

# Modeling the Development of Vocational Competence: a Psychometric Model for Economic Domains

Viola Katharina Klotz<sup>1</sup>  · Esther Winther<sup>1</sup> ·  
Dagmar Festner<sup>2</sup>

Received: 14 January 2015 / Accepted: 7 August 2015 /  
Published online: 25 August 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** This article discusses the development of vocational competence through economic vocational educational training (VET) from a theoretical and psychometric perspective. Most assessment and competence models tend to adopt a state perspective toward assessments of competence and carve out different structures of competence for diverse vocational domains. However, the order and at what stages of development these identified structures actually occur remains uncertain. This study therefore moves beyond a static perspective to denote changes in competence over the duration of vocational training, using item response theory-based scaling and a cross-sectional database of 877 economic apprentices. The resulting four-stage psychometric model represents a systematization of the development of vocational competence, characterized by the degree of occupational specificity and different forms of cognitive processing. This proposed psychometric model can be used to inform educational researchers and practitioners about the different stages of competence development, such that they can both assess and teach economic competence more effectively.

**Keywords** Competence development · Vocational learning · Learning progression · Item response theory

---

✉ Viola Katharina Klotz  
klotz@die-bonn.de

<sup>1</sup> German Institute for Adult Education (DIE), Heinemannstraße 12-14, 53175 Bonn, Germany

<sup>2</sup> University of Paderborn, Warburger Str. 100, 33098 Paderborn, Germany

## Theoretical Background

### Research Context and Research Questions

A detailed understanding of the stages through which vocational competence develops is essential for assessing learners' proficiency and designing vocational programs and training materials that can facilitate the acquisition of such competence. In any such endeavor, it also is essential to identify, for each stage, which capacities a learner has acquired and which more sophisticated capacities he or she thus is in a position to attain. Existing qualitative research offers support for the idea of an expert-novice paradigm, based on situated performance and experiential learning, in the vocational domains of chess playing and air force piloting (e.g., Dreyfus and Dreyfus 1980), nursing (Benner 1984, 2004; Benner et al. 1996), management (Worthy 1996), social work (Ryan et al. 1995), and computer programming (Campbell et al. 1992; Chmiel and Loui 2004). However, no quantitative evidence affirms this approach with psychometric modeling. Despite advancements in psychometric theory and procedures (e.g., Pellegrino et al. 2001; Pellegrino 2012; Wilson 2005, 2008) as well as tremendous progress in the past decade by models of vocational competence structures and levels (e.g., Nickolaus et al. 2008; Seeber 2008; Winther and Achtenhagen 2009), modeling of competence *development* remains in its infancy. One key issue is the potential to conceive of competence development in two conceptually distinct ways, though these ways are not contradictory and can be combined (Wilson 2009). First, it may entail a change in cognitive structures, in the sense of an integration of cognitive elements (in psychometric contexts, usually referred to as factors) due to cognitive cross-linking or a differentiation of various cognitive elements due to the creation and contouring of new cognitive elements (e.g., Gschwendtner 2011; Nickolaus 2011; Klotz and Winther 2015; Klotz 2015). Second, it can be an increase of proficiency with respect to solving tasks with qualitatively described degrees of difficulty (in psychometric contexts, usually referred to as levels). The approach surely affects the corresponding measurements used. The former approach relies on factor analyses, whereas the latter models and locates competence in one coherent scale, as we detail in this contribution.

We focus on the latter approach to model different aspects of competence as a coherent, composed score that represents a vector of competence (DiBello et al. 2007; Zhang and Stout 1999) and to track vocational competence development as learning progression on this unidimensional scale via a cross-sectional design. Our objective is to confront theoretical considerations of vocational development (Dreyfus and Dreyfus 1980; Gelman and Greeno 1989) with psychometric procedures and thereby establish an empirically validated developmental model that can account for both situational and process-related aspects of vocational competence development.

When using the term “competence”, we understand it in line with Mulder et al. (2006), as the capability to perform by using knowledge, skills, and attitudes that are integrated in the professional repertoire of the individual; it is therefore reasonable to deduce cognitive structures from the solution of authentic situations (performance), assuming adequate item design and psychometric procedures (e.g., Chomsky 1965; Shavelson 2008; Wilson 2008). However, in this contribution we specifically consider

knowledge and ability, not attitude-related aspects of vocational competence in terms of motivation and volition.<sup>1</sup>

Moreover, when we use the term “development”, we refer to developmental changes on a cohort level exclusively. Our cross-sectional research design only permits the examination of collective developmental trends in competence development over the span of vocational education and training (VET) *between* cohorts. As our sample contains entirely different cases for the beginning and the end of vocational training, our cross-sectional design is thus inappropriate for studying individual developmental patterns *within* cohorts and is unable to resolve issues of causal order (e.g., Menard 2002, p. 29).

With respect to this approach towards competence development our research endeavor refers to the question “how does apprentices’ vocational competence develop during VET on a collective level?”. In line with Dreyfus and Dreyfus (1980), we examine if learners in vocational education and training (VET) follow an aggregate trajectory, from having mainly non-situational, decontextual preknowledge to gaining domain-specific competence that enables them to handle vocation-specific situations proficiently. With respect to cognitive processing, we further predict that learners during VET move from conceptual to procedural and finally to interpretational competence.

## Developmental Assumptions

Most discussions of the *development* of vocational competence rely on two theoretical expectations: (1) the expert–novice paradigm (Dreyfus and Dreyfus 1980) or (2) the idea of domain-linked and domain-specific competences (Gelman and Greeno 1989). These two theoretical approaches are not contradictory; rather, the latter seeks to model a more detailed specification of the underlying cognitive processes that appear in the former theory’s general description of the stages of vocational development.

