpretation of scores on a 12-item instrument proposed to measure teachers’ beliefs of self-efficacy, a central concept in the social cognitive theory of psychology. A hypothesized hierarchal four-factor structure of data was confirmed by confirmatory factor analysis. Expected relationships between scores on the instrument and satisfaction with professional performance, satisfaction with student interaction, work satisfaction, satisfaction with the career choice, experiencing the profession as living up to in beforehand held expectations, and turnover were confirmed by correlation analyses. Regression analyses confirmed expected partial mediation effects of perceived professional demands on the relationship between teachers’ self-efficacy and health. In summary, all psychometric evaluations provided evidence for the validity of the proposed interpretation of scores on the instrument. In addition, through logical analyses regarding the format of items, the source of item content, and relations to a previously
validated measurement of teachers’ self-efficacy, the validity of the instruments content was verified. Though this study could conclude that the instrument provides a valid measure of teachers’ self-efficacy based on evaluations of internal structure, relations to other variables, and instrument content, future studies evaluating the validity of response processes, and the predictive validity of the instrument, would be valuable to further assure this conclusion.

Keywords: Psychometrics, self-efficacy, social cognitive theory, teachers, validation
Measuring teachers’ self-efficacy
A validation study of a self-efficacy instrument used in the Prospective Analysis of Teachers’ Health

Elin Frögéli

Introduction

Social cognitive theory

Human beings are proactive agents with the ability to exercise control over their thoughts, motivation, and action. This is the basic presumption of the social cognitive theory of psychology, a theory for understanding human behavior by combining thoughts from the behaviorist theory and the social learning theory (Bandura, 1977). According to the social cognitive theory, human action, thought and motivation is determined within a triadic system of reciprocal influences between person, behavior and environment (Bandura, 1989). The person is assumed to affect behavior and environment through five basic capabilities (Bandura, 1989). (1) Symbolizing capability enables humans to successfully react, change and adapt to the environment around them by transforming visual experiences into internal cognitive models. Actions symbolically represented in the mind can be tested and evaluated based on earlier experiences before a decision of which action to execute is made. (2) Forethought capability enables self-regulation of behavior based on outcome expectations. Courses of actions with expected positive outcomes may be pursued and those with expected negative outcomes can be avoided. Consequently, the future can act causally on the present by being cognitively represented by forethought. (3) Through observational learning capability skills can be acquired via the observation of others. In this way the danger of failing when attempting an act never before performed is reduced. The more complex the act is, the larger is the risk of failing, and the greater is the dependence on observational learning from models. For observational learning to take place, models must be perceived as competent and akin regarding characteristics of relevance for the performance. (4) Through self-regulatory capability behavior can be regulated based on evaluations of incongruence between performance outcomes and internal self-set standards. The self-regulatory capability plays a central role in the setting of higher goals that drive learning. (5) Self-reflective capability, finally, is the capability to think about and analyze experiences and thought processes. Through self-reflection humans can generate knowledge about themselves and their environment. According to the social cognitive theory, the most valuable type of knowledge that humans can derive from self-reflection are beliefs about ability to exercise control over events that affect their lives. These beliefs are called self-efficacy beliefs, and they are the most important determinants of human action (Bandura, 2001).

Self-efficacy beliefs

Self-efficacy beliefs refer to peoples’ beliefs about their capabilities to exercise control over events that affect their lives (Bandura, 1989). A person’s sense of self-efficacy is based on perceived capabilities to organize and execute the courses of action required to produce desired results (Bandura, 1977). If capabilities appear insufficient, the action will not be executed (Bandura, 2001). In addition, beliefs of self-efficacy determine the amount of effort that will be put into execution, the perseverance in the face of difficulty, and
whether or not failures are motivating or demoralizing (Bandura, 1989). Given appropriate level of skills, no other human characteristic is more important to determine which course of action is executed, and what will be the result of the performance (Bandura, 1977). Beliefs of self-efficacy influence action through cognitive, motivational and affective processes (Bandura, 1989). (1) Cognitive processes enable people to predict the occurrence and outcome of events, and bring about the means necessary to act and to exercise control. Self-efficacy plays a central role in goal-setting and commitment to goal attainment. Goals drive action by providing a standard to which one can measure and evaluate performance. Self-efficacy beliefs furthermore influence action by affecting whether or not cognitions are optimistic or pessimistic, and self-aiding or self-hindering. People with a high sense of efficacy visualize successful performance outcomes that guide their behavior. People with low sense of efficacy visualize negative outcomes that undermine the execution of actions. (2) Self-efficacy beliefs also influence action by determining peoples’ level of motivation. People typically believe themselves to be more efficacious than they really are. This misperception of efficacy is positive since goal-setting will be ambitious, and effort will be put into executions accordingly. As a consequence, action outcomes are likely to be positive, strengthening self-efficacy and future performance. (3) Finally, through influencing thought processes, beliefs of self-efficacy functions as cognitive mediators of stress reactions. When performing or anticipating actions for which perceived efficacy is high, stress reactions are low and actions are successfully executed. When performing or anticipating actions for which self-efficacy is low, stress reactions will be high because of the perceived threat of failing the execution and receiving undesired consequences. These disturbing reactions bring about a change in the course of action to avoid the stressor, or impair the level of performance if execution cannot be avoided (Bandura, 1989).

In addition to affecting action through cognitive, motivational and affective processes, self-efficacy beliefs also affect which actions that will be executed by influencing a person’s selection of environment (Bandura, 1989). Just as people choose courses of actions they believe they can control, they choose environments and situations that they perceive themselves as capable of handling efficaciously (Bandura, 1989).

*The specific nature of beliefs of self-efficacy*

Beliefs of self-efficacy are not global and general but task- and context-specific (Bandura, 1997). The same person may have a high sense of self-efficacy to write a speech, but a low sense of self-efficacy to deliver the very same speech to a public audience. The effect of self-efficacy beliefs on performance therefore cannot be understood without considering the situation and the constraints of the action in question. Beliefs of self-efficacy are also specific to the different roles that people hold (parent, partner, professional et cetera) (Bandura, 1997). A social worker may feel confident about his or her capability to handle difficult interpersonal relations when at work and in the role of a professional, but less so when with friends and family outside of the professional role.

Though their context and task specific nature, beliefs of self-efficacy to execute diverse courses of actions can co-vary (Bandura, 1997). This may be the case when different domains of actions are regulated by similar sub-skills (Bandura, 2006). The ability to diagnose task demands, to construct and evaluate alternative courses of action, to set goals, and to initiate and sustain effort are self-regulatory skills that guide different types of successful performances and may cause variable beliefs of self-efficacy to co-vary. Co-variation of self-efficacy beliefs may also be the result if the beliefs are developed at the same time and space (Bandura, 2006).
Development of self-efficacy beliefs

Beliefs of capability to successfully use available skills to execute actions are moldable and develop through time and experiences (Bandura, 1977). The development of efficacy expectations is primarily dependent on the cognitive evaluation of information about personal capabilities from four different sources. (1) By successfully executing a challenging task a person receives direct information about personal capability. Enactive mastery experiences provide the most potent information for development of efficacy beliefs. (2) By observing competent models act and be rewarded for their performance, indirect information about personal capability is gained. The more alike the person and model are in task-relevant characteristics, the more valuable the observation will be for the development of self-efficacy beliefs. (3) Verbal reassurance about one’s level of personal capability to execute an action strengthens self-efficacy beliefs when provided by someone who is trusted and viewed as competent for the action in question. (4) If the execution of an action is accompanied by feelings of stress that is interpreted as a sign of low personal capability this will hinder the performance and have negative effects on the development of self-efficacy beliefs. Psychological and emotional arousals thus influence the development of self-efficacy beliefs by affecting performance and evaluations of efficacy information (Bandura, 1977).

Measuring self-efficacy

Self-efficacy is a psychological construct; an informed, scientific idea that is developed to describe or explain behavior (Cohen & Swederlik, 2004; Netemayer, Bearden, & Sharma, 2003). Psychological constructs, also called latent variables, cannot be observed, measured or evaluated directly (Brown, 2006; Cohen & Swederlik, 2004; Netemayer et al., 2003; Raykov & Marcoulides, 2006). Therefore, assessment thereof is dependent on evaluations of specific behaviors that are believed to be indicative of the construct in question (Cohen & Swederlik, 2004; Raykov & Marcoulides, 2006). These indicators are usually measured using instruments such as tests, scales and self-report (Brown, 2006; Raykov & Marcoulides, 2006). It is assumed in test theory that a single score on an instrument based on heterogeneous indicators (i.e. items) reflect whatever is common among the items (i.e. in this case, the construct of self-efficacy). The instruments are thus expected to measure data of a one-dimensional nature (Gustafsson & Åberg-Bengtsson, 2010). Beliefs of self-efficacy are measured by self-report scales (Bandura, 2006). Items on the scales inquire about beliefs of capability to coordinate skills to attain desired goals in particular domains and circumstances. The strength of efficacy beliefs is recorded on a unipolar 100-point scale with 10-unit intervals. The lowest record of 0 signifies that the respondent believes he or she cannot do the task in question at all. A record of 50 signifies that the respondent is moderately certain that he or she can successfully execute the task, and a record of 100 indicates that the respondent is highly certain of his or her capability (Bandura, 2006).

Self-efficacy in relation to other variables

As the core of human agency, beliefs of self-efficacy are expected to be central for psychological adjustment, psychological problems, and physical health (Maddux, 2002). Self-efficacy has, amongst other things, been found to relate to: individual performance (Bandura, 1977), goal-setting and the level of stress and depression experienced in threatening or taxing situations (Bandura, 1997), experienced depression in relation to chronic illness (Sacco et al., 2005), the success of therapeutic interventions (Bandura, 1977), the performance of school children (Bandura, 1997; Klassen, 2010), life-style choices (Ayotte, Margrett, & Hicks-Patrick, 2010; Bandura, 1997; Ferrier, Dunlop, &
Blanchard, 2010), career choices (Miller, Roy, Brown, Thomas, & McDaniel, 2009), professionals’ performance (Bandura, 1997; Cherniss, 1980a, 1993), collaborative practice of professional working units (Le Blanc, Schaufeli, Salanova, Llorens, & Nap, 2010), and the consequences of repeated or prolonged exposure to external demands (i.e. stressors) (Bandura, 1997; Cherniss, 1980a, 1993). These relationships should however not necessarily be interpreted as implying that self-efficacy causally effect the other variables. It is possible that the causality goes in the other direction (i.e. from the other variable to self-efficacy), or that another variable causally effect self-efficacy as well as other variables.

