Self-Regulated Learning as a Competence
Implications of Theoretical Models for Assessment Methods

Joachim Wirth und Detlev Leutner
Duisburg-Essen University, Germany

Abstract. Lively research on self-regulated learning has produced a great number of models of self-regulated learning competence and it is still a challenge to integrate them within a single coherent framework. However, such a framework is necessary for, among other reasons, the development of valid assessment methods. We argue that one common characteristic of all models is that they consider the competence to make solid comparisons as a key competence of self-regulated learning. However, the kind of comparisons and the kind of standards used for these comparisons differ between models. The same is true for assessment methods. Valid assessment methods also have implemented comparisons and they also differ concerning the kind of comparison and the kind of standards used for assessment. In order to categorize both, models as well as assessment methods, we propose to distinguish between component models and process models of self-regulated learning. Component models imply the use of offline standards for assessment whereas process models imply the use of online standards. Both offline and online standards can be either quantitative or qualitative. We show that using qualitative standards leads to a higher validity of the assessment than using quantitative standards. This advantage of qualitative standards can be shown for both offline standards as well as online standards.

Keywords: self-regulated learning, competence model, assessment, measurement.

Introduction

The competence of self-regulated learning is regarded as a necessary prerequisite for lifelong learning (e.g., Commission of the European Community, 2000; Spiel & Schober, 2002) and is mentioned in almost all German school curricula (Artelt, 2006; Artelt, Baumert, & Julius-McElvany, 2003). Unfortunately, however, there is no commonly accepted definition of self-regulated learning competence. One reason for this lack of consensus is that the lively research on self-regulated learning over the last 30 years has produced a great number of models that differ to a greater or lesser extent (e.g., Boekaerts, 1999; Pressley, Borkowski, & Schneider, 1989; Winne & Hadwin, 1998; Zimmerman, 2000). These models have served as the theoretical foundation for the development of assessment methods that, as a logical result, also differ to a greater or lesser extent. Thus, it remains a challenge for research in the field of self-regulated learning to integrate the different models within a single coherent framework or theory. Such a framework or theory is necessary for the development of valid methods assessing self-regulated learning competence.

One attempt to catalog and synthesize the different models has been made by Winne and Perry (2000; see also Thillmann, 2007). According to these authors, models of self-regulated learning can be categorized as component models or as process models. Component models describe self-regulated learning in terms of different learner competencies that foster self-regulated learning and that are considered as relatively enduring attributes of the person. One of the most prominent component models of self-regulated learning is Boekaerts' (1997) six-component model. Boekaerts describes components of self-regulated learning competence in terms of different types of prior knowledge. She distinguishes between prior knowledge about cognitive aspects of self-regulation and prior knowledge about motivational aspects of self-regulation. Boekaerts emphasizes that both types of prior knowledge are equally important for the competent self-regulation of learning. Both types can be further differentiated on three levels, (a) the domain-specific level, (b) the strategic level, and (c) the goal level. Without going into the details of this model, the six-component model clearly aims to describe self-regulated learning competence in terms of the prerequisites a learner needs for competent self-regulation. This perspective on prerequisites is characteristic of component models and has important implications for assessment. For example, assessment methods based on component models can be administered independently of any learning process. If the competence of self-regulated learning is seen as a set...
of enduring learner attributes, these attributes should be assessable at any time.

Component models describe the competence of self-regulated learning without specifying in which phase of the learning process the different components of competence are needed. In contrast, process models aim to describe the (ideal) process of self-regulated learning in terms of the properties of a series of phases or events. For example, Zimmerman (1998) postulates three cyclical phases of self-regulation, (a) forethought, (b) performance or volitional control, and (c) self-reflection, each of which can be characterized by the specific demands they make on the learner. Thus, process models describe typical requirements that have to be met in different phases of the cyclic learning process, but they do not specify the competencies needed to meet these requirements.

The focus on phases and phase-specific requirements also has important implications for the development of assessment methods. It suggests (a) that different assessment methods are needed for the requirements typical of the different phases, and (b) that the assessment of self-regulated learning must take place during the phases of a learning process.