Dreyfus and Dreyfus’s (1980) model suggests what a practitioner has achieved at each of five distinct stages and what higher order skills he or she is then ready to pursue, from being (1) a novice, to showing (2) competent and (3) proficient behavior and finally, after tremendous vocational experience, to acting as (4) an expert or even (5) a master. The first three stages seemingly might occur during initial vocational training; the last two stages require a long-term perspective, such as when advanced professionals develop intuition and are intensely absorbed in their work activities. Dreyfus and Dreyfus (1980) describe the first three stages as an expansion of novices’ non-situational, decontextual preknowledge, which grows to include relevant knowledge about aspects, specific guidelines, and action maxims, such that it acquires an increasingly organized shape as specific knowledge. The newly acquired knowledge also is situational: It comprises meaningful elements of vocational experience and is stored to provide the learner with a basis for deciding how to act in similar situations (Dreyfus and Dreyfus 1980).

This conceptualization corresponds with Gelman and Greeno’s (1989) conception of domain-linked and domain-specific competence. In line with Gelman and Greeno

---

<sup>1</sup> However, such aspects also influence task solutions and therefore presumably are integrated in our measurement approach to some (unknown) extent.

(1989), we assume domain-linked and domain-specific aspects of vocational competence, such that the former is generally relevant in decontextualized form to the economic domain, and the latter is highly situational and reflective of vocation-specific aspects, guidelines, and action maxims. The theoretical foundation of this conceptualization lies in the idea that learners acquire specific rules of practice through a set of general principles that support selective attention and guide further learning in a given domain. The domain-linked category precisely refers to key skills, or knowledge and ability that is general but also relevant for solving vocational problems. The domain-linked category therefore refers to the notion of occupational key skills, which comprise knowledge about general methods and strategies (Rychen and Salganik 2003). In an economic domain, concepts such as literacy and numeracy represent this type of non-specific preknowledge (OECD 2003; Winther and Achtenhagen 2009). An example for domain-linked competence would be the ability to perform simple exchange-rate calculations. Those calculations as such do not require any specific vocational knowledge or ability, but can be coped with by simply applying the general mathematical concept of “rule of three”, which learners are already familiar with from their general school education. Domain-specific competence instead entails specific occupational knowledge and skills, including occupation-specific contents and job- or enterprise-specific rules and skills (Oates 2004), which are highly situated. In an economic context, such knowledge might consist of legal accounting rules for preparing a balance sheet, for example, or the specific rules of a purchase contract, which are newly acquired during VET.

Gelman and Greeno (1989) further describe three kinds of cognitive processes that *complement* these categories: conceptual, procedural, and interpretational. Together, these competences represent an action schema for performing vocational tasks (Gelman and Greeno 1989; Greeno et al. 1984). First, conceptual competence implies an understanding of the principles in the domain and corresponds to factual knowledge that can be translated into an action schema. Second, procedural competence is an understanding of the principles of action, which usually takes the form of knowledge applications, such as ways to operate with facts, structures, knowledge nets, and their corresponding elements. Third, interpretational competence focuses on appropriate strategic decision-making processes that reflect a grounded interpretation of the results obtained through conceptual and procedural competence. This last category therefore entails the appropriate application of conceptual and procedural competence and constitutes the most complex and difficult ability (Shavelson 2008).

A combination of the forms of cognitive processing with different types of occupational knowledge (domain-linked and domain-specific) in the sense that these two aspects are complemented with each other allows cognitive psychology to account for the situated nature of vocational learning. In line with Billett (1994), we argue that novices do not necessarily lack cognitive ability. Rather, in most cases they lack specific knowledge and experience within a particular domain (Glaser 1990), which would otherwise enable them to conceptualize and categorize problems and deploy their cognitive structures more effectively (Billett 1994). This perspective also suggests a distinct understanding of domains, compared with formal academic disciplines (Billett 1994), in that domains are shaped by social circumstances within a particular vocational “community of practice” (Lave and Wenger 1991, p. 29), rather than reflecting an inert set of principles. The development of vocational competence in a

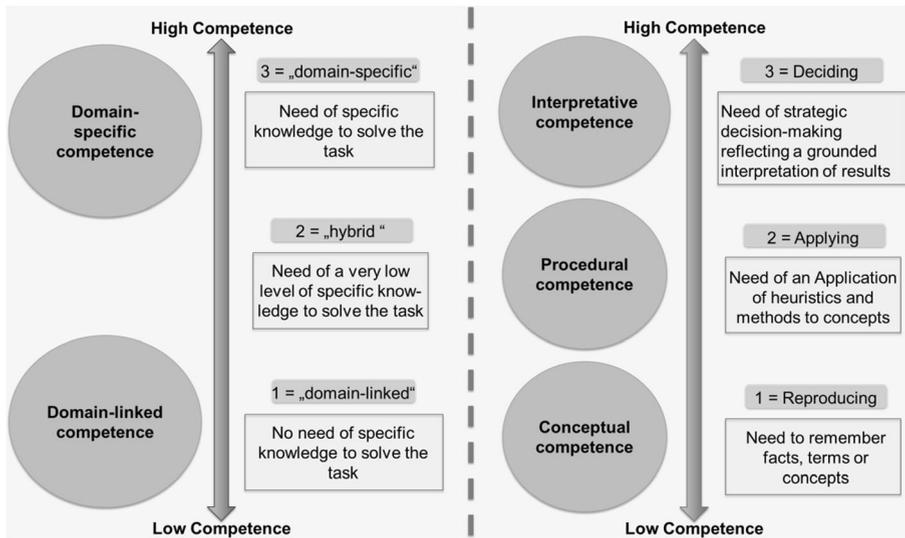
vocational domain thus is premised on specific problems and rules of practice (Billett 1994; Brown et al. 1989; Lave 1990). The exigency of situated knowledge and ability in a specific vocational domain links the two research approaches we have introduced.