**Self-efficacy in relation to occupational health**

The prolonged and repeated exposure to stressors, and the detrimental effects thereof, has been extensively studied within the field of occupational health psychology where it has been related to the processes of burnout and turnover (Cherniss, 1980a, 1993; Gustavsson, Hallsten, & Rudman, 2010; Karasek & Theorell, 1990; Schaufeli, Leiter, & Maslach, 2009). High demands, low control and lacking social support are stressors that repeatedly have been shown to have detrimental effects on health in the professional population (Cherniss, 1980a, 1993; Karasek & Theorell, 1990; Schaufeli et al., 2009). Stressors that are perceived as uncontrollable have proven to be especially detrimental to health (Miller, Chen, & Zhou, 2007). In addition, Cherniss (1980a) found that many professionals during a period at the beginning of their careers experienced high levels of role-related stress and anxiety. The primary stressor for these new professionals was a sense of insecurity about competence and uncertainty about performance. Even though they had gone through many years of formal schooling, numerous of the new professionals did not perceive themselves as competent enough to successfully handle the various demands of the profession. Cherniss came to call this a “crisis of competence”. To cope with the stressful situation the new professionals typically worked much overtime, adopted less ambitious goals, restricted their personal involvement in their jobs, shifted responsibility for shortcomings from themselves to factors out of their control, became less idealistic and trusting and more “objective” and “professional”, and increased their concern of self-protection and self-enhancement. These means of coping with the stressors of the new profession often lead to a process of burnout characterized by emotional exhaustion, depersonalization and strengthened feelings of inefficacy, as well as turnover. Professionals who were successful in dealing with the stressors of the new profession, and who did not enter a burnout process, characteristically had a more realistic perception of their level of competence and the demands that they were to encounter in their new profession (Cherniss, 1980a). Cherniss (1980a) concluded, as did Bandura (1977), that when people believe that they can cope effectively with stressors, the situations are not perceived as threats and instead are dealt with effectively. Accordingly, as an essential mediator of psychological and physiological processes in response to stressors, self-efficacy has been proposed to be a central component in the relation between prolonged exposure to stressors and health (Bandura, 1997; Cherniss, 1993). The perception of competence based on cognitive processing of available resources protect against ill health because physiological reactions are short-termed, adaptive, and provide resources that enable effective coping (McEwen, 1998).

**Professionals’ self-efficacy beliefs**

Building on these conclusions about the relation of self-efficacy, perceived demands, and health, it has been suggested that the development of a sense of competence is one of the most important tasks for new professionals to be able to perceive demands of the
profession as manageable challenges and avoid ill health (Cherniss, 1980a). Beliefs of professional self-efficacy refer to beliefs about capabilities to organize and execute the courses of actions required to produce desired results within the frame of the profession (Tschannen-Moran & Woolfolk Hoy, 2001). Beliefs of self-efficacy are most malleable early in learning (Bandura, 1977), and once established the beliefs appear to remain unchanged unless persuasive and conflicting evidence causes them to be re-evaluated (Bandura, 1997). Ensuring the development of a strong sense of efficacy early in the professional career can therefore be expected to have longstanding positive effects (Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998). This makes the development of professionals’ beliefs of self-efficacy an interesting intervention target to ensure job satisfaction and to battle burnout and turnover. The theoretical framework of professionals’ self-efficacy is however lacking concerning the development of self-efficacy beliefs during the initial period of the career. It would therefore be valuable to follow a group of professionals in education, training, and during their early years in the working field with the objective to map the development of self-efficacy. Interesting questions to answer concern the new professionals’ level of self-efficacy when leaving the educational setting, how beliefs of self-efficacy develop in the transition from formal education to working life, and how different situations affect beliefs of self-efficacy early in the professional career.

In the prospective longitudinal PATH-study (Prospective Analysis of Teachers’ Health) roughly 2000 Swedish teachers were followed with annual surveys from their second-last year of formal training to their third year in the professional field (Gustavsson, Kronberg, Hultell, & Berg, 2007). The surveys contained questions about teaching preparation, professional situation, family situation, coping behavior, physical and mental health, demographics, et cetera. The surveys also contained 12 items intended to measure teachers’ self-efficacy. The PATH-study with its large sample and longitudinal nature thus provides the possibility to expand the knowledge of professionals’ self-efficacy by analyzing the development in the transition from formal education to the working field in a cohort of Swedish teachers.

**Teachers’ self-efficacy beliefs**

Teachers’ sense of self-efficacy refers to teachers’ beliefs about their capabilities to bring about desired outcomes of student engagement and learning, even among students who may be unmotivated or difficult (Tschannen-Moran & Woolfolk Hoy, 2001). Teachers’ behavior in the classroom, the goals they set as teachers and their effort to reach those goals are predicted to be affected by their beliefs of self-efficacy (Tschannen-Moran et al., 1998). In addition, teachers’ beliefs of self-efficacy are expected to have consequences beyond those related to the individual professional’s own performance and health. Indeed, teachers’ sense of efficacy have been found to be positively related to students learning, controlling for actual competence (Bandura, 1997), teachers’ level of experimentation in teaching, and teachers’ willingness to try new strategies in order to better meet the needs of students and facilitate learning (Tschannen-Moran et al., 1998). In addition, teachers who have a high sense of efficacy have been shown to create mastery experiences for their students in order to strengthen students’ perception of their personal abilities (Bandura, 1997), and persists in helping students who have trouble learning (Tschannen-Moran et al., 1998) viewing learning problems as challenges that can be mastered by extra effort and ingenuity (Bandura, 1997). Teachers’ self-efficacy have furthermore been associated with teachers’ experience of stress when students are not behaving properly, and the way teachers criticize students for making errors (Tschannen-Moran & Woolfolk Hoy, 2001; Tschannen-Moran et al., 1998; Tsouloupas, Carson, Matthews, Grawitch, & Barber, 2010), as well as teachers’ organizational and planning
skills, fairness, clarity, enthusiasm about teaching, and commitment to the profession (Tschannen-Moran et al., 1998). Teachers with a lower sense of efficacy are more likely to leave the profession (Tschannen-Moran & Woolfolk Hoy, 2001). This is true for experienced teachers and teachers at the beginning of their careers (Bandura, 1997; Tschannen-Moran & Woolfolk Hoy, 2001). Novice teachers who feel efficacious in their professional role express higher satisfaction about their work, and a more positive attitude about staying in the field of teaching than new teachers who doubt their professional capabilities (Tschannen-Moran et al., 1998). Teachers’ professional sense of efficacy is furthermore negatively related to the level of stress experienced by both novice and senior teachers (Bandura, 1997). Swedish researchers have found that teachers’ assessment of self-efficacy is positively related to students’ learning (Skolverket, 2006). The higher the Swedish teachers’ self-efficacy, the more successful they are at motivating students (Skolverket, 2006). Teachers’ evaluation of self-efficacy is also related to the extent to which they find the learning environment to be positive or negative, and their perception of disturbing sounds and activity (Skolverket, 2006).

**Measuring teachers’ self-efficacy beliefs**

As previously stated, self-efficacy is a psychological construct that can only be measured indirectly by assessing behavior that is assumed to be indicative of the construct (Cohen & Swederlik, 2004; Netemayer et al., 2003). Furthermore, beliefs of self-efficacy are task specific, and measurements must relate to this specific nature (Bandura, 2006). It is however possible to obtain a measure of professionals’ self-efficacy by a combined assessment of beliefs of self-efficacy to execute a number of diverse tasks that are of central importance to the profession (Tschannen-Moran & Woolfolk Hoy, 2001). A recently developed instrument called the Ohio State teacher efficacy scale (OSTES) measures teachers’ self-efficacy with 24-items. The items tap a wide variety of significant tasks within the profession. For instance, there are items inquiring about the perceived capability to adjust lessons to individual needs, handle learning difficulties and motivate student engagement and interest. The content of all items in OSTES have been assessed and approved as relevant by a panel of teachers. The items form a hierarchal three-factor structure. The three factors efficacy for student engagement, efficacy for instructional strategies and efficacy for classroom management all load on a common factor assumed to be the core concept of self-efficacy (Tschannen-Moran & Woolfolk Hoy, 2001).

The OSTES is thoroughly evaluated in relation to American conditions of the teaching profession (Tschannen-Moran & Woolfolk Hoy, 2001). However, it is not validated in a Swedish context. In fact, no instrument or items to measure teachers’ beliefs of self-efficacy adapted to Swedish conditions were available when the PATH-survey was constructed. Therefore, aiming at contributing to the theoretical framework of self-efficacy, the 12-items intended to measure teachers’ self-efficacy were developed (Gustavsson et al., 2007). The content of the items were explicitly designed to measure novice teachers’ self-efficacy. Before analyses of the development of self-efficacy and the effects of efficacy expectations early in the career can be conducted on data of the PATH-study it must however be ensured that the instrument is a valid and reliable measurement of teachers’ sense of self-efficacy in Sweden. Consequently, the purpose of this study was to carry out a process of validation of the instrument developed within the PATH-study to measure Swedish teachers’ sense of self-efficacy.

**The process of validation**

Validity refers to “the degree to which evidence and theory supports the interpretations of test scores entailed by proposed uses of tests” (American Psychological
The process of validation is a process of hypothesis testing where information are sought and evaluated in order to accumulate evidence for the validity of the proposed interpretations of test scores. The first step in the validation process is providing the interpretation of test scores (American Psychological Association, 1999). In this particular case where the instrument is supposed to measure teachers’ self-efficacy the proposed interpretation is that a high score on the instrument reflects a high sense of professional self-efficacy, whereas a low score reflects a low sense of efficacy. The higher the perceived self-efficacy, the higher the likelihood that a course of action will be chosen, initiated and successfully executed (Bandura, 2006). The rational for the development of this instrument of teachers’ self-efficacy within the PATH-study is to make possible an increase of the knowledge of professional self-efficacy by examining the development in the transition from formal education to working life.

**Validity evidence based on the internal structure**

An analysis of the internal structure of an instrument is an evaluation of the extent to which the relations between the items of the instrument comply with the theory of the construct on which the proposed interpretations are based (American Psychological Association, 1999). Multiple item instruments intended to measure psychological constructs, such as the PATH-instrument, must be carefully developed in order to ensure the relationship between the test score and the construct in question (Cohen & Swederlik, 2004; McDonald, 1999; Netemayer et al., 2003; Raykov & Marcoulides, 2006). According to classic test theory, the score is supposed to reflect the correlation of items that is assumed to be explained by the construct (Schmidt & Embretson, 2003). When controlling for the construct, the items should share no additional variance (Gorsuch, 2003). In addition, items are expected to have uncorrelated errors that are specific, as well as random (Schmidt & Embretson, 2003). This type of instruments using heterogeneous items to assess one-dimensional data are recommended to be addressed using hierarchal factor models (Gustafsson & Åberg-Bengtsson, 2010).

The items included in the PATH-instrument is assumed to measure teachers’ self-efficacy by assessing beliefs of self-efficacy in 12 different areas of central importance to the teaching profession.

Given the theory of self-efficacy, the characteristics of the PATH-instrument, and the assumptions about data from classic test theory, in this study it was hypothesized that the relations of scores on the PATH-instrument are reflected by a hierarchal four-factor model. All items were expected to be related to each other reflecting the one-dimensional structure of self-efficacy. In addition, based on item content, four sub factors reflecting efficacy for instructional strategies, efficacy to give special support to individual students, efficacy for classroom management, and efficacy for teacher-parent interaction were expected.