The different models of self-regulated learning make it difficult to find a commonly accepted definition of what can be regarded as a competence of self-regulated learning. In order to formulate at least a working definition of self-regulated learning competence, we combined the core features of component models and process models of self-regulated learning. We suggest that the competence of self-regulated learning can be defined as a learner’s competence to autonomously plan, execute, and evaluate learning processes, which involves continuous decisions on cognitive, motivational, and behavioral aspects of the cyclic process of learning (Boekaerts, 1999; Pressley et al., 1989; Winne & Hadwin, 1998; Zimmerman, 2000). These decisions are related to a variety of metacognitive demands that arise during the learning process. Thus, the competence of self-regulated learning can be regarded as a structured composition of different competence components, each of which aims to meet one of these metacognitive demands.

The different models of self-regulated learning (i.e., component models and process models) differ widely in terms of the metacognitive demands they emphasize, but at least three classes of metacognitive demands are incorporated in all models (Schreiber, 1998; Zimmerman, 2000). These three classes can be labeled goal setting, planning, and monitoring.

Goal setting requires the learner’s competence to decide what needs to be learned. In order to make this decision, the learner must be able to analyze the conditions and constraints of the learning task (task conditions) and to activate and evaluate his or her prior knowledge about the domain and about appropriate learning strategies (cognitive conditions; Winne & Hadwin, 1998). By comparing the internal cognitive conditions with the external task conditions, the learner identifies differences between his or her actual and desired knowledge. Thus, learning goals can be seen as the result of a learner’s decision to reduce one or more of these differences with respect to a certain criterion.

Planning the learning process requires the learner’s competence to decide how to reach the learning goal. It requires the learner to choose a learning strategy that is appropriate to transform his or her current knowledge into the desired knowledge. In order to be able to choose a learning strategy, the learner needs to activate his or her knowledge about learning strategies. For each activated learning strategy, the learner then has to analyze the conditions under which it is applicable (conditional knowledge; Paris, Lipson, & Wixson, 1983) and to compare these conditions with the characteristics or conditions of the learning task. A learning plan can, thus, be seen as the result of a learner’s decision for one or more of the activated learning strategies, the conditions of which are met by the task conditions.

Monitoring is considered the key component of self-regulated learning competence (Butler & Winne, 1995; Winne, 1996). It comprises two subcomponents (Schreiber, 1998; Zimmerman, 2000). First, it requires the learner’s competence to continuously keep track of what he or she is doing during the learning process and of the outcome of the learning activities (Bandura, 1986). This observation subcomponent should be as objective as possible and, therefore, independent from the learning goals. Secondly, monitoring requires the learner’s competence to evaluate whether the learning activities executed correspond with the learning strategies planned, and whether the learning outcome observed corresponds with the learning goals. Thus, there are two results of the monitoring process: The first concerns whether there is a discrepancy between the observed and the planned learning activities; the second concerns whether there is a discrepancy between the actual and the desired knowledge.

Self-regulated learning is regarded as a cyclic process (Winne & Hadwin, 1998; Zimmerman, 2000). That means that if a learner detects any discrepancy during monitoring, he or she again has to decide on the learning goals and the learning plan. To refine his or her learning goals and/or learning plan, the learner again has to make the comparisons and decisions described above, bearing in mind that both the task conditions and the cognitive conditions will probably have been changed by the learning activities engaged in thus far.

The cyclic feature is typical of all kinds of regulation processes. In psychological research, for example, it has been described in terms of the well-known test-operate-test-exit loop (TOTE; Miller, Galanter, & Pribram, 1965). The test components of this loop are characterized by the comparison of an actual state with a desired state. Likewise, comparisons must be drawn to meet the metacognitive demands of self-regulated learning. Learning goals result from the comparison of what a learner currently knows and what he or she desires to know. Learning plans result
from a comparison of the learning strategies currently available to a learner and the learning strategies that will lead to the desired learning goal under the actual task conditions. Monitoring detects discrepancies by comparing the observed learning activities and the observed learning outcome with the desired learning activities defined by the learning plan and the desired learning outcome defined by the learning goals. Thus, one of the key competencies needed to meet all the metacognitive demands of self-regulated learning is the competence to make solid comparisons between an actual state and a desired state or, in other words, to compare empirical outputs with theoretical standards. Methods assessing self-regulated learning competence must take this key component of self-regulation into account. In other words, a valid assessment of self-regulated learning competence must implement comparisons of actual and desired states or comparisons of empirical output values and theoretical standard values.