### Theoretical Outline of the Learning Progression

The concept of learning progression is undergoing swift development, though it is really just the latest manifestation of a much older idea, namely, the concept of regular progress in acquiring a certain body of knowledge or professional practice (Wilson 2009). Learning progressions describe the sequential ordering of a set of vocational skills or pieces of knowledge, from least to most complex, and they describe successive levels of proficiency in a domain that reflect theoretical descriptions and examples of expected performance for a given level of competence. They also accord with a cognitive-constructivist view of vocational learning, in which competence builds on itself incrementally within the individual (Johnson and Tymms 2011; Phillips 1995). To establish the ordering of fields of actions within a vocational domain (e.g., nursing, banking), a structuring theoretical framework is required, which Wilson (2005, 2008) refers to as a “construct map”. Construct maps can be derived from theoretical research into the underlying cognitive structure of the vocational domain, as well as from expert judgments about what constitutes higher or lower levels of competence; they also might be informed by empirical research into how students perform in practice or during simulated tasks (Pellegrino et al. 2001; Wilson 2009). By implication, such sequences entail stages through which learning progresses. A straightforward way to regard the relationship of a construct map with a learning progression is to recognize how the levels of the construct map link to the levels of the learning progression, such that the construct map serves as a “skeleton” of the body of learning progressions (Wilson 2009). The learning progression then provides information about qualitative and quantitative progress; successive levels represent the successive stages of sophistication in the learner’s ability, and movement within a single level indicates that learners have become more sophisticated with respect to that level (e.g., wider applicability) (Wilson 2009).

With respect to an application of this idea of a learning progression to our research endeavor, we use the extent of content specificity (domain-linked versus domain-specific competence as well as a middle ranged option if the task was partially solvable without any specific knowledge) as well as the level of cognitive processing (conceptual, procedural and interpretative) as difficulty generating features to conceptualise a construct map of vocational competence (see Fig. 1) on which by implication learning progressions should be trackable.

Connecting the construct map with our theoretical developmental framework of Dreyfus and Dreyfus’ (1980), we assume that learning progressions in vocational domains are characterized by the extent of content specificity as well as the level of cognitive processing. More precisely, within this model we expect that novices expand their domain-linked competence, to include relevant domain-specific knowledge about facts and action maxims. We further assume that novices develop their cognitive abilities with respect to vocational tasks, to include procedural and interpretational ability at the end of their training. Therefore we expect an upward movement on our construct map for both the extent of content specificity as well as the level of cognitive processing.



**Fig. 1** Construct maps of specificity and cognitive processing

For research in vocational competence, such a heuristic appears promising, especially for the efforts (1) to explicitly describe increased proficiency or qualitative changes in performance ability in various (occupational) situations, which relates closely to a latent notion of competence with respect to performance (e.g., Chomsky 1965), and (2) to develop learning progressions through backward analysis (e.g., Taylor 2013) by first identifying an ideal point of desired ability and then the skills required to move from a naive (novice) to an advanced (depending on the vocational assessment aim, namely, to be competent, proficient, or expert) level. A learning progression thus is criterion based by nature, which is important for teaching and assessing vocational competence, because in many instances, some standard of vocational ability that indicates the quality of a certain set of vocational activities is of interest, rather than a mere test score.

### Assessment Design

The design of a vocational learning progression requires either strong familiarity with the vocational area or interactions with vocational experts to determine which vocational activities are more or less complex and how the capacity to achieve them might develop over time (Duncan et al. 2009; Plummer and Krajcik 2010). Ideally, the progression model reflects the structures of the vocational curriculum and instruction (Pellegrino 2012) and includes examples of student work at each level (Mislevy and Haertel 2006; Taylor 2013; Wilson 2005). To apply these design-guidelines as part of the operationalization process, the assessment tasks designed to measure economic knowledge and skills in the commercial sector refer to curriculum-based economic situations that represent job-related skills in the field of economics and business management and were implemented with a paper-and-pencil test. The item format of all tasks was open ended. The tasks were developed binary (and coded 0 for “none or a wrong solution” and 1 for “a correct answer”) or partial credit (and then scored with 0

for “none or a wrong solution”, 1 for “a partially correct answer” and 2 for “a fully correct answer”). In our design, vocational beginners had to complete 34 tasks, advanced vocational learners had to complete 28 tasks, and 16 items were identical for both groups allowing for a design-based linking of the two groups via ConQuest.

To model learning progression in line with our theoretical framework, we developed 28 items to assess domain-linked and 18 items to assess domain-specific competence. An example domain-linked task would be an exchange rate calculation, which is part of the vocational curriculum but can be completed simply by applying a general numerical skill such as the rule of three. Domain-specific tasks instead demand specific vocational knowledge, such as requiring the correct application of economic contract theory to a given situation (see item 6 in the Appendix). Learning different vocational activities also involve varying cognitive demands, as noted by Gelman and Greeno (1989). To clarify these processes, we designed 15 items to reflect conceptual, 21 items for procedural, and 10 items to indicate interpretative competence.<sup>2</sup> Conceptual items target learners’ declarative knowledge in the economic domain; procedural items require learners to apply their knowledge by analyzing a given situation (e.g., item 6 in the Appendix); and the interpretative items demand that learners make strategic decisions on the basis of their declarative and procedural knowledge (see item 5 in the Appendix). As items for interpretative competence are more difficult/complex they require more test time. So with respect to test time interpretational items are adequately represented. We also designed more items on a procedural level as the stress of the vocational curriculum lies in an application of vocational knowledge which would be “procedural competence”.

Our model is somewhat taxonomic, in the sense of a cumulative ordering. Solving domain-specific tasks certainly requires some domain-linked competence (economic numeracy and literacy). Furthermore, with respect to cognitive processing, reasoned decision making based on an interpretation of attained results usually requires reflection on the concepts and procedures used. For example, item 5 in the Appendix requires the highest level of vocational competence in terms of specificity, as well as with respect to cognitive processing. Here learners need to make a decision for one supplier by analyzing and weighing several competing aspects like social and environmental conditions as well as aspects of pricing and the production process. Meanwhile other items require only specificity or cognitive processing at a lower level. To validate our test design as an adequate categorization of the developed items in our model, we asked 24 vocational experts (12 experts for each item) to rate all tasks on a three-point Likert scale with respect to an authentic item design as well as with respect to their estimation of domain-specificity and cognitive processing.

## Methodology

In quantitative terms, learning progressions can be validated with latent class analysis (Steedle and Shavelson 2009) or Rasch item difficulty estimates (Corcoran et al. 2009;

<sup>2</sup> Regarding the 15 items to reflect conceptual competence, 5 items were domain-linked and 10 domain-specific. From the 21 items for the measurement of procedural competence 16 items were domain-linked and 5 domain-specific. Finally, the 10 items depicting interpretative competence were comprised by 7 domain-linked and 3 domain-specific items.