**Validity evidence based on relations to other variables**

Relationships between scores on the instrument and scores on other tests expected to measure the same construct, a related construct, or a different construct, provide a second source of validity evidence (American Psychological Association, 1999). Based on the knowledge of the construct, hypotheses about expected relationships can be formulated and examined (American Psychological Association, 1999).

In this study, scores on the PATH-instrument were hypothesized to be positively related to satisfaction with professional performance (i.e mastery), satisfaction with student interaction, work satisfaction, satisfaction with the career choice, and experiencing the profession as living up to in beforehand held expectations. On the contrary, beliefs of self-efficacy were
expected to be negatively related to intention to quit the profession. Relations to mastery and satisfaction with student interaction were expected to be especially strong because they are conceptually closer to the general construct of self-efficacy and the specific construct of teachers’ self-efficacy than the other variables.

In addition, teachers’ beliefs of self-efficacy as measured by the PATH-instrument were hypothesized to be related to health via the effect of beliefs of self-efficacy on perception of professional stressors.

Validity evidence based on test content

The content of an instrument must adequately relate to the construct the instrument is proposed to measure (American Psychological Association, 1999). The content of the instrument concerns the themes, wording and format of the items included (American Psychological Association, 1999). Validity evidence based on test content may be accumulated through logical or empirical analyses (American Psychological Association, 1999). Validity evidence based on test content may also be obtained from expert judgments of the relation between the items of the instrument and the construct in question (American Psychological Association, 1999). Experts can furthermore value the representativeness of the items included in the instrument, and whether or not they are suitable to adequately examine all aspects of the construct.

Validity evidence based on response processes

Response processes is a final source of information to investigate to ensure that the proposed interpretation of scores is valid (American Psychological Association, 1999). This information can be gathered by observing, monitoring or interviewing respondents about thought processes and strategies for responding, and how the meaning of items was interpreted (American Psychological Association, 1999). In the case of self-efficacy beliefs, it is important to conclude that the instrument really measures respondents’ perceived capabilities to exercise control. The assessment of teachers’ sense of efficacy must reflect what teachers believe they can do in a given situation, not what they believe they will do (Bandura, 2006). However, no information was available about the response processes of the respondents in the PATH-project. Consequently, this source of validity evidence was not examined in this validation study.

Purpose

The purpose of this study was to examine information about properties of the PATH-instrument regarding its internal structure, relations to other variables, and content. The aim of this examination was to conclude if scores on the instrument validly can be interpreted as a measure of Swedish teachers’ beliefs of self-efficacy.

Method

Subjects

For the present study data were taken from the second follow-up of the PATH-study, one year after graduation. Specifically, data of those 1,489 respondents answering a minimum of 8 out of the 12 items inquiring about perceived professional capabilities were used (53% of the total number of 2798 respondents at baseline). The median age for the selected sample was 28 years and 86.4% were women.
An attrition analysis using logistic regression with attrition (versus responding) as dependent variable; and sex, age, and age of students (younger versus older), during the first wave of measurement as independent variables showed that males (OR=0.75; \( p < .001 \)) and younger participants (OR=0.98; \( p < .001 \)) were more likely to not participate in the second follow-up of the PATH-study. However, the amount of explained variance in attrition was only 1.6%, indicating that it was not very likely that this had any considerable effect on the generalizability of the results of the study.

**Ethical concerns**

The Research Ethics Committee at Stockholm’s University had granted permission to carry out the PATH study. All participants had given consent and were explicitly informed about their freedom to terminate their participation in the study at any time if so was desired. Before analyses of data were conducted, identification numbers of participants were transformed into specific study numbers, making the identity of participants unknown.

**Instrument**

The PATH-instrument purposed to measure teachers’ beliefs of self-efficacy consists of 12 items. The reliability of the items on the instrument is high as indicated by a Cronbach’s \( \alpha \) of .92. Each item inquires about respondents’ certainty to successfully perform task that are central to the teaching profession. Respondents rate their level of certainty on a scale ranging from 0 (cannot do at all) to 100 (highly certain can do). The content of the items is based on the competencies required to successfully function as a teacher in Sweden as defined by a panel of experts at the Swedish Ministry of Education and Research.

**Data analysis**

The process of validation is a process of hypothesis testing where information about properties of an instrument is evaluated. The purpose of the validation process is to accumulate evidence for the validity of the proposed interpretations of test scores (American Psychological Association, 1999). Three sources of validity evidence for the proposed interpretation of scores on the PATH-instrument were analyzed in this study: internal structure, relations to other variables, and content.

**Validity evidence based on the internal structure**

Validity evidence for proposed interpretations of scores based on internal structure is provided if proven that items of the instrument relate to one another as expected (American Psychological Association, 1999). In this study, the hypothesis that relations of scores on the instrument may be reflected by a hierarchal four-factor model was examined by confirmatory factor analysis using LISREL Version 8.80 (Jöreskog & Sörbom, 2006). The 12 items (translated to English), and the expected factor structure, is presented in Table 1.
Table 1. Expected factors in the hierarchical four-factor structure

<table>
<thead>
<tr>
<th>Factors</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1: Efficacy for instructional strategies</td>
<td>1 … use your knowledge of the subjects so that students learn and develop?</td>
</tr>
<tr>
<td></td>
<td>2 … organize and carry out work so that each student develops according to his or her potential?</td>
</tr>
<tr>
<td></td>
<td>3 … analyze and evaluate student learning and development?</td>
</tr>
<tr>
<td>Factor 2: Efficacy to give special support to individual students</td>
<td>4 … give special support to pupils with learning difficulties of any kind?</td>
</tr>
<tr>
<td></td>
<td>6 … give special support to pupils who live in a socially difficult situation of any kind?</td>
</tr>
<tr>
<td></td>
<td>12 … motivate students who show a lack of interest in their studies?</td>
</tr>
<tr>
<td>Factor 3: Efficacy for classroom management</td>
<td>5 … create a good working climate in the student group?</td>
</tr>
<tr>
<td></td>
<td>7 … actively discourage bullying, harassment and abuse among students?</td>
</tr>
<tr>
<td></td>
<td>11 … deal with unexpected demands that affect the teaching situation?</td>
</tr>
<tr>
<td>Factor 4: Efficacy for teacher-parent interaction</td>
<td>8 … carry out development discussions in order to promote students' cognitive and social development?</td>
</tr>
<tr>
<td></td>
<td>9 … lead parent-teacher meetings that invite the parents to participation and engagement?</td>
</tr>
<tr>
<td></td>
<td>10 … carry out discussions with parents who are rooted in some sort of problem with student?</td>
</tr>
</tbody>
</table>

A confirmatory factor analysis is a theory-driven statistical method for examining relations among items based on the linear model. In order to conduct a confirmatory factor analysis, the investigator must have a firm theoretical idea of the expected relations of items. Based on this idea, a measurement model is formulated indicating how many factors that exist in data, which items related to which factors, if items are related to one another, et cetera (Brown, 2006; McDonald, 1999; Netemeyer et al., 2003; Raykov & Marcoulides, 2006). As defined in Brown (2006), a factor is “an unobservable variable that influences more than one observed measure [indicator or item] and that accounts for the correlations among these observed measures [indicators or items]” (p. 13). As is assumed in classical test theory, and evaluated in the method of factor analysis, when controlling for the shared variance of items explained by the common factor there should remain no correlation between the items (Brown, 2006; Raykov & Marcoulides, 2006). The hypothesized measurement model, and rules for variance and covariance of linear combinations of variables, makes it possible to create a reproduced variance-covariance matrix where the theoretically expected relations among items are represented by different parameters. In the next step of the confirmatory factor analysis, the parameter estimates (factor loadings, error variance, factor variance, et cetera) that most precisely reproduce the variance and covariance of observed data are sought using one of a number of estimation methods. The aim of confirmatory factor analysis is to find estimates for the parameters in the proposed measurement model that produce a reproduced variance-covariance matrix that resembles the data variance-covariance matrix as closely as possible. If it is possible to interpret the relations among observed data as proposed in the measurement model, parameter estimates may be computed, and the model is said to fit the data. The better the model fit the observed data, the stronger the support for the proposed interpretation of scores on the instrument (Brown, 2006; Raykov & Marcoulides, 2006).
Data analyzed by confirmatory factor analysis should be on interval level at least (Brown, 2006; Raykov & Marcoulides, 2006). Because items in the PATH-instrument are ordinal variables without origins or units of measurement polychoric correlations and asymptotic covariance were estimated using PRELIS Version 2.80 before the confirmatory factor analysis was performed, as suggested by Jöreskog (2005). When doing so it is assumed that each ordinal variable (i.e. item) has an underlying continuous scale ranging from $-\infty$ to $+\infty$ that can be used in the confirmatory factor analysis instead of the ordinal variable. A metric is assigned to the ordinal variable by the underlying continuous variable (Jöreskog, 2005). Thus, by using polychoric correlations to investigate the relationships among ordinal variables, biased model fit statistics due to violations of the assumptions of the method may be avoided (Brown, 2006; Flora, Finkel, & Foshee, 2003).

It is recommended that the evaluation of the significance of the model fit is based on multiple sources of information about the model (Brown, 2006; Raykov & Marcoulides, 2006). An often used indicia of the resemblance of the reproduced variance-covariance matrix and the data variance-covariance matrix is $\chi^2$. Other common statistics for this evaluation are SMRM (standardized root mean square residual; the average difference between the correlations observed in the data variance-covariance matrix and the correlations predicted by the proposed measurement model), RMSEA (root mean square error of approximation; the extent to which a model fits reasonably well to the data variance-covariance matrix) and CFI (comparative fit index; the ratio of improvement in noncentrality when moving from the null to a considered model, to the noncentrality of the null model) (Brown, 2006). SMRM is an indicator of absolute fit, RMSEA an indicator of parsimony, and CFI and indicator of comparative fit. A good model fit is indicated by a SMRM value of .08 or lower. RMSEA should have a value of .06 or lower. It has however been suggested that RMSEA values up to .08 may be interpreted as adequate model fit (Brown, 2006), and that values ranging from .08 to .10 suggest mediocre fit not to be rejected (MacCallum, Browne, & Sugawara, 1996). CFI values should be .95 or higher to indicate a good model fit (Brown, 2006). These statistics provide global indications of model fit based on the difference between the data variance-covariance matrix and the reproduced variance-covariance matrix (Brown, 2006; Raykov & Marcoulides, 2006). Specific information about how adequately the model’s parameter estimates reproduces each variance and covariance in data is provided by the residual matrix of the confirmatory factor analysis model. By examination of the standardized and unstandardized residuals information of localized areas of ill-fit may be attained which is important for the model evaluation (Brown, 2006; Raykov & Marcoulides, 2006). Localized ill fit may be the result if items relate more strongly to another factor than the one proposed in the measurement model, as well as if items share variance that is not explained by the common factor.