Indeed, methods assessing self-regulated learning are based on such comparisons. They compare a learner’s observed behavior with a theoretical standard. Depending on the theoretical model on which the assessment methods are based, however, the theoretical standards with which the learner’s behavior is compared differ. The question of how to define and assess self-regulated learning competence is, thus, the question of how to model and operationalize the comparison of the learner’s observed behavior with a standard behavior.

In the following, we describe the different approaches to assessing the competence of self-regulated learning, or at least some components of it, and we give a typical example of each approach. We discuss the approaches in terms of how they implement comparisons of empirical values (observed behavior) with theoretical values (theoretical standards), and how the kind of implementation affects the validity of the assessment method.

**Assessing Self-Regulated Learning**

We categorize the different assessment approaches according to the kind of standards of comparison they use. On the one hand, we distinguish between offline standards and online standards (Desoete, Roeyers, & De Clercq, 2003). **Offline standards** are defined independently of an actual learning process. They are absolute values that are the same for all learners tested. They are usually applied to assess whether learners possess the different competencies regarded as prerequisites of competent self-regulation of learning (e.g., knowledge about learning strategies). Offline standards are theoretically based on component models of self-regulated learning (Thillmann, 2007). Component models (e.g., Boekaerts, 1997, 1999; Pintrich, 1999, 2000) describe a number of different competencies, all fostering self-regulated learning. However, component models do not refer directly to an actual process of self-regulated learning. In contrast, **online standards** are continuously defined and redefined during a learning process, in response to the continuously changing conditions (Wirth & Leutner, 2006). These standards change continuously during the learning process, and they differ interindividually. Online standards are usually applied to assess the process of regulation itself. That is, they do not assess whether learners possess the competencies required for self-regulated learning, but whether they are able to apply their self-regulatory competencies when working on a specific learning task.

Online standards are theoretically based on process models of self-regulated learning that describe an ideal cycle of self-regulatory phases during learning (e.g., Winne & Hadwin, 1998; Wirth, 2004, 2005; Zimmerman, 2000).

On the other hand, we distinguish between quantitative standards and qualitative standards. The underlying assumption of **quantitative standards** is that the more students execute certain (cognitive or metacognitive) activities, the better they are able to regulate their learning. They are used for “the more . . . the better” comparisons. **Qualitative standards** are based on definitions of activities that are assumed to be of high regulatory quality. They are used for “the better the fit . . . the better” comparisons. Definitions of high-quality activities are derived from theoretical models and/or from expert ratings. The underlying assumption of these standards is that the better the fit between the learners’ activities and what are specified to be high-quality activities, the better they are able to regulate their learning.

**Offline Standards**

**Quantitative Offline Standards**

Most of the test instruments commonly used to assess self-regulated learning are questionnaires and inventories using quantitative offline standards (Biggs, 1978; Entwistle, 1988; Pintrich, Smith, Garcia, & McKeachie, 1991; Weinstein, 1987; Wild, Schiefele, & Winteler, 1992). These instruments aim to assess the use of learning strategies by defining different classes of learning strategies (see Baumert & Köller, 1996) and administering a number of questionnaire items per class. Each item asks respondents to rate whether they tend to use a specific learning strategy. For example, learners have to rate their agreement with the item “When reading I try to relate the material to what I already know” (Pintrich et al., 1991) on a 7-point Likert scale ranging from *not at all true of me* (1) to *very true of me* (7).

The standard of this kind of assessment method is defined as the theoretical maximum rating across all items. That is, respondents receive the highest score if they rate all of the positively formulated items as 7 and all of the negatively formulated items as 1. This kind of standard represents the *maximum view* that it is crucial for competent self-regulation of learning to know as many learning strat-
egies as possible and to apply them at all times. In other words, the more strategies learners know and the more they use them, the better their score.