Johnson and Tymms 2011). However, only the latter approach locates the items in a developmental scale, such that it offers a potentially more powerful tool for exploring learning progressions (Pellegrino et al. 2001). This method also connects more readily to qualitative expert judgments through Hartig's (2007) procedure, such that we can derive a well-reasoned, criterion-based model of learning progression. Specifically, IRT-based models assume a single trait, described by an interval (logit) scale (Wright and Stone 1979), on which both the person's ability and item difficulty can be estimated; negative values indicate below-average and positive values above-average difficulty or competence. The chance of success thus depends on the difference between the person's ability and the item's difficulty. The measured ability of the person does not depend on which items are attempted (in psychometric literature, this useful characteristic is referred to as "objective specificity"; Rasch 1977). If the items represent understanding of certain ideas, the order of difficulty can be inferred to represent the order in which learning progresses (Johnson and Tymms 2011)—or for our purposes, how competence gains in a certain vocational domain are achieved.

To analyze open-ended items, we used a multidimensional random coefficient multinomial logit model (MRCML) (Adams et al. 1997) and assessed the resulting polytomous database of varying scaling within the program framework of the ConQuest software (Wu et al. 1997).

To identify progress in the "mass" of ability, as well as with respect to different competence qualities, we asked 24 experts to rate the specificity and cognitive processing level associated with our items. Using Hartig's (2007) method, we then identified thresholds on the logit competence scale by regressing experts' estimation of the difficulty of each task (the degree of specificity and cognitive processing as dummy-coded variables) on the item difficulty of the logit scale (for a more detailed description, see Hartig 2007; Hartig et al. 2012). Thus we estimated thresholds for the two groups, vocational beginners (group 1: 0.0–0.5 years of initial vocational training) and advanced learners (group 2: 1.3–2.8 years of initial vocational training). A downward shift in the level of difficulty in a comparison of group 1 with group 2 would mean that the learners must have progressed in that competence quality (as determined by expert judgments), because the items on that level became easier for them to solve.

## Data Acquisition

With respect to assessment, we used a cross-sectional design, though with this design, we cannot estimate or explain individual differences within the cohort. On the other hand, longitudinal data would have caused learning effect issues, such that test results might improve not due to increased competence but rather due to learners' recognition of tasks they have previously seen (e.g., Hoffman et al. 2011; Salthouse and Tucker-Drob 2008). The bulk of these cross-sectional data were gathered from visits to four vocational schools across Germany in Munich, Hanover, Bielefeld, and Paderborn in 2013. The schools were selected as they had the most industrial apprentices within their respective region. It is noteworthy that the data were acquired within the German VET system which is constituted by the duality of (1) company-based training programs (about 3 days per week, provided by the private firm sector) accompanied by (2) school-based components (about 2 days per week, provided by the public sector).

Within this VET structure apprentices acquire practical and theoretical knowledge with respect to their specific training occupation. Table 1 presents the sample, classified into the two groups of vocational beginners ( $n_1=398$ ) and advanced learners ( $n_2=479$ ), and their basic characteristics. Though the data is a non-random-sample, the two subsets are remarkably similar with respect to the distributional characteristics of all collected variables but on average differ in their consumed instructional VET time (Average years of initial VET) and respectively their age.

During these firsthand observations, with respect to test motivation we observed that the students engaged very well with the instrument—probably because it was presented as a preparation for their final vocational examination, and we promised and provided feedback, which also likely explains the low rate of missing values (3.37 %). The solutions were corrected and coded according to a detailed assessment scoring guide (Wilson 2008). Two graders corrected 16 % of all 877 tests independently, to estimate the accuracy of the correction process. The intraclass correlation coefficient (ICC) indicated a satisfactory degree of scoring objectivity (ICC [3:1]=0.916).<sup>3</sup>

## Results

The data yielded good item fit-indices for a unidimensional model ( $.90 \leq \text{wMNSQ} \leq 1.17$ ;  $0.1 \leq \text{T-values} \leq 2.30$ ;  $-2.125 \leq \text{Item Thresholds (difficulty)} \leq +3.158$ ;  $0.057 \leq \text{measurement error} \leq 0.089$ ) and neglectable correlations close to zero between the resulting error terms in a unidimensional model,<sup>4</sup> suggesting that our construct can be considered a unidimensional vector score.

We then proceeded with an application of the Hartig-Method (Hartig 2007; 2012). To estimate the threshold change in different qualities of competence over the vocational trajectory, we first determined the item difficulty for vocational beginners. Regressing expert judgments of specificity and cognitive processing on item difficulty yielded the regression model for group 2 that we present in Table 2.

As this table shows, the item characteristics that significantly contributed to item difficulty were moderate and high levels of specificity, as well as cognitive processing that requires interpretative decision making (see Fig. 1).<sup>5</sup> The effect sizes reveal that the item characteristic of specificity is primarily what makes items difficult for learners ( $\beta=0.252$  for moderate and  $\beta=0.666$  for high degree of specificity). The quality of cognitive processes has a medium-sized effect ( $\beta=0.564$ ). The Hartig (2007) method uses these effect sizes to develop a model-based description of the thresholds for competence-based level models. Specifically, the method adds up the unstandardized regression coefficients in a theoretically reasonable order: The intercept ( $-2.012$ ) forms the first threshold, indicating the difficulty of

<sup>3</sup> This ICC formula is used when all subjects are rated by the same raters (Two-way-mixed; absolute agreement (unjusted); single measure) (e.g., Shrout and Fleiss 1979).

<sup>4</sup> The correlations between the error terms of a unidimensional model ( $0.001 \leq \text{correlations of error terms} \leq 0.221$ ) were calculated as a Q-matrix in the software environment R.