Validity evidence based on relations to other variables

Validity evidence for the proposed interpretation of scores based on relations to other variables is obtained if proven that scores on the instrument relate to other variables in a manner expected from theory and previous research. In this study, hypothesized relationships between the PATH-instrument and other variables were investigated by analyses of correlation and multiple regressions using PASW Statistics Version 18.

A correlation analysis is a statistical method that examines the relationship between two variables based on the linear model (Cohen & Swederlik, 2004). The correlation coefficient ($r$) indicates the degree of co-variation of the variables and varies from $-1$ to $+1$. The more the coefficient $r$ deviates from zero, the stronger is the relationship (Cohen & Swederlik, 2004). A coefficient of .10 is indicative of a weak correlation, .30 is indicative of a medium correlation, and .50 is indicative of a strong correlation (Cohen, 1988). The
significance of the correlation coefficient can be examined by testing the probability of the attained value given the case of a null correlation between the variables (Cohen & Swederlik, 2004). However, the evaluation of the significance of correlation coefficients are dependent of sample size and it is thus important to take into consideration that with a large enough sample all coefficients will pass as statistically significant (Allison, 1999). In this study, the significance of the difference between two correlation coefficients was examined using the Fisher r-to-z transformation 1.

The meaning of the relationship is further interpreted by the positive or negative sign that accompanies the correlation coefficient. A positive sign indicates that the value of one variable increases as the value of the other increases. A negative sign indicates that the value of one variable decreases as the value of the other increases. No conclusions about causality can be drawn from analyses of correlation (Cohen & Swederlik, 2004).

The correlation statistics Pearson’s r was used in the analyses in this study (Cohen & Swederlik, 2004). This statistics assumes data to be at least on interval level, but is often used for analyses of ordinal items such as those in the PATH-instrument (Allison, 1999). Teachers’ beliefs of self-efficacy as measured by the PATH-instrument were hypothesized to be positively related to variables concerning satisfaction with professional performance (i.e. mastery), satisfaction with student interaction, work satisfaction and career choice satisfaction, as well as experiencing the profession as living up to in beforehand held expectations. On the contrary, beliefs of self-efficacy were expected to be negatively related to intention to quit the profession. These variables were operationalized using five scales (Mastery, Interaction, Work satisfaction, Met expectations, and Intention to quit) computed from single items from the PATH-survey. Some items had to be reversed before the scales could be computed to ensure the psychometric properties of the scales. Content, source, range, mean, standard deviation and Cronbach’s α of the scales are presented in Table 1 in Appendix.

Regression analysis is a statistical method used to study relationships between a single dependent variable and one or more independent variables (Allison, 1999). The difference from correlation analysis lies in the ability of multiple regression analysis to include more than one independent variable. As correlation analysis, multiple regression analysis assumes that the value of the dependent variables is a linear function of the value of the independent variable plus random error. The variance of the random error is assumed to be independent on the variance of the dependent variable, which is acknowledged as the assumption of homoscedasticity. Furthermore, the random errors of different individuals in the sample are assumed to be uncorrelated and the random errors across all individuals are assumed to have a normal distribution. In addition, the possibility of the problem of multicollinearity needs to be addressed when performing multiple regression analyses. Multicollinearity refers to high correlations between independent variables that make it impossible to separate their effects on the dependent variable. The best measure of multicollinearity is the statistics called tolerance associated with each independent variable. This statistic should not drop below .40 (Allison, 1999). If these assumptions are not violated, the regression coefficients called slope (b) and intercept may be calculated using the ordinary least square method. If there is no relationship between the independent and the dependent variables, the coefficient b will equal zero. Testing the probability of this coefficient given a hypothesized effect of zero thus provides an estimate for how good the regression model is. An evaluation of the regression model can also be done by dividing the coefficient b with the standard error (SE). If this value is over 1.96 the relation between the variables can be interpreted as statistically significant. As the correlation coefficient r, the

1 http://faculty.vassar.edu/lowry/rdiff.html
analysis of the significance of the coefficient \( b \) is dependent on sample size and with a large enough sample all coefficients will appear significant. In these cases, the standardized coefficient \( \beta \) (beta) is useful to determine the magnitude of the results since it is not dependent on sample size. The value of \( \beta \) indicates the change in the dependent variable in standard deviation units with a change of one standard unit in the independent variable. Standardized coefficients of less than .05 are considered to be very small and should not be interpreted as meaningful (Allison, 1999).

Since multiple independent variables can be entered in the regression analysis the method can be used to examine if one variable mediates the effect of another variable on the dependent variable (Allison, 1999). In this study, the hypothesis that teachers’ beliefs of self-efficacy as measured by the PATH-instrument were related to health via the effect of self-efficacy on perception of professional stressors was examined by regression analysis as a part of the validation process. This is a three-phase analysis. Figure 1 is intended to facilitate the understanding of the nature of this analysis. First an analysis evaluating the effect of teachers’ self-efficacy beliefs on health is performed and coefficients are obtained representing the total effect (\( c \)) of the independent variable (teachers’ self-efficacy beliefs) on the dependent variable (health). Thereafter the relationship (\( a \)) between teachers’ self-efficacy beliefs and perception of professional stressors is examined to ensure that the mediator is dependent on the independent variable. Finally, the analysis of the effect of teachers’ self-efficacy beliefs on health is performed again, but this time the variable perception of professional stressors is included in the analysis and controlled for. The consequence of controlling for perception of professional stressors is that the regression coefficients tell the effect of teachers’ self-efficacy beliefs on health controlling for any possible effect of perception of professional stressors on health. This effect is represented by \( c' \) in the model. If the value of the \( c' \) coefficient is lower than \( c \) this indicates that perception of professional stressors (to some extent) mediates the effect of teachers’ self-efficacy beliefs on health. Put differently, this indicates that teachers’ self-efficacy beliefs affect health through perception of professional stressors (Allison, 1999). A useful measure of the mediation effect size is the proportion of the total effect that is mediated (\( ab/c \)) (MacKinnon, 2008). This ratio constitutes the proportion of the total effect of an independent variable on a dependent variable that is explained by the mediated effect. The significance of the mediated effect can be evaluated by a Sobel test\(^2\) (MacKinnon, 2008).

![Figure 1. Mediation of the effect of teachers’ self-efficacy on health by perception of stressors](http://www.danielsoper.com/statcalc/calc31.aspx)
PASW Statistic Version 18 was used to control data to ensure the assumptions of homoscedasticity and normally distributed residuals for the independent variable, the mediators and the independent variables. In addition it was concluded that tolerance varied from .53 to .80 and multicollinearity thus was not a problem.

The dependent variable health was operationalized using the Early career burnout scale (ECB) measuring burnout by a one-dimensional sequential-developmental model (Gustavsson et al., 2010) and the Scale of work engagement and burnout (SWEBO) measuring state moods associated with work engagement and burnout (Hultell & Gustavsson, 2010). As two measures of health is used in separate analyses results may be replicated and conclusions more reliable. Content, source and range of the two scales, as well as the mean, standard deviation and Cronbach’s α are presented in Table 2 in Appendix.

The mediator variables in the analyses (i.e. perception of stressors within the profession) were operationalized using six scales concerning qualitative and quantitative work demands, role ambiguity, role conflict, social support, and control. These scales were computed from single items from the PATH-survey inquiring about acknowledged stressors in the professional context such as having too much to do and too little influence. Some items had to be reversed before the scales could be computed to ensure the psychometric properties of the scales. Content, source and range of the items included in the six mediation scales, as well as the mean, standard deviation and Cronbach’s α of the computed mediator scales are presented in Table 3 in Appendix.

*Validity evidence based on test content*

Validity evidence for the proposed interpretations of instrument scores based on the instruments content is obtained if proven that items adequately tap and assess self-efficacy beliefs in all of the areas of the profession. Beliefs of self-efficacy are not global but specific for different types of executions (Bandura, 2006). Accordingly, items on self-efficacy measurement need to assess perception of capabilities in diverse situation. Items must be precise in terms of situational demands and circumstances. A measurement where items are formulated in general terms, not explicitly defining the demands that must be handled, will be of little explanatory and predictive value. Finally, the range of the content of the included items must be sufficient to investigate professionals’ beliefs of self-efficacy in all important areas of the profession (Bandura, 2006).

Evaluations of the validity of the interpretation of scores on the PATH-instrument based on test content was made by logical analyses regarding the format of the items, the development of the items, and the way the content of the items relate to the teaching profession in Sweden, as well as the American measurements of teachers’ self-efficacy OSTES.

**Results**

*Validity evidence based on the internal structure*

Validity evidence for the proposed interpretation of the instrument score based on the internal structure was provided by confirmatory factor analysis. The hypothesized hierarchical four-factor structure was confirmed to be an appropriate interpretation of data. Satorra-Bentler Scaled $\chi^2(50)$ was 595.06, $p < .001$, SMRM was .048 and CFI was .98, indicating a good model fit. RMSEA was .087, 90 percent CI (.081-.094), indicating a mediocre model fit. No changes that would make for a better model fit was suggested based
on residual and modification indices, further strengthening the conclusion that the proposed hierarchal four-factor model was a valid interpretation of observed data.

Figure 2 illustrates the confirmed factor structure. Values associated with the paths are standardized factor loadings. All factor loadings were statistically significant \( (p < .001) \) and of large sizes ranging form .71 to .88. As expected, items tse1, tse2 and tse3 loaded on a common factor (Instruct; Efficacy for instructional strategies). Items tse4, tse6, and tse12 formed factor two (Support; Efficacy to give special support to individual students). The factor efficacy for classroom management (Class_m) was confirmed for items tse5, tse7 and tse11. Finally, the latent factor of efficacy for teacher-parent interaction (Parent_m) was confirmed for items tse8, tse9, and tse10. The highest factor loadings were found for item tse2 (.88) and item tse8 (.86).

All four factors loaded highly on the common higher order factor of teachers’ self-efficacy (TSE) confirming the one-dimensional structure. Factor one (Instruct) had a factor loading of .86. Factor two (Support) and factor three (Class_m) had the highest factor loadings (.96 and .99, respectively). Factor four (Parent_m) had the lowest factor loading with a coefficient of .82. In summary, the confirmatory factor analysis provided evidence for the one-dimensional structure of data and supported the validity of the interpretation of scores on the instrument as proposed by the theory of self-efficacy.
19

Figure 2. Hierarchal four-factor model with parameter estimates ($p < .001$). $N = 1489$. tsel-12 represents the 12 items of the instrument presented in Table 1. Instruct = Efficacy for instructional strategies, Support = Efficacy to give special support to individual students, Class_m = Efficacy for classroom management, Parent_m = Efficacy for teacher-parent interaction. TSE = Teachers’ self-efficacy.