Questionnaires and inventories on learning strategies have been subject to much criticism in the last 10 years (e.g., Artelt, 2000), not least because their correlations with learning outcomes have been found to be very low, indicating low validity (e.g., Veenman, 2005). This low validity was primarily ascribed to the fact that respondents filling in learning-strategy questionnaires have to reflect retrospectively on their learning strategy use. For example, learners responding to the item cited above have to recall past learning situations and “count” how often they related the learning material to their prior knowledge in those situations. It is obvious that the result depends on the learning situations recalled. The sample of recalled learning situations is probably not representative for all of a learner’s learning situations. For example, learning situations from yesterday are much more easily recalled than learning situations that took place a month ago. Thus, reflections of this kind are highly error-prone and can easily result in inaccurate and invalid ratings (Lompscher, 1994). In addition to these problems, however, it can be assumed that the kind of standard that these questionnaires and inventories use for scoring—a maximum value in the sense of a quantitative offline standard—negatively affects the validity of these assessments. There is no theoretical model corroborating the maximum view that it is crucial to know as many learning strategies as possible and to apply them whenever possible (Leopold, den Elzen-Rump, & Leutner, 2007; Leutner & Leopold, 2006). Learning can be regulated very well with knowledge of just a few strategies, as long as they are appropriate under the current task conditions, and the learner is proficient in applying them. Thus, quantitative offline standards do not have a solid theoretical foundation, and must, therefore, be considered as standards of low validity.

**Qualitative Offline Standards**

Schlagmüller and Schneider (2007) have developed an instrument assessing respondents’ knowledge of learning strategies when learning from texts. Comparable instruments are available for the areas of mathematics (developed by Artelt, see Ramm et al., 2006) and science (Thillmann, 2007). In contrast to the questionnaires and inventories described above, these tests use a qualitative offline standard. Respondents first read a description of a specific learning task. They are then presented with five to seven detailed descriptions of learning approaches of differing strategic quality. Respondents are asked to rate the learning approaches from very good (1) to deficient (6), according to their appropriateness for accomplishing the learning task.

The definitions of the standards are based on expert ratings. Learning strategy experts in the respective field also rate the quality of the different learning approaches presented. Paired comparisons are used to determine whether the experts, on average, rated each learning approach to be better than, the same as, or worse than every other learning approach. These paired comparisons serve as a qualitative offline standard. The same paired comparisons are computed for each respondent and compared with the experts’ paired comparisons. Respondents receive the highest score if their paired comparisons are the same as the experts’ paired comparisons. In other words, the better the correspondence between the respondents’ ratings and the experts’ ratings, the higher their score.

This kind of assessment method does not assess how often students use a certain learning strategy (quantitative standard), but whether they are aware of the conditions under which certain learning strategies are applicable and appropriate (qualitative standard). This kind of standard, thus, represents an optimum view rather than a maximum view.

The validity of assessment methods using qualitative offline standards seems to be high. Schlagmüller and Schneider (2007) reported correlations with external criteria from \( r = .41 \) (intelligence) to \( r = .48 \) (reading literacy). Artelt, Demmrich, and Baumert (2001) compared learning strategy tests using quantitative offline standards and a learning strategy test using a qualitative offline standard in terms of their predictive power for reading literacy. They used a structural equation model that included interest and reading self-concept as predictors of reading literacy. It emerged that learning strategy tests with quantitative offline standards had only very limited predictive effects on reading literacy (\( \beta = .05 \) in this model, whereas the learning strategy test with qualitative offline standards had very strong predictive power (\( \beta = .48 \)).

Although there is only a short history of tests using qualitative offline standards in the field of self-regulated learning, there is convincing empirical evidence that using qualitative instead of quantitative offline standards improves test validity. To date, however, these kinds of tests are only available for assessing (conditional) knowledge about learning strategies. Yet learning strategy knowledge is just one of the components of self-regulated learning competence. Further research is necessary to examine whether qualitative offline standards are also appropriate for assessing other competence components of self-regulated learning (e.g., goal setting or monitoring).

In a project funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) we are currently taking a number of steps in this direction. For example, we aim to develop and to evaluate different assessment methods using qualitative offline standards to gauge the competence component of goal setting. Respondents are given a science text and instructed to learn its content. In order to assess their competence to analyze the task conditions, we ask them before working with the text to rate the text with respect to typical attributes that make its content difficult or easy to learn (e.g., whether or not the text contains many unknown words; whether or not the text structure is complex). The respondents’ ratings are then com-
pared with experts’ ratings. Respondents receive a high score if their ratings correspond with the experts’ ratings. We also ask the respondents to rate their content-specific prior knowledge (cognitive conditions). They are then administered a content-specific knowledge test, and receive a high score if their (subjective) rating corresponds with their (objective) score in the content-specific knowledge test. In other words, even if respondents have only limited content-specific prior knowledge, they can receive a high score on this competence component of self-regulated learning if they also rate their prior knowledge to be very low.