<sup>5</sup> It should be noted that due to multicollinearity not all of the original levels of cognitive processes and domain specificity outlined in our conceptual construct map made it into the final regression model, even though all aspects were investigated. According to the Hartig-method only the most strong dummy characteristics in terms of effect size are taken into consideration when forming levels. This does not necessarily indicate that the other aspects are completely irrelevant, but that those have to be deduced back by item content of the respective resulting stages/levels (indicated not by our own interpretation but by classification of the experts).

**Table 1** Sample description ( $N=877$ )

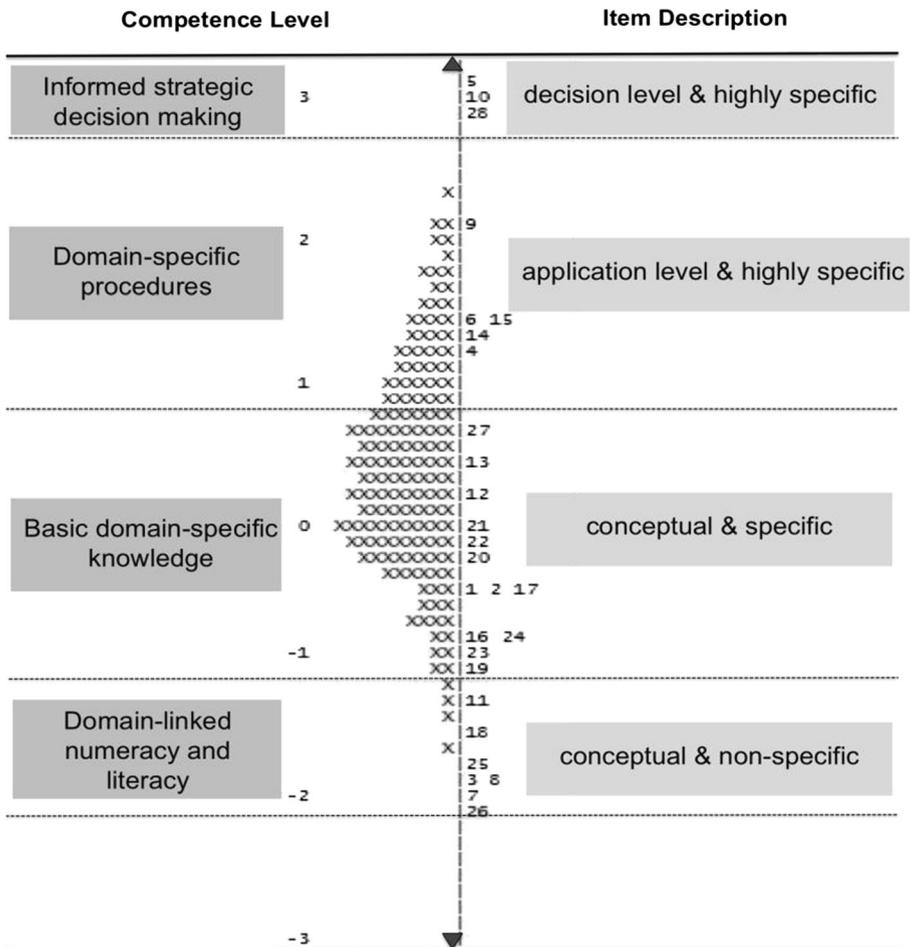
Sample Characteristic	Group 1 ( $n_1=398$ )	Group 2 ( $n_2=479$ )
Average years of initial VET	Ø 0.3 years (= beginners)	Ø 2.0 years (= advanced)
Age	Ø 19.5 years old	Ø 21.1 years old
Gender	Female: 56 %	Female: 59 %
Migration background	20 %	18 %
Educational career	Secondary modern school: 1 % Junior high school: 25 % Advanced technical college: 22 % High school business diploma: 6 % General high school diploma: 46 %	Secondary modern school: 1 % Junior high school: 25 % Advanced technical college: 25 % High school business diploma: 9 % General high school diploma: 40 %
Company size (number of employees)	< 10: 2 % 10–49: 13 % 50–249: 35 % ≥ 250: 18 % ≥ 500: 32 %	< 10: 1 % 10–49: 10 % 50–249: 34 % ≥ 250: 20 % ≥ 500: 35 %

an item if it contains no difficulty-generating characteristics. The next threshold results from adding the next difficulty characteristic. If an item is of moderate specificity but indicates no cognitive processing difficulty, the threshold for this level of competence would be  $-2.012 + 0.837 = -1.175$ . Adding the unstandardized regression coefficient for a high degree of specificity yields a threshold value of 0.893. When all characteristics apply (i.e., the item is highly specific and requires interpretation), the generated threshold is 2.767, which represents the maximum level of vocational competence, namely, domain-specific rule understanding combined with the cognitive ability to apply interpretative competence to make grounded, strategic decisions. Figure 2 shows all four generated competence levels. It is noteworthy that all the items fall into the categorizations predicted by the expert ratings, with the exception of item 11, for which the experts predicted a moderate specific level.

Through inferences from the item characteristics to the respective competence levels, we propose that the first level can be described as domain-linked numeracy and literacy. Learners at this competence level are new to the domain and have only non-specific knowledge. This does not mean that domain-linked knowledge is generally easier. However the domain-linked contents described in the vocational curriculum for industrial apprentices and for a lot of other economic apprenticeships span relatively

**Table 2** Regression model of item difficulty for advanced vocational learners (group 2)

Adjusted $R^2=0.768$	$\beta$ (unstd.)	$\sigma$	$\beta$ (std.)	T	$p$
Intercept	-2.012	0.347		-5.798	0.000
Specificity (2=moderate specificity)	0.837	0.426	0.252	1.965	0.061
Specificity (3=high specificity)	2.068	0.409	0.666	5.053	0.000
Cognitive processing (3=interpretation)	1.874	0.334	0.564	5.605	0.000