Validity evidence based on relations to other variables

Hypotheses about scores on the PATH-instrument intended to measure teachers’ self-efficacy in relation to other variables were supported by the correlation analyses performed in this study. Teachers’ beliefs of self-efficacy as measured by the PATH-instrument were positively related to the four scales assessing mastery, interaction, work satisfaction, and met expectations. On the contrary, teachers’ beliefs of self-efficacy were negatively related to the scale assessing intention to quit. Coefficients $r$ are presented in Table 2. All correlation coefficients were statistically significant at the .001 level (2-tailed). The magnitude of the coefficients indicates small, moderate and strong effects, ranging from (-.25 to .52. The strongest correlation coefficient was found for the scale Mastery, followed by Interaction. This was expected since beliefs about personal agency and control (i.e. mastery) are at the core of self-efficacy. In addition, these beliefs in the context of student interaction are the centre of teachers’ self-efficacy. Test of the significance of the difference
between correlation coefficients showed that the correlation between the PATH-instrument and Mastery was significantly stronger than the correlation between the PATH-instrument and Interaction (z = 2.83, \( p < .005 \) one-tailed) and the scales Work satisfaction, Met expectations, and Intention to quit (z = 6.05, \( p < .001 \); z = 7.44, \( p < .001 \); z = 8.73, \( p < .001 \), one-tailed, respectively). The correlation for Interaction was significantly stronger than the correlation for Work satisfaction, Met expectations and Intention to quit (z = 3.21, \( p < .001 \); z = 4.72, \( p < .001 \); z = 5.9, \( p < .001 \), one-tailed, respectively). In summary, scores on the instrument were related to other variables in the manner that was expected based on the theory of self-efficacy, providing evidence for the validity of the proposed interpretation of scores.

Table 2. Correlations \( r \) between the PATH-instrument and scales representing satisfaction with professional performance (i.e. mastery) (N = 1485), satisfaction with student interaction (N = 1484), work and career choice satisfaction (N = 1484), experienced concordance of expectations and professional experiences (N = 1476), and intention to quit the profession (N = 1483).

<table>
<thead>
<tr>
<th>Scale</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PATH-instrument</td>
</tr>
<tr>
<td>PATH-instrument</td>
<td>-</td>
</tr>
<tr>
<td>Mastery</td>
<td>-</td>
</tr>
<tr>
<td>Interaction</td>
<td>-</td>
</tr>
<tr>
<td>Work satisfaction</td>
<td>-</td>
</tr>
<tr>
<td>Met expectations</td>
<td>-</td>
</tr>
<tr>
<td>Intention to quit</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: All correlations were statistically significant, \( p < .001 \) (2-tailed).

Regression analyses confirmed the hypothesis that beliefs of self efficacy as measured by the PATH-instrument were related to health via the effect of self-efficacy on perception of professional stressors. In the first part of the regression analyses it was concluded that scores on the instrument were significantly related to judgments of health as measured by SWEBO and ECB (total effect indicated by the c in Figure 1, see estimates for the parameters in the c column for all variables in Table 3 and 4). Coefficient \( \beta \) ranged from .35 to .36 (\( p < .001 \)) for the relation between teachers' self-efficacy and SWEBO in the five models, indicating that the change of one standard unit in teachers' self-efficacy would cause a change of .35-.36 standard units in SWEBO. In all five models using ECB as dependent variable coefficient \( \beta \) between teachers' self-efficacy and ECB ranged from .35 to .36 (\( p < .001 \)).

In the second part of the regression analyses it was showed that perception of the professional stressors role ambiguity, role conflict, qualitative- and quantitative demands, social support and control were dependent on the judgment of professional efficacy (arrow a in Figure 1, see estimates for the parameters in the a column for all variables in Table 3 and 4). Coefficient \( \beta \) between teachers’ self-efficacy and the mediator variables in the ten models were statistically significant and varied from .30 to .43 (\( p < .001 \)) for all five.
variables. Perceived qualitative demands and control had the weakest relation to teachers’ self-efficacy in the models using SWEBO as dependent variable ($\beta (1440) = .30, p < .001$ for qualitative demands; $\beta (1424) = .30, p < .001$ for control). Perceived control had the weakest relation to teachers’ self-efficacy in the models using ECB as dependent variable ($\beta (1444) = .30, p < .001$). Perceived role ambiguity had the strongest relation to self-efficacy in models with dependent variable SWEBO, as well as ECB ($\beta (1435) = .42$, and $\beta (1455) = .43, p < .001$, respectively).

In the last step of the analyses it was confirmed that how professional stressors were perceived partially mediated the effect of teachers’ beliefs of self-efficacy on health in all ten models. This is indicated by the lower estimates in the c’ column compared with the c column. In Table 3 and 4 the proportion of the total effect that is mediated is presented in the ab/c column. Sobel tests showed that the mediation effect was statistically significant for all models. The relationship between teachers’ self-efficacy and SWEBO when controlling for the effect of mediator variables is represented by $\beta$ values ranging from .23 to .28 ($p < .001$). The relationship between teachers’ self-efficacy and ECB when controlling for the effect of the mediator variables is represented by $\beta$ values from .23 to .29 ($p < .001$). In summary, regression analyses showed that beliefs of self-efficacy affect health by affecting how professional stressors are perceived as expected by the theory. These results provide evidence for the validity of the interpretation of scores on the instrument.

Table 3. The effect of teachers’ self-efficacy beliefs on health as measured by SWEBO controlling for the variables role ambiguity, role conflict, quantitative demands, qualitative demands, social support, and control, respectively.

<table>
<thead>
<tr>
<th>Variables</th>
<th>a</th>
<th>b</th>
<th>SE</th>
<th>$\beta$</th>
<th>b</th>
<th>SE</th>
<th>$\beta$</th>
<th>c</th>
<th>SE</th>
<th>$\beta$</th>
<th>c’</th>
<th>SE</th>
<th>$\beta$</th>
<th>ab/c</th>
<th>Sobel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role ambiguity</td>
<td>.12</td>
<td>.02</td>
<td>.18</td>
<td>.25</td>
<td>.01</td>
<td>.42</td>
<td>.14</td>
<td>.01</td>
<td>.36</td>
<td>.11</td>
<td>.01</td>
<td>.28</td>
<td>.22</td>
<td>6.44</td>
<td></td>
</tr>
<tr>
<td>Role conflict</td>
<td>.25</td>
<td>.02</td>
<td>.35</td>
<td>.21</td>
<td>.01</td>
<td>.36</td>
<td>.14</td>
<td>.01</td>
<td>.36</td>
<td>.09</td>
<td>.01</td>
<td>.23</td>
<td>.35</td>
<td>10.30</td>
<td></td>
</tr>
<tr>
<td>Quantitative demands</td>
<td>.10</td>
<td>.02</td>
<td>.15</td>
<td>.19</td>
<td>.01</td>
<td>.32</td>
<td>.14</td>
<td>.01</td>
<td>.35</td>
<td>.12</td>
<td>.01</td>
<td>.31</td>
<td>.13</td>
<td>5.21</td>
<td></td>
</tr>
<tr>
<td>Qualitative demands</td>
<td>.20</td>
<td>.02</td>
<td>.30</td>
<td>.22</td>
<td>.01</td>
<td>.30</td>
<td>.14</td>
<td>.01</td>
<td>.35</td>
<td>.10</td>
<td>.01</td>
<td>.24</td>
<td>.32</td>
<td>9.60</td>
<td></td>
</tr>
<tr>
<td>Social support</td>
<td>.18</td>
<td>.02</td>
<td>.30</td>
<td>.22</td>
<td>.02</td>
<td>.33</td>
<td>.15</td>
<td>.01</td>
<td>.36</td>
<td>.11</td>
<td>.01</td>
<td>.27</td>
<td>.27</td>
<td>8.51</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>.14</td>
<td>.02</td>
<td>.25</td>
<td>.21</td>
<td>.02</td>
<td>.30</td>
<td>.14</td>
<td>.01</td>
<td>.35</td>
<td>.11</td>
<td>.01</td>
<td>.28</td>
<td>.21</td>
<td>7.56</td>
<td></td>
</tr>
</tbody>
</table>

Note: All estimates (regression coefficients as well as Sobel estimates) are statistically significant, $p < .001$ (2-tailed). $N = 1411-1476$.  

---

Table 4. The effect of teachers’ self-efficacy beliefs on health as measured by ECB controlling for the variables role ambiguity, role conflict, quantitative demands, qualitative demands, social support, and control, respectively.

<table>
<thead>
<tr>
<th>Variables</th>
<th>(a) b</th>
<th>(SE)</th>
<th>(\beta)</th>
<th>(B) b</th>
<th>(SE)</th>
<th>(\beta)</th>
<th>(c) b</th>
<th>(SE)</th>
<th>(\beta)</th>
<th>(c') b</th>
<th>(SE)</th>
<th>(\beta)</th>
<th>(ab/c)</th>
<th>Sobel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role ambiguity</td>
<td>.12</td>
<td>.02</td>
<td>.18</td>
<td>.30</td>
<td>.02</td>
<td>.43</td>
<td>.17</td>
<td>.01</td>
<td>.36</td>
<td>.13</td>
<td>.01</td>
<td>.28</td>
<td>.22</td>
<td>6.66</td>
</tr>
<tr>
<td>Role conflict</td>
<td>.25</td>
<td>.02</td>
<td>.36</td>
<td>.25</td>
<td>.02</td>
<td>.36</td>
<td>.17</td>
<td>.01</td>
<td>.35</td>
<td>.11</td>
<td>.01</td>
<td>.23</td>
<td>.36</td>
<td>10.23</td>
</tr>
<tr>
<td>Quantitative demands</td>
<td>.10</td>
<td>.02</td>
<td>.15</td>
<td>.27</td>
<td>.02</td>
<td>.39</td>
<td>.17</td>
<td>.01</td>
<td>.35</td>
<td>.14</td>
<td>.01</td>
<td>.29</td>
<td>.17</td>
<td>5.46</td>
</tr>
<tr>
<td>Qualitative demands</td>
<td>.20</td>
<td>.02</td>
<td>.30</td>
<td>.29</td>
<td>.02</td>
<td>.41</td>
<td>.17</td>
<td>.01</td>
<td>.35</td>
<td>.11</td>
<td>.01</td>
<td>.23</td>
<td>.35</td>
<td>10.01</td>
</tr>
<tr>
<td>Social support</td>
<td>.18</td>
<td>.02</td>
<td>.30</td>
<td>.28</td>
<td>.02</td>
<td>.36</td>
<td>.17</td>
<td>.01</td>
<td>.36</td>
<td>.12</td>
<td>.01</td>
<td>.25</td>
<td>.30</td>
<td>9.28</td>
</tr>
<tr>
<td>Control</td>
<td>.14</td>
<td>.02</td>
<td>.25</td>
<td>.25</td>
<td>.02</td>
<td>.30</td>
<td>.17</td>
<td>.01</td>
<td>.35</td>
<td>.13</td>
<td>.01</td>
<td>.27</td>
<td>.22</td>
<td>7.90</td>
</tr>
</tbody>
</table>

Note: All estimates (regression coefficients as well as Sobel estimates) are statistically significant, \(p < .001\) level (2-tailed). \(N = 1411-1476\).