Likewise, we aim to develop and to evaluate different tests using qualitative offline standards to gauge the competence component of monitoring. For example, we ask respondents to set themselves specific learning goals before reading the science text. After they have read the text, we ask them to rate whether they have achieved their learning goals. They are then administered a content-specific knowledge posttest. By comparing their performance in the posttest with their predefined learning goals, we determine how far they have achieved their learning goals, and we use the result of this comparison as the qualitative offline standard for evaluating the students’ own rating. Respondents receive a high score if their own rating of the degree to which they have reached their learning goals corresponds with the qualitative offline standard. In other words, even if respondents have not reached their own learning goals, they can receive a high score on this competence component of self-regulated learning if they also rate their goal attainment to be low.

Online Standards

Assessment methods using (quantitative or qualitative) offline standards draw on component models of the competence of self-regulated learning; they aim at testing components that are regarded as prerequisites of self-regulated learning. However, assessments using offline standards do not assess whether students actually use their competencies when working on a specific learning task. Methods aiming to assess the application of self-regulatory competencies have to use online standards. Therefore, they have to draw on process models of self-regulated learning. Given that both task conditions and cognitive conditions change continuously during the learning process, online standards also change continuously (Wirth & Leutner, 2006). Against this background, methods that assess aspects of the learning process online are required. Computer-based methods are commonly used for online assessments. Wirth (2004, 2005) presented a computer-based assessment of self-regulated learning that combines quantitative and qualitative online standards. In the following, we describe this approach, focusing first on the quantitative aspect of the standard and then describing the use of the qualitative aspect for comparing respondents’ observed behavior with the standard.

Quantitative Online Standards

Wirth (2004, 2005) examined learners’ self-regulation as they learned to direct an interactive system developed by Funke, Töpfer, and Wagener (1998, see Figure 1). Specifically, their task was to learn how to direct a computer-simulated rocket from one planet to another planet. After landing on the second planet, respondents had to learn how to direct a space buggy from the rocket to a diamond hidden on the planet. Learning in interactive learning environments like this requires the learner to pursue two goals simultaneously. On the one hand, the learner has to generate new feedback information about the system by interacting with the system. The learner manipulates input variables of the system and receives feedback about the effect of his or her manipulation. In the system used by Wirth (2004, 2005), learners manipulated the input variables by clicking on one of the 20 action buttons presented in the upper panel of the screen. Feedback was given by 20 signals in the lower panel. Every time the learner manipulated a specific input variable for the first time, he or she was given specific feedback information for the first time. This feedback is considered as new information generated by the learner’s interaction with the system.

On the other hand, the learner has to integrate the feedback information into his or her knowledge structure. For example, the learner can manipulate the same input variables (e.g., click on the same button) again and again, always receiving the same feedback information. This kind of a rehearsal strategy supports learning in the sense that it increases the probability that the learner will be able to access and apply this information (almost automatically) on later occasions (Anderson, 1982).

The activities of generating new information differ from the activities of integrating information; thus, learners always have to decide whether they want to generate or to integrate information. It was assumed that competent self-regulated learners start by generating information, but very soon turn to integrating, so that they do not forget the feedback information they have already generated. In order to test this specific hypothesis about competent self-regulation of the learning process, Wirth (2004, 2005) categorized every mouse click in terms of whether it generated feedback for the first time (generating mouse click) or whether it produced feedback that had already been generated (integrating mouse click). The ratio of integrating mouse clicks to generating mouse clicks was then computed (ratio_{observed} = integrating mouse clicks/generating mouse clicks). This ratio_{observed} was used as an index of a respondent’s observed behavior.