**Fig. 2** Difficulty model for advanced vocational learners (group 2) (Note: Each X represents 3 learners; WLE-Reliability=0.847; EAP/PV Reliability=0.834)

restricted content and are located on a low competence level compared with PISA standards (see a curricular analysis by Winther et al. 2013). Furthermore this level of competence comprises knowledge and ability that learners should already have acquired at their general school, which makes it almost natural that this is the easiest level for learners. The second level, basic domain-specific knowledge, implies that learners have acquired a moderate level of domain-specific knowledge and internalized the basic rules of a community of practice. We refer to the third level as domain-specific procedures. Learners who reach this level can apply highly domain-specific knowledge in procedural situational settings. Finally, on the fourth level of informed strategic decision making, learners can apply highly specific rule knowledge in complex decision-making processes that require a sound interpretation of a vocational situation, as well as the applied rules and heuristics. This model serves as a blueprint to describe how advanced learners spread over the four categories and how difficult it is for them to surpass a certain threshold and reach a higher qualitative level of competence. In our

sample of advanced vocational learners, almost all of them reached at least basic domain-specific knowledge, and many are at the domain-specific procedures level. However, only five persons in the data set reached the level of informed strategic decision making, prompting us to speculate that this level is a curricular ideal (e.g., AKA 2011), rarely reached in practice by vocational learners by the end of their VET.

By implication, the sequences obtained from this blueprint model should represent stages through which learning progresses (Wilson 2009). To model a learning progression, we applied the same thresholds to vocational beginners, just starting their initial vocational training. With this approach, we can infer that the observed changes in item difficulty between the two groups result from changes in their vocational competence. By regressing the expert judgments of specificity and cognitive processing on item difficulty, we obtained the model for group 1 in Table 3.

Comparing the effect sizes as relative distances on the logit scale between the regression models for beginning and advanced vocational learners reveals that for the vocational beginners, domain-specific items represent an even greater source of difficulty; these learners appear to lack domain-specific competence at the beginning of their VET. To identify any absolute drops in difficulty—or, following the idea of a learning progression, absolute increases in competence—we next compared the blueprint model of advanced vocational learners (group 2) as a reference point with the threshold locations on the logit scale for group 1 (see Table 4).

According to Paek (2002), absolute differences on the logit scale less than 0.426 are negligible. Differences up to 0.638 indicate medium-sized effects, and those higher than this level indicate a strong effect size for a learning progression. The first threshold, indicating domain-linked numeracy and literacy competence, shifted down by 0.989 in group 2, so it indicated a strong learning progression. This interesting finding suggests that during the VET, learners not only acquire new domain-specific knowledge but also increase their general abilities, such as numeracy and literacy. It also challenges the notion that vocational competence progression consists solely of cognitive movement from general to specific knowledge. Instead, it appears that vocational, domain-specific understanding stimulates the acquisition of general abilities, possibly due to problem-based learning, as suggested by management learning research pertaining to business and management (e.g., Kolb and Kolb 2009). The second and third thresholds shifted by 1.170 and 1.163, respectively, revealing even larger effects for the increase of domain-specific competence and specific analytic skills during VET. For the final, informed strategic decision-making level, we observed only a small, non-significant effect (0.129), so learners did not appear to improve significantly with respect to that competence level. This finding is not surprising, considering that only five learners in advanced group

**Table 3** Regression model of item difficulty for vocational beginners (group 1)

Adjusted R <sup>2</sup> =0.733	Beta (unstd.)	$\sigma$	Beta (std.)	T	<i>p</i>
Intercept	-1.023	0.280		-3.653	0.002
Specificity (2=moderate specificity)	1.018	0.357	0.423	2.851	0.011
Specificity (3=high specificity)	2.061	0.415	0.837	4.971	0.000
Cognitive processing (3=interpretation)	0.840	0.338	0.330	2.485	0.023

**Table 4** Competence level threshold movement, as an indication of learning progression

Competence level	Group 1 (beginners)	Group 2 (advanced)	Learning Progression	Effect Size
1. Domain-linked numeracy and literacy	-1.023	-2.012	0.989	high
2. Basic domain-specific knowledge	-0.005	-1.175	1.170	high
3. Domain-specific procedures	2.056	0.893	1.163	high
4. Informed strategic decision making	2.896	2.767	0.129	–

2 reached this level. A comparison at this early stage of lifelong vocational development may not be appropriate yet, because this level of competence likely arises at a later stage (i.e., expert stage in Dreyfus and Dreyfus's [1980] model), when an expert gains experience and thus can control complex mechanisms in more situation-specific and intuitive ways. In Fig. 3 we depict the results from Table 4 regarding how learners develop their vocational competence. The absolute drops in the level of difficulty from group 1 to group 2 (grey to black thresholds) reflect the learning progression over the initial vocational trajectory (black to grey thresholds).

This model allows us to account for the total increase, as well as interpret the movement with respect to the different qualities of vocational competence. The results of the regression models and the threshold movement together suggest that vocational learners mostly increase their proficiency in domain-specific competence and, to a lesser extent, their capacity for cognitive processing. The model also represents a criterion-based reference of the absolute quality of competence that learners develop during VET. That is, vocational learners start with domain-linked cognitive bases of numeracy and literacy (group 1), which get strengthened during their VET and expanded as well as integrated with domain-specific elements. They can link this newly acquired, specific knowledge with their cognitive processing capacities and thereby perform specific economic tasks on not only conceptual but also procedural levels (group 2).

## Theoretical and Educational Significance

Vocational development on an aggregate level in the economic domain largely depends on a combination of two aspects: the acquisition of a set of principles specific to the vocational community ( $\beta=0.918$ ) and demand for cognitive processing to complete vocational activities ( $\beta=0.564$ ) (see Gelman and Greeno 1989). We moreover conclude from our results that learning about specific vocational cases also can contribute significantly to abstract knowledge (i.e., threshold shifts of 0.989 for domain-linked numeracy and literacy), beyond the widely accepted Piagetian view that proficiency increases as a learner moves from the concrete to the abstract.