Validity evidence based on test content

Validity evidence for the proposed interpretation of the scores on the instrument based on test content was provided by logical analyses of the format and content of items. The format of items and scales needs to be tailored to properly measure beliefs of self-efficacy (Bandura, 2006). The questions and response scale in the PATH-instrument were of the format presented by Bandura (2006) in his text *Guide for constructing self-efficacy scales*.

The content of the items were chosen to measure the capabilities that Swedish teacher students are to possess in order to receive their diploma. These capabilities are referred to as the requirements for diploma\(^4\) and are supposed to tap the capabilities that are necessary to hold to successfully function as a teacher in Sweden (Per Klingbjer, Swedish Ministry of Education and Research, personal correspondence). These requirements are addressed to the universities, and the purpose thereof is to ensure a certain level of education and preparation for the attending profession (Per Klingbjer, Swedish Ministry of Education and Research, personal correspondence). The specific capabilities included are chosen by a panel of experts based on their collected knowledge of the teaching profession in Sweden on order of the Swedish Ministry of Education and Research (Per Klingbjer, Ministry of Education and Research, personal correspondence). The panel consists of teachers, professors and principals from different Swedish universities, representatives of the Swedish National Agency for Higher Education and the two Swedish teacher unions (Lärarförbundet and Lärarnas Riksförbund) (Per Klingbjer, Swedish Ministry of Education and Research, personal correspondence). Student representatives from Sveriges Förenade Studentkårer are also included (Per Klingbjer, Swedish Ministry of Education and Research, personal correspondence). The essential responsibility of a Swedish teacher is to make possible the cognitive and social development of students at their individual levels (Skollagen 1985:1100). This is the highest order requirement for diploma, followed by specific capabilities necessary to realize this basic obligation (Högskoleförordningen SFS 1993:100). As already defined, teachers’ sense of efficacy refers to teachers beliefs about their capabilities to bring about desired outcomes of student engagement and learning, even

\(^4\) Examensmål
among students who may be unmotivated or difficult (Tschannen-Moran & Woolfolk Hoy, 2001). The concordance between this definition and the content of the Swedish requirements for diploma is obvious. Basing the content of an instrument intended to measure Swedish teachers’ beliefs of self-efficacy on the requirements for diploma issued by the Swedish government thus seems rational.

The content of the instrument was further evaluated in relation to OSTES. As can be seen by comparing Table 1 (in section Data analysis) and Table 4 in Appendix the similarity of the items included in the PATH-instrument of teachers’ sense of efficacy and in the OSTES is striking. The items on the OSTES have been examined by a panel of in-service teachers and approved as tapping abilities significant for successful teaching (Tschannen-Moran & Woolfolk Hoy, 2001). There are only two major differences between the PATH-instrument and the OSTES. First, the items on OSTES are twice as many and of a more specific nature (Tschannen-Moran & Woolfolk Hoy, 2001). Since beliefs of self-efficacy are task specific this may imply that the OSTES gives a more accurate measure of the construct. However, the three factors included in OSTES are represented in the PATH-instrument as well, and given the relation to the Swedish requirements for diploma, the PATH-instrument with its 12 items was expected to adequately assess all essential capabilities of the teaching profession in Sweden. Second, the instrument in the PATH-survey contains three questions regarding the parent-teacher relationship that is not included in the OSTES. Relations to parents of students is a central aspect of the Swedish teaching profession (Högskoleförordningen SFS 1993:100). The questions inquiring about the teacher-parent interaction are thus necessary for the PATH-instrument to provide a valid measure of teachers’ beliefs of self-efficacy in Sweden. Despite these differences, the apparent parallels between items on OSTES and the PATH-instrument were interpreted as providing support for the validity of the content of the PATH-instrument. In conclusion, evidence was found for the validity of the proposed interpretation of the PATH-instrument based on evaluations of various properties of its format as well as content.

Discussion

In the prospective longitudinal PATH-project a cohort of Swedish teachers were followed with annual surveys during a five years period in the transaction from formal education through the early years in the working field (Gustavsson et al., 2007). Aiming at contributing to the framework of professionals’ self-efficacy an instrument supposed to measure teachers’ self-efficacy was included in the annual surveys. The purpose of this study was to investigate the validity of the proposed interpretation of scores on the instrument. Out of recommendation of the American Psychological Association (American Psychological Association, 1999) three sources of validity evidence were investigated. Confirmatory factor analysis confirmed that a hierarchal four-factor model appropriately reflected the internal structure of data ($\chi^2 (50) = 595.06, p < .001$, SMRM = .048, CFI = .98, and RMSEA = .087 (.081-.094)). Expected relationships between scores on the instrument and other variables were confirmed by correlation analyses ($r$ ranging from -.25 to .52, $p < .001$). Regression analyses confirmed expected mediation effects ($\Delta\beta$ ranging from .04 to .13, $p < .001$). In addition, logical analyses regarding the format of items, the source of item content, and relations to the previously validated measurement of teachers’ self-efficacy OSTES, further verified the validity of the interpretation of scores. In summary, this study found support for the validity of the proposed interpretation of scores on the PATH-instrument intended to measure teachers’ self-efficacy.
Next, results and conclusion of this study will be presented in more detail following the same structure as previous sections of this paper. Theoretical and methodical reflections are integrated and presented together with propositions of future research.

**Validity evidence based on internal structure**

The internal structure of the instrument was evaluated by confirmatory factor analysis. Factor analysis is a useful method to assess the internal structure of instruments purposed to measure psychological constructs by investigating the relations of the items assumed to be indicative of the construct. As confirmatory factor analysis is a theory driven method it provides a strict evaluation of the validity of the proposed interpretation of scores based on internal structure (Brown, 2006). Measurements with heterogeneous items assumed to be indicative of a common latent one-dimensional construct are best addressed using hierarchal factor models (Gustafsson & Åberg-Bengtsson, 2010). In this study, a hierarchal four-factor model was hypothesized and four different types of goodness of fit statistics ($\chi^2$, RMSEA, SMRM and CFI) confirmed it to be an appropriate interpretation of data. Standardized factor loadings between the items and the four sub-factors were high, ranging from .71 to .88. Standardized factor loadings of the relation between the four factors and the higher order factor ranged from .82 to .99 supporting the interpretation of teachers' self-efficacy as a one-dimensional construct. It has previously been concluded that teachers’ self-efficacy may be measured as a hierarchal construct by combining self-efficacy beliefs for diverse capabilities that are of central importance to the professional performance (Tschannen-Moran & Woolfolk Hoy, 2001, 2007; Tschannen-Moran et al., 1998). The higher order four-factor structure of data from the PATH-instrument presented in this study is thus in line with theory and previous research.

Additional analysis of the confirmed hierarchal four-factor model showed that factor two (Efficacy for instructional strategies) and factor three (Efficacy to give support to individual students) had the highest factor loadings on the common factor of teachers’ self-efficacy. It could, based on this result, be suggested that these two lower-order factors alone constitutes a sufficient measure of teachers’ self-efficacy. However, it is not possible to claim that the six items forming these two factors (tse4, tse6, tse12, and tse5, tse7, tse11, respectively) adequately assess the full range of the teaching profession. Therefore, such a two-factor instrument would not provide a valid measure of teachers’ self-efficacy. This way of reasoning clearly points to the necessity of including multiple sources of instrument information in a validation study.

Confirmatory factor analysis assumes linear relations among data and therefore need data to be recorded on interval scales (Raykov & Marcoulides, 2006). A limitation in the interpretation of results thus arises because items intended to measure psychological phenomena often are limited to ordinal scales. In this study, this was recognized and handled by using polychoric correlation matrices instead of the traditional correlation matrices, as suggested by Jöreskog (2005).

The major limitations of the result of the confirmatory factor analysis presented in this study concern the possibilities of missing items and alternative models. That data from the PATH-instrument may validly be interpreted as measuring teachers’ beliefs of self-efficacy as shown in this study do not ensure that the full range of the construct is included in the instrument. This problem of construct underrepresentability is a major threat to construct validity (Gustafsson & Åberg-Bengtsson, 2010). There may be indicators of teachers’ self-efficacy that should be included in the instrument but that are not. It is possible that including these items would make for an even better model fit. Future advances in the theory of teachers’ self-efficacy may bring about ideas of additional indicators to include in the instrument. Furthermore, that it could be concluded that the proposed measurement
model was an appropriate interpretation of data does not exclude the possibility that there may be other ways of interpreting data that may be equally appropriate, or better. These limitations add to the necessity of evaluating multiple aspects of instruments in validation processes.

Validity evidence based on relations to other variables

Validity of score interpretations was examined by correlation analyses relating the scores of the instrument to scales computed by items from the PATH-survey. All correlation coefficients were statistically significant ($p < .001$) and of the expected directions, providing support for validity of the proposed interpretation of scores on the PATH-instrument. Coefficients $r$ varied from -.25 for the scale Intention to quit, to .52 for the scale Mastery. As expected based on the theory of self-efficacy and teachers’ self-efficacy (e.g. Bandura, 1977; Tschannen-Moran & Woolfolk Hoy, 2001), the scales assessing mastery and satisfaction with student interaction had coefficients of particularly large magnitude. The scale Mastery is conceptually closer to self-efficacy (i.e. the ability to exercise control over events that affect one’s life) than the other scales, as it inquire about satisfaction with quality and quantity of professional performance, as well as satisfaction with problem solving ability. The scale Interaction is likewise related to teachers’ self-efficacy as it includes items inquiring about how rewarding, exciting, likable and annoying the teachers find their student interaction. Teachers’ self-efficacy is defined as “teachers beliefs about their capabilities to bring about desired outcomes of student engagement and learning, even among students who may be unmotivated or difficult” (Tschannen-Moran & Woolfolk Hoy, 2001). The construct has previously been associated with teachers’ willingness to try new instructional strategies, and persistence in helping students who have trouble learning (Tschannen-Moran et al., 1998). Self-efficacy of teachers has also been related to teachers’ enthusiasm about teaching, as well as teachers’ commitment to the profession (Tschannen-Moran et al., 1998). Teachers with a lower sense of efficacy have been shown to be more likely to leave the profession (Tschannen-Moran & Woolfolk Hoy, 2001). Teachers that feel efficacious in their professional role express higher satisfaction about their work, and have a more positive attitude about staying in the field of teaching than teachers that doubt their professional capabilities (Tschannen-Moran et al., 1998). As this study may confirm these relations for the scores on the PATH-instrument, evidence is provided for the validity of the instrument as a measure of teachers’ self-efficacy.