Within every system state that students visited, they had to choose from a set of 20 possible generating or integrating actions. That is, they had to choose from a number of action
buttons that they had not clicked on before in this specific system state and from a number of buttons they had already clicked on in this specific system state. Each of these action buttons was classified in terms of whether a respondent had already clicked on it within the specific system state (integrating option) or not (generating option). A ratio of the number of integrating options to the number of generating options was computed (ratio\text{standard} = \text{integrating options}/\text{generating option}). An index of a respondent's observed behavior exactly matching this ratio (ratio\text{observed} = \text{ratio standard}) indicates that a learner's behavior does not differ from the behavior that can be expected for any respondent given the options currently available independent from any regulatory attempts. That is, this respondent-specific ratio represents a behavior that is totally random, and it must be assumed that this behavior is not at all regulated by the learner. This ratio\text{standard} served as an online standard.

A respondent's observed behavior was compared with this theoretical standard by computing the logarithm of the ratio of the respondent's ratio\text{observed} to the respondent's ratio\text{standard} (logor = \ln(\text{ratio observed}/\text{ratio standard})). The resulting index was called logor (“log odds ratio”). The more the logor score differed from zero, the more a respondent's observed behavior differed from random behavior. That is, it can be assumed that the more a logor score differed from zero, the more the learning behavior was regulated by the learner.

For theoretical reasons, it must be assumed that a score such as the logor score that only values how much a learner's behavior differs from a random behavior, irrespective of the sign of the score (i.e., a score using only a quantitative online standard) is of low validity. For example, consider a learner who decides to click repeatedly on just one button throughout the whole learning process. Obviously, this results in a learning behavior that is very different from random behavior and is, therefore, evaluated as highly self-regulated. It can be assumed that this learner will gain knowledge about the feedback received by clicking on this specific button. But, of course, he or she will not learn anything about the feedback available by clicking on any of the other buttons. Thus, the learning outcome will be very

Figure 1. User interface of the interactive computer-simulated learning environment used by Wirth (2004, 2005).
low, and the quality of the self-regulation of the learning process must also be considered very low. As a consequence, it does not seem sufficient to consider only the quantitative aspects of an online standard; the qualitative aspects must also be considered.

Qualitative Online Standards

The $\log_{\alpha}$ score indicates not only how much learners regulate their learning, but also whether learners focus on generating new information or whether they try to integrate information. A $\log_{\alpha}$ score less than zero indicates that a learner generated more new information than expected from random behavior. A $\log_{\alpha}$ score greater than zero indicates that a learner integrated more information than expected from random behavior. As mentioned above, it was assumed that competent self-regulated learners start by generating new information, but very soon turn to integrating, so that they do not forget the feedback information they have already generated. Thus, there is a theoretically defined optimal sequence of regulation from generating to integrating.

In order to test whether learners followed this theoretically optimal sequence of regulation, Wirth (2004) divided the learning process into time intervals and computed the $\log_{\alpha}$ scores for each time interval. A good fit of a respondent’s learning behavior with the optimal sequence of regulation was expressed by negative $\log_{\alpha}$ scores for the first time intervals and positive $\log_{\alpha}$ scores for later time intervals. Latent growth curve models of the learner’s behavior revealed that learners whose behavior corresponded with the theoretically defined optimal sequence of regulation gained more knowledge about the interactive system than learners whose behavior was either highly regulated, but not in the theoretically optimal way, or not regulated at all. Thus, this combination of a quantitative and a qualitative online standard allowed a valid assessment of self-regulated learning competence.

Conclusion

There are a number of theoretical models of self-regulated learning. These models differ in the aspects of self-regulated learning emphasized. Furthermore, they differ in whether they describe self-regulated learning in terms of a number of competence components seen as prerequisites for the competent self-regulation of a learning process (component models) or whether they describe the process of self-regulation in terms of an optimal sequence of phases (process models). Although the different models of self-regulated learning differ markedly in some aspects, they also share some commonalities. All models incorporate at least the three metacognitive demands of goal setting, planning, and monitoring. Furthermore, all models (more or less explicitly) identify the competence to make solid comparisons as a key competence of self-regulated learning.

The methods used to assess self-regulated learning also implement comparisons as a crucial feature of the scoring procedure. However, they differ in terms of the kind of standards they use for these comparisons. These standards are determined by the theoretical model that test developers use to define what a test intends to assess. For example, component models imply the use of offline standards, whereas process models imply the use of online standards.