Regarding its educational significance, our developmental model might be applied as a grounded framework for criterion-based assessment in economic domains (e.g., for occupational licensing). With this model, learners could be assessed in terms of their quantitative proficiency and with regard to a desired competence standard. Therefore, our model constitutes a criterion-based reference of the absolute quality of competence developed

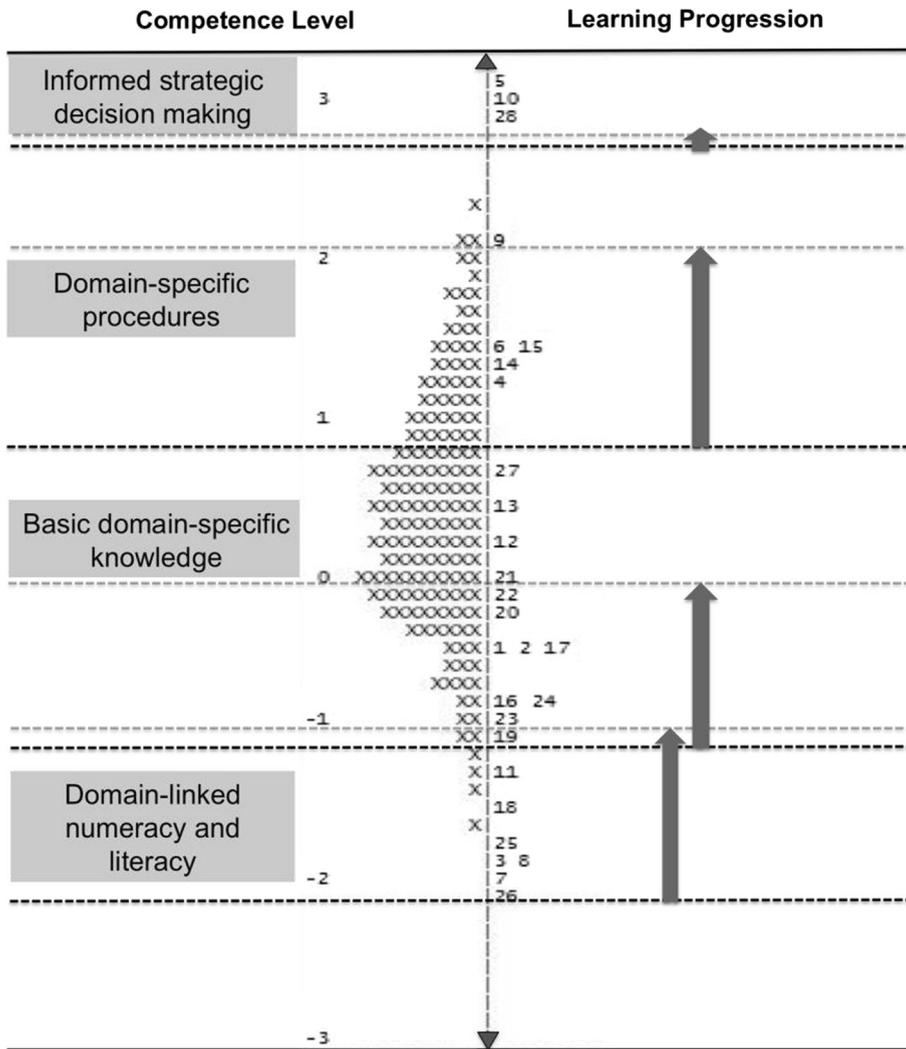


Fig. 3 Learning progression

during VET and can be used to evaluate learning results. The training implications of this developmental description also are straightforward: The different stages constitute a systematic, successive vocational transformation of cognitive structures. The designers of VET programs and vocational teachers could use the respective descriptions and examples for each stage (tasks) to facilitate learners' advancement to higher stages and avoid leaving them unchallenged or, alternatively, introducing later stage tasks prematurely, which might impede further advancement (e.g., Hacker 1986, 2003).

However, the theoretical and educational significance of our model is limited by several shortcomings of our research design. Generally speaking, we acknowledge that the average learning path we model reflects current curriculum practices or the status quo, and it might change if workplace and school curricula were to change. Furthermore, though the total sample size for our study is substantial for testing group

hypotheses, it was not drawn randomly. Moreover, the cross-sectional nature of our design means that we cannot examine individual variations with respect to potentially distinct development paths. The obvious recognition of workplace complexities and individual vocational trajectories cautions against accepting the notion of a “typical” learning progression implied by an aggregate learning path as presented here. Rather, many factors affect students’ learning, and variability across individuals is inevitable.

For vocational competence, this article thus portrays one perspective among the many possible ways that vocational development could be conceived of and captured. With respect to our applied IRT-based method of modeling learning progressions, we believe it is a promising new approach and way forward for vocational educational science, in that it can describe learning outcomes while also clarifying vocational learning processes. Starting with an average path can help define developmental “differences” and explicate the causes of variability in longitudinal designs, which in turn could yield important insights into vocational learning. A new systematic field of research, to advance understanding of vocational learning (e.g., roles of workplace and vocational school characteristics, curriculum and instruction that account for individual differences), thus opens for further investigation.

**Acknowledgments** This research was supported by the German Research Foundation (DFG), within the subproject “Competence-oriented assessments in VET and professional development” (Wi 3597/1–2), in the larger framework of the priority program “Competence Models for Assessing Individual Learning Outcomes and Evaluating Educational Processes” (SPP 1293).

**Conflict of Interest** The authors declare that they have no conflict of interest.

## Appendix 1

# *Ceraforma Keramik AG*



Since its foundation in 1982, the Ceraforma Keramik AG has developed into an expanding and globally active industrial enterprise having their head office in Aachen, Germany. The company is involved in the production of ceramic goods, such as china and porcelain for tableware and vases or sanitary ware.

In the past, the management of Ceraforma Keramik realized that the four divisions—procurement logistics, production, human resource management as well as marketing and sales—used to operate too independently from each other, which caused disturbances in the performance process and led to customer complaints. In response to these problems, so-called *horizontal teams* were established consisting of work members from different company divisions.

You have been employed with Ceraforma Keramik in such a horizontal team since the beginning of this year. Here the allocated customer orders are being handled in all business processes ranging from the receipt of orders to the settlement of accounts. Ms Kenk, the team leader, Mr Friebel and Ms Hoffmann, the new trainee, are your colleagues in the horizontal team.



## Business Process 1

### **Situation:**

Your team just received a new customer enquiry. Your colleague, Mr Friebel, shows you the following e-mail which arrived on 30 March 20... at 10:17.