Correlation analyses, as confirmatory factor analyses, assume linear relationships among data and thus need data to be at least on interval level. Ordinal variables of the kind used in the PATH-survey are however often analyzed using this method. It is assumed that respondents have interpreted the ordinal response scale as an interval scale where an increase or decrease of one unit means the same no matter the position on the scale. In many cases this is rational and has no large consequences on the results (Allison, 1999). One should however be aware of this assumption and way of conduct.

The major limitation of the results of correlation analyses presented in this study concern uninvestigated relations. Although five scales with diverse content were significantly related to scores on the teacher’ self-efficacy instrument, there still remain important relationships that could not be investigated in this study due to lack of data. Based on the theory of Bandura (1977) a measurement of self-efficacy should relate to (a) the likelihood that teachers choose to engage in versus avoid tasks within the teaching profession, (b) the effort they will put into the execution of the task, (c) how long they will persist in their efforts when faced with difficulties, and finally (d) their success in performance. For instance, it would have been valuable to relate scores on the instrument to objective indices of teachers’ performance outcomes. Similarly, it would have been
interesting to examine if scores on the instrument relate to the performance of the teachers’
students. If scores on the instrument correctly reflect teachers’ beliefs of self-efficacy they
are expected to be positively related to students learning (Bandura, 1997; Skolverket, 2006)
and teachers’ willingness to try new strategies in order to better meet the needs of students
and facilitate learning (Tschannen-Moran et al., 1998). Future examinations of that kind
may provide additional confirming or disconfirming evidence for the validity of the
proposed interpretation of scores on the instrument.

Hypotheses about relations of scores on the PATH-instrument and other variables
were further evaluated with regression analysis. The analyses confirmed the hypothesis that
beliefs of self-efficacy as measured by the PATH-instrument were related to health via the
effect of self-efficacy on perception of professional stressors. The reliability of this result is
supported since it was replicated for models using SWEBO as well as ECB as dependent
variable. In the field of occupational health psychology it has been proposed that
professionals’ self-efficacy is a key concept in the relation between taxing professional
situations and ill health such as burnout (Cherniss, 1980a, 1993). A firm belief in personal
capability to control stressors is expected to counteract ill health by affecting the extent to
which stressors are perceived as threats that evoke arousal. That these previous findings and
theoretical assumptions could be reflected in this study support the interpretation of the
PATH-instrument as providing a valid measure of teachers’ self-efficacy.

The investigated mediators (perception of six types of acknowledge stressors within
the profession; (Cherniss, 1980b; Karasek & Theorell, 1990)) accounted for 17 to 35
percent of the total effect of teachers’ self-efficacy on health. In other words, the mediators
examined in this study did only partially explain the relation between teachers’ beliefs of
self-efficacy and health. However, this does not inflict on the conclusion about the validity
of the proposed interpretation of scores on the instrument. According to the theory, and as
presented in the introduction of this paper, beliefs of self-efficacy affect the human agent
via cognitive, motivational, and physiological processes, as well as via selection of
environment (Bandura, 1989). The mediation effect of perception of stressors here
investigated thus only account for (part of) the cognitive processes by which the effect of
self-efficacy is expressed. Therefore, data included in these analyses were not expected to
account for the total effect of self-efficacy on health. Data concerning affective processes,
motivational processes, and selection of environment would have been interesting to
include in the analyses. Possibly this could broaden the understanding of the relation
between scores on the PATH-instrument and health, and contribute to the investigation of
the validity of the instrument. Unfortunately, those data were not available. In addition, it
would have been interesting to investigate if teachers’ beliefs of self-efficacy as measured
by the PATH-instrument mediate the effect of objective professional stressors on health as
expected by the theory of self-efficacy. However, since all data in the PATH-survey are
self-report, no objective measures were available. When studying relations of variables and
mediation effects with multiple regression analyses a discussion of the included variables,
such as this, is always valid. Typically the researcher aim at describing the total effect of an
independent variable on an independent variable by mediation variables (Allison, 1999). If
the total effect is only partially explained by the mediators, some mediator variable (or
variables) is missing, and the explanatory value of the mediation model is limited. The
purpose of the mediation analyses in this study was however not to explain the total relation
between teachers’ self-efficacy and health. Instead, it was specifically to investigate if
perception of professional stressors affected by beliefs of self-efficacy mediate the relation
of scores on the PATH-instrument and assessments of health, as proposed in the field of
occupational health psychology (Cherniss, 1980a, 1993). In conclusion, even though the
regression analyses only showed partial mediation effects the results are in line with the
theory of self-efficacy and previous research. The mediation analyses thus confirmed the hypotheses, and provided additional support for the interpretation of scores on the PATH-instrument based on investigation of relations to other variables.

As is the case with correlation analyses, multiple regressions are often used to evaluate relations among ordinal level variables even though the method assumes data to be recorded on interval scales (Allison, 1999). In this study, assumptions about homoscedasticity, normally distributed residuals and multicollinearity were investigated and it was concluded that they were not violated, assuring the value of the results.

Since all statistical analyses conducted in this validation study are based on evaluations of relations of items, a general methodological discussion about the fact that all data are self-report is warranted. Answers on the items are given by the same person at the same time. This may result in a same-source bias and inflated correlation coefficients if data on variables share variance due to some other factor than the variable in question. On the other hand, errors are included in the computation of coefficients. This may cause estimates to be lower than the real-life relations among variables (Schmidt & Embretson, 2003), and faulty conclusions as a result of self-report data thus should not be a problem in this study.

Also important to note is that when analyzing data by correlations, as is done in this study regarding relations of the PATH-instrument and other variables (e.g. mastery, satisfaction with student interaction, perception of professional demands, health), no conclusions about causality may be drawn. It is not possible to decide, based on analyses of co variation, which one of two variables that causally effect the other, or if there is a third variables effecting both variables.

Validity evidence based on test content

By logical analyses the content of the items included in the PATH-instrument was judged as appropriate to provide a valid measure of teachers’ self-efficacy based on the format of questions and the response scale, the relations to the requirements of the teaching profession in Sweden, and the measurement OSTES, validated for assessing teachers self-efficacy in America. The primary limitation of this conclusion originates in the content of the Swedish requirement of diploma. In this study it was assumed that these requirements issued by the Swedish Ministry of Education and Research adequately tap the full range of capabilities that is necessary to hold to successfully function as a teacher in Sweden. This assumption seems rational since the requirements are developed by a group of experts on the Swedish teaching profession. However, there may be capabilities that are necessary to hold that are not included in the document of requirements, and thus was not included in the PATH-instrument. For instance, there may be capabilities that teachers acquire in the working field but are not expected to hold as they graduate from their formal education. These capabilities would not be included in the requirements for diploma. Should this be the case the instrument cannot be regarded as a valid measure of in-service teachers’ beliefs of self-efficacy. Future studies may address this limitation by discussing the content of the instrument with a panel of experts out of the perspective of the professional field and not the educational field.

In summary, in this study, based on analyses of internal structure, relations to other variables, and instrument content, evidence was found for the validity of the proposed interpretation of scores on the PATH-instrument intended to measure teachers’ beliefs of self-efficacy in Sweden.

A last recommended source of evidence for the validity of the interpretation of scores based on response processes (American Psychological Association, 1999) could not be evaluated in this study. Future analyses of information concerning response processes may
provide additional reassurance of the validity of the proposed interpretation of scores. For the instrument to be a valid measure of teachers’ self-efficacy it is of greatest importance to conclude that the instrument really measures respondents’ perceived capabilities to exercise control within their profession (Bandura, 2006). The assessment must reflect what teachers believe they can do in a given situation, not what they believe they will do (Bandura, 2006). Interviewing a number of teachers about their thought processes as they answer the questions may be a useful approach to investigate this final source of validity evidence.

No predictive analyses were conducted in this study, which may be its major limitation. Predictive validity is argued to be the most important psychometric information to present when reporting about measurement instruments (Rick, Briner, Daniels, Perryman, & Guppy, 2001). The longitudinal PATH-study contains data that makes it possible to conduct such analyses in the future. Predictive validity evidence for the proposed interpretation of instrument score would be obtained if proven that teachers low in self-efficacy at the beginning of their career at a later point in time are less content with their professional performance, experience higher level of professionally related stress and ill-health, and choose to leave the profession to a greater extent than teacher high in self-efficacy. This type of longitudinal analyses does however awake an additional number of considerations relevant to the validity of the proposed interpretation of scores. This study has provided evidence for the validity of the proposed interpretation when respondents are one year out of their formal education. Further analyses are necessary to claim the validity of the proposed interpretation of scores when teachers are still in education, as well as later in their professional career. In addition, it needs to be concluded that the proposed interpretation of the scores on the instrument is equally valid for teachers working with students of all ages. The content and demands of the profession is not the same for teachers working with small children as for teachers working with older children. It is thus possible that the instrument provides a better measure for teacher of one group or the other. To ensure that the instrument provides a valid score of teachers’ self-efficacy for both professional groups is important. The same can be said for all kinds of groups (sex, age, et cetera). This reasoning relates to the generalizability of the results. The original sample of teacher students included in the PATH-study was representative for the population cohort of teacher students in Sweden at that time (Theorin, 2006). To investigate the generalizability of the results of this validation study, an attrition analysis was performed using logistic regression with attrition (versus responding) as dependent variable; and sex, age, and age of students (younger versus older), during the first wave of measurement as independent variables. The results showed that males (OR=0.75; \( p < .001 \)) and younger participants (OR=0.98; \( p < .001 \)) were more likely to not participate in the second follow-up of the PATH-study. However, the amount of explained variance in attrition was only 1.6%, indicating that it was not very likely that this had any considerable effect on the results of the study. In conclusion, the sample used in this study to evaluate the validity of interpretations of scores on the instrument intended to measure teachers self-efficacy was representative for the Swedish population of newly educated teacher students. Future analyses of validity in relations to other groups is however necessary to conclude that the results may be generalized to the full professional teacher population.

Summary

In this study three sources of evidence for the validity of interpretations of scores on an instrument intended to measure teachers’ beliefs of self-efficacy in the prospective longitudinal PATH-study was examined. The instrument includes 12 items and the internal consistency as indicated by Cronbach’s \( \alpha \) is .92. A confirmatory factor analysis confirmed
that scores on the instrument rightfully can be interpreted as measuring teachers’ self-efficacy as a hierarchical four-factor construct. Correlation analyses and regression analyses provided support for the proposed interpretation of scores based on relations to other variables. Logical analyses of the instrument format, content of the items, and relation to the Ohio State teacher efficacy scale provided supporting evidence based on instrument content. In summary, it was concluded that scores on the instrument included in the longitudinal PATH-study constitute a valid measure of Swedish teachers’ beliefs of self-efficacy.

The primary limitations of this study concern analyses that could not be conducted because of lack of data (such as analyses relating scores on the instrument to objective measures of teachers’ performance outcomes, and analyses of response processes) as well as the lack of predictive analyses. Future analyses of these kinds may provide additional support for the validity of the interpretations of scores on the instrument.