The use of qualitative standards requires a model of self-regulated learning that defines self-regulatory aspects in terms of optimal values (optimum view) rather than maximum values (maximum view). Qualitative standards are not necessarily high; they may be very low. For example, when assessing whether learners are able to estimate their own content-specific knowledge, the standard value is low if the learner’s content-specific knowledge is low. In contrast, quantitative standards are maximum values. For example, a qualitative standard concerning learning strategy knowledge represents the maximum view that it is always appropriate to have as many learning strategies as possible available and to use them whenever possible.

Comparison of Different Types of Standards

Comparison of assessment methods applying offline standards with assessment methods applying online standards does not reveal either kind of standard to be superior or inferior to the other. Both can lead to highly valid assessments of the aspect of self-regulated learning competence under investigation. Whether test developers use offline or online standards depends primarily on the theoretical model. Offline standards may have economic advantages over online standards; they can be used in paper-and-pencil tests that can be administered parsimoniously. Online standards depend on the use of computer-based assessments to reflect the continuously changing conditions of both the task conditions and cognitive conditions. This probably explains why most assessment methods of self-regulated learning use offline standards (Biggs, 1978; Entwistle, 1988; Pintrich et al., 1991; Schlagmüller & Schneider, 2007; Weinstein, 1987; Wild et al., 1992), although offline standards are also used for computer-based tests (e.g., Jamieson-Noel & Winne, 2003).

Comparison of quantitative standards with qualitative standards reveals the great advantage of qualitative standards in terms of the validity of the resulting assessment methods. This advantage applies to both qualitative offline standards and qualitative online standards. The reason seems to be that the theoretical models underlying qualitative standards are more elaborated than the models underlying quantitative standards. For example, a model assuming that a specific learning strategy is always help-
ful (maximum view) provides only a rough estimation of the appropriateness of a specific learning strategy for a specific learning task. In contrast, a model assuming that a specific learning strategy is optimal under specific task conditions and cognitive conditions (optimum view) can provide a much more accurate estimation if the conditions of a specific learning task are known. Assuming that a specific learning strategy is optimal only under certain conditions is a much more elaborated theory than assuming that a learning strategy is appropriate, without specifying under which conditions. It seems reasonable that qualitative standards of comparisons that are based on more elaborated theoretical models lead to more valid comparisons and assessments than do quantitative standards that are based on less elaborated models. First empirical evidence supporting this view has been reported by Artelt, Demmrich, and Baumert (2001) and Wirth (2004, 2005).

A substantial number of assessment methods have been developed in the last 20 years in the field of self-regulated learning. Although research on self-regulated learning is quite lively and productive, most of these methods suffer from a lack of validity (e.g., Veenman, 2005). Furthermore, most are restricted to assessing the use of learning strategies. The crucial prerequisite for valid assessment is an elaborated theoretical model. Theoretical models of self-regulated learning have become increasingly refined in the last 10 years, thus, affording the opportunity to develop better assessment methods, for example, by using qualitative standards. Empirical findings on the validity of recently developed methods using qualitative standards (Artelt, see Ramm et al., 2006; Schlagmüller & Schneider, 2007; Thillmann, 2007) are promising. However, these assessment methods are also restricted to (conditional knowledge of) learning strategies. From the perspective of a component model of self-regulated learning, conditional knowledge of learning strategies is just one of several competence components necessary for competent self-regulation. There is clearly a need to develop and evaluate methods assessing further components of self-regulated learning competence. It seems to be possible to use qualitative standards for these developments. Experiences to date with qualitative standards suggest that this is indeed a promising avenue of research.

Acknowledgments

The preparation of this paper was supported by grant WI 2663/4 from the German Research Foundation (DFG) in the Priority Program “Competence Models for Assessing Individual Learning Outcomes and Evaluating Educational Processes” (SPP 1293) and by grant LE 645/8 in the DFG Research Unit “Research on Teaching and Learning Science” (FOR 511). Joachim Wirth is now affiliated with Ruhr University Bochum, Germany.

References


Joachim Wirth

Department of Educational Research
Ruhr-University Bochum
P.O. Box 102148
D-44721 Bochum
Germany
Tel. +49 234 32-28728 / 32-22728
E-mail joachim.wirth@rub.de