5 After repeated negotiations the company Ceraforma accepts the order from the DIY Bauhannes at the price stipulated by Mr Schwienert. Receipt of confirmation of the order by email is on Friday, 6 April 20... You have been informed that there is no sufficient quantity of quartz crystal on stock to execute the order. You are therefore required to order 25 t of new quartz crystals. You then contacted various suppliers by mail and you received the

emails below from Mineral Seifert AG from Aachen and Tam-Quarz Ltd. from South Africa:

	<b>An:</b>	horizontalteam3@ceraforma.de
	<b>Kopie:</b>	
	<b>Blindkopie:</b>	
	<b>Betreff:</b>	Unser Angebot für Sie

Lieber Herr Friebe,

wir freuen uns, dass wir Sie wieder einmal von unseren Produkten und Leistungen überzeugen konnten.

Aufgrund unser langjährigen Geschäftsbeziehungen, können wir Ihnen zu Ihrer Anfrage folgende Konditionen anbieten:

Produkt: reiner Quarz, Bergkristall, weiß  
 Preis/Menge: 500,00 EUR/t inkl. MwSt.  
 Zahlungsbedingungen: 10 Tage 3 % Skonto; 60 Tage netto Kasse  
 Bezugskosten: 100,00 EUR pauschal/Lieferung  
 Lieferzeit: 3 Werktage ab Bestelleingang  
 Angebot ist gültig: bis zum 15.04.20..

Wir würde uns freuen, wenn Sie sich erneut für unsere Produkte und unseren Service entscheiden würden.

Einen schönen Tag noch und freundliche Grüße

Jörg Schewe  
 Vertrieb  
 Mineral Geifert AG Aachen

	<b>An:</b>	horizontalteam3@ceraforma.de
	<b>Kopie:</b>	
	<b>Blindkopie:</b>	
	<b>Betreff:</b>	RE: Angebotsanfrage

Dear Sir/ Madam,

In reply to your enquiry dated xy.20.. we are pleased to make the following offer:

- white mountain quartz crystal: € 450.00/ ton
- minimum order quantity: 10 tons
- shipping charges: € 13.00/ 100kg

Since this is your first order, we allow a quantity discount of 3% per ton for orders exceeding 30 tons.

This offer is valid until April 15. 20.. Please consider that transportation by ship might takes up to a month.

We look forward to hearing from you soon!

Sincerely, (alternative: Yours sincerely/ Kind regards)

J. Stones  
 Tam-Quarz, South Africa  
 Mail: stones@tam.za





- Dreyfus, S. E., & Dreyfus, H. L. (1980). *A five-stage model of the mental activities involved in directed skill acquisition*. Berkeley: University of California, Berkeley.
- Duncan, R. G., Rogat, A. D., & Yarden, A. (2009). A learning progression for deepening students' understandings of modern genetics across the 5th-10th grades. *Journal of Research in Science Teaching*, 46(6), 655–674.
- Gelman, R., & Greeno, J. G. (1989). On the nature of competence: Principles for understanding in a domain. In L. B. Resnick (Ed.), *Knowing and learning: Essays in honor of Robert Glaser* (pp. 125–186). Hillsdale: Erlbaum Associates.
- Glaser, R. (1990). Re-emergence of learning theory within instructional research. *American Psychologist*, 45(1), 29–39.
- Greeno, J. G., Riley, M. S., & Gelman, R. (1984). Conceptual competence and children's counting. *Cognitive Psychology*, 16, 94–143.
- Gschwendtner, T. (2011). Die Ausbildung zum Kraftfahrzeugmechatroniker im Längsschnitt. Analysen zur Struktur von Fachkompetenz am Ende der Ausbildung und Erklärung von Fachkompetenzentwicklungen über die Ausbildungszeit. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 25, 55–76.
- Hacker, W. (1986). Arbeitspsychologie—Psychische Regulation von Arbeitstätigkeiten. In E. Ulich (Ed.), *Schriften zur Psychologie*. Bern: Huber.
- Hacker, W. (2003). Action Regulation Theory: A practical tool for the design of modern work processes? *European Journal of Work and Organizational Psychology*, 12(2), 105–130.
- Hartig, J. (2007). Skalierung und Definition von Kompetenzniveaus. In B. Beck & E. Klieme (Eds.), *Sprachliche Kompetenzen. Konzepte und Messung* (pp. 83–99). Beltz: Weinheim.
- Hartig, H., Frey, A., Nold, G., & Klieme, E. (2012). An application of explanatory item response modeling for model-based proficiency scaling. *Educational and Psychological Measurement*, 72(4), 665–686.
- Hoffman, L., Hofer, S. M., & Sliwinski, M. J. (2011). On the confounds among retest gains and age-cohort differences in the estimation of within-person change in longitudinal studies: A simulation study. *Psychology and Aging*, 26, 778–791.
- Johnson, P., & Tymms, P. (2011). The emergence of a learning progression in middle school chemistry. *Journal of Research in Science Teaching*, 48(8), 849–877.
- Klotz, V. K. (2015). *Diagnostik beruflicher Kompetenzentwicklung: Eine wirtschaftsdidaktische Modellierung für die kaufmännische Domäne*. Wiesbaden: Springer Gabler.
- Klotz, V. K., & Winther, E. (2015). Kaufmännische Kompetenz im Ausbildungsverlauf—Befunde einer pseudolängsschnittlichen Studie. *Empirische Pädagogik*, 29(1), 61–83.
- Kolb, A. Y., & Kolb, D. A. (2009). Experiential learning theory: A dynamic, holistic approach to management education and development. In S. J. Armstrong & C. V. Fukami (Eds.), *Sage handbook of management learning, education and development* (pp. 42–68). London: Sage.
- Lave, J. (1990). The culture of acquisition and the practice of understanding. In J. W. Stigler, R. A. Shweder, & G. Herdt (Eds.), *Cultural Psychology* (pp. 259–286). Cambridge: Cambridge University Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Menard, S. (Ed.). (2002). *Longitudinal research*. Thousand