Within the field of occupational health psychology it has been proposed that the development of a strong sense of self-efficacy is essential for new professionals to successfully handle the stressors of their professions. Knowledge about the nature of the development of these beliefs has however been lacking. Now that it has been concluded that the instrument included in the PATH-surveys provides a valid measure of newly educated teachers’ beliefs of self-efficacy in Sweden it is possible to investigate the development of self-efficacy beliefs in the transition from formal education to the professional field within the PATH-study. Increased knowledge of professionals’ self-efficacy may bring about ideas about interventions targeting the development of these important beliefs in the battle against job dissatisfaction, turnover and ill health.

References


Appendix

Table 1. Item content, source and range of scales used in correlation analyses.

<table>
<thead>
<tr>
<th>Scale name</th>
<th>Items</th>
<th>Source</th>
<th>Range</th>
<th>M</th>
<th>SD</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mastery</td>
<td>I am satisfied with the quality of the work I do</td>
<td>(Dallner et al., 2000)</td>
<td>1-5 (1= very rarely or never; 5=very often or always)</td>
<td>4.05</td>
<td>0.61</td>
<td>.75</td>
</tr>
<tr>
<td></td>
<td>I am satisfied with the amount of work I do</td>
<td>(Dallner et al., 2000)</td>
<td>1-5 (1= very rarely or never; 5=very often or always)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I am satisfied with my ability to solve problems at work</td>
<td>(Dallner et al., 2000)</td>
<td>1-5 (1= very rarely or never; 5=very often or always)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>I think it is very rewarding to work with my students.</td>
<td>(Lait &amp; Wallace, 2002)</td>
<td>1-5 (1= very rarely or never; 5=very often or always)</td>
<td>4.35</td>
<td>0.54</td>
<td>.73</td>
</tr>
<tr>
<td></td>
<td>I am very excited to be working with my students.</td>
<td>(Lait &amp; Wallace, 2002)</td>
<td>1-5 (1= very rarely or never; 5=very often or always)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I get very annoyed at my students.</td>
<td>(Lait &amp; Wallace, 2002)</td>
<td>1-5 (1= very rarely or never; 5=very often or always)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I dislike working with my students.</td>
<td>(Lait &amp; Wallace, 2002)</td>
<td>1-5 (1= very rarely or never; 5=very often or always)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work satisfaction</td>
<td>I am satisfied with my career choice</td>
<td>(Blau, Paul, &amp; St John, 1993)</td>
<td>1-5 (1= strongly agree; 5=strongly disagree)*</td>
<td>4.00</td>
<td>0.87</td>
<td>.83</td>
</tr>
<tr>
<td></td>
<td>If I had to choose today, I would not choose to work as a teacher</td>
<td>(Blau et al., 1993)</td>
<td>1-5 (1= strongly agree; 5=strongly disagree)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Even if I was not dependent on my salary I would still continue to work as a teacher</td>
<td>(Blau et al., 1993)</td>
<td>1-5 (1= strongly agree; 5=strongly disagree)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I wish I had chosen another career</td>
<td>(Blau et al., 1993)</td>
<td>1-5 (1= strongly agree; 5=strongly disagree)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met expectations</td>
<td>My experience of this work has been more positive than I originally expected</td>
<td>(Lait &amp; Wallace, 2002)</td>
<td>1-5 (1= strongly disagree; 5=strongly agree)</td>
<td>3.68</td>
<td>0.88</td>
<td>.68</td>
</tr>
<tr>
<td></td>
<td>By and large, this job is not what I thought it would be</td>
<td>(Lait &amp; Wallace, 2002)</td>
<td>1-5 (1= strongly disagree; 5=strongly agree)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>My job has not lived up to the expectations I had when I first started working</td>
<td>(Lait &amp; Wallace, 2002)</td>
<td>1-5 (1= strongly disagree; 5=strongly agree)*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Intention to quit
I often think of changing profession (Hellgren, Sjöberg, & Sverke, 1997) 1-5 (1=strongly disagree; 5=strongly agree) 1.41 0.76 .82
I am actively looking for work outside teaching (Hellgren et al., 1997) 1-5 (1=strongly disagree; 5=strongly agree)
I would as soon as possible like to leave the teaching profession (Hellgren et al., 1997) 1-5 (1=strongly disagree; 5=strongly agree)
* reversed items

Table 2. Items, source, range, $M$, $SD$ and $\alpha$ of the scales used as dependent variables in the regression analyses.

<table>
<thead>
<tr>
<th>Scale name</th>
<th>Items</th>
<th>Source</th>
<th>Range</th>
<th>$M$</th>
<th>$SD$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWEBO</td>
<td>In the last two weeks at work I have been feeling: Decrepit</td>
<td>(Hultell &amp; Gustavsson, 2010)</td>
<td>1-4 (1=all the time; 4=not at all)</td>
<td>3.39</td>
<td>0.53</td>
<td>.91</td>
</tr>
<tr>
<td></td>
<td>In the last two weeks at work I have been feeling: Undecided</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>In the last two weeks at work I have been feeling: Exhausted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>During the past two weeks I have towards my work been feeling: Indifference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>During the past two weeks I have towards my work been feeling: futility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>During the past two weeks I have towards my work been feeling: Resignation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>During the last two weeks when I have worked I have been: out of focus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>During the last two weeks when I have worked I have felt: Restless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>During the last two weeks when I have worked I have felt: distractibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECB</td>
<td>There are days when I feel tired even before I go to work.</td>
<td>(Gustavsson et al., 2010)</td>
<td>1-4 (1=entirely accurate; not accurate at all)</td>
<td>3.04</td>
<td>0.62</td>
<td>.87</td>
</tr>
<tr>
<td></td>
<td>It happens more and more often that I talk about my work in a disparaging way.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I need more time to relax now than before to recover from work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lately, I have done my work more and more mechanically, without using the brain.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>At work, I feel often emotionally leached.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Over time one loses a deeper interest in the work.
I feel more and more indifferent to work.

Table 3. Items, source, range, $M$, $SD$ and $\alpha$ of the scales used as mediators in the regression analyses.

<table>
<thead>
<tr>
<th>Scale name</th>
<th>Items</th>
<th>Source</th>
<th>Range</th>
<th>$M$</th>
<th>$SD$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role ambiguity</td>
<td>I'm getting conflicting information from two or more people in my work. I am forced to do things in my work that should be done in a different way. I carry out my work in a manner that is accepted by one colleague, but not others. It often happens that I get instructions or directives that contradict each other.</td>
<td>(Sverke &amp; Hellgren, 2002)</td>
<td>1-5 (1=very often or always; 5=very rarely or never)</td>
<td>3.89</td>
<td>0.89</td>
<td>.85</td>
</tr>
<tr>
<td>Role conflict</td>
<td>It is clearly stated what is expected of me in my work.* I have a clear understanding of the tasks of my post.* I think my work directives are diffuse and unclear.</td>
<td>(Sverke &amp; Hellgren, 2002)</td>
<td>1-5 (1=very often or always; 5=very rarely or never)</td>
<td>3.58</td>
<td>0.91</td>
<td>.79</td>
</tr>
<tr>
<td>Quantitative demands</td>
<td>I have enough time to complete my tasks.* It happens that I have to work under high time pressure. I have too much to do in my work. Unreasonable demands are put on me in my work.</td>
<td>(Sverke &amp; Hellgren, 2002)</td>
<td>1-5 (1=very often or always; 5=very rarely or never)</td>
<td>3.19</td>
<td>0.91</td>
<td>.87</td>
</tr>
<tr>
<td>Qualitative demands</td>
<td>I have too great a responsibility in my work. I have tasks that I think are too difficult to manage. My work contains elements that demands too much of my capacity.</td>
<td>(Sverke &amp; Hellgren, 2002)</td>
<td>1-5 (1=very often or always; 5=very rarely or never)</td>
<td>3.83</td>
<td>0.88</td>
<td>.78</td>
</tr>
<tr>
<td>Social support</td>
<td>If you need, do you get support and help with your work from your colleagues?* If you need, are your co-workers then willing to listen to problems related to your work?* Can you get appreciating feedback for your performance at work from your immediate supervisor?* If you need, do you get support and help with your work from your immediate supervisor?*</td>
<td>(Dallner et al., 2000)</td>
<td>1-5 (1=very often or always; 5=very rarely or never)</td>
<td>3.73</td>
<td>0.80</td>
<td>.82</td>
</tr>
</tbody>
</table>
If you need, is your immediate supervisor then willing to listen to problems related to your work?*

Control

I have sufficient influence in my work.*
I can decide how I want to do my work.*
There is room for me to take own initiatives in my work.*

(Sverke & Hellgren, 2002)

1-5 (1=very often or always; 5=very rarely or never)

3.99 0.75 .81

*reversed items

Table 4. OSTES factor structure.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1: Efficacy for instructional strategies</td>
<td>1. To what extent can you use a variety of assessment strategies?</td>
</tr>
<tr>
<td></td>
<td>2. To what extent can you provide an alternative explanation or example when students are confused?</td>
</tr>
<tr>
<td></td>
<td>3. To what extent can you craft good questions for your students?</td>
</tr>
<tr>
<td></td>
<td>4. How well can you implement alternative strategies in your classroom?</td>
</tr>
<tr>
<td></td>
<td>5. How well can you respond to difficult questions from your students?</td>
</tr>
<tr>
<td></td>
<td>6. How much can you do to adjust your lessons to the proper level for individual students?</td>
</tr>
<tr>
<td></td>
<td>7. To what extent can you gauge student comprehension of what you have taught?</td>
</tr>
<tr>
<td></td>
<td>8. How well can you provide appropriate challenges for very capable students?</td>
</tr>
<tr>
<td>Factor 2: Efficacy for classroom management</td>
<td>9. How much can you do to control disruptive behavior in the classroom?</td>
</tr>
<tr>
<td></td>
<td>10. How much can you do to get children to follow classroom rules?</td>
</tr>
<tr>
<td></td>
<td>11. How much can you do to calm a student who is disruptive or noisy?</td>
</tr>
<tr>
<td></td>
<td>12. How well can you establish a classroom management system with each group of students?</td>
</tr>
<tr>
<td></td>
<td>13. How well can you keep a few problem students from ruining an entire lesson?</td>
</tr>
<tr>
<td></td>
<td>14. How well can you respond to defiant students?</td>
</tr>
<tr>
<td></td>
<td>15. To what extent can you make your expectation clear about student behavior?</td>
</tr>
<tr>
<td></td>
<td>16. How well can you establish routines to keep activities running smoothly?</td>
</tr>
<tr>
<td>Factor 3: Efficacy for student engagement</td>
<td>17. How much can you do to get students to believe they can do well in schoolwork?</td>
</tr>
<tr>
<td></td>
<td>18. How much can you do to help your students value learning?</td>
</tr>
<tr>
<td></td>
<td>19. How much can you do to motivate students who show low interest in schoolwork?</td>
</tr>
<tr>
<td></td>
<td>20. How much can you assist families in helping their children do well in school?</td>
</tr>
</tbody>
</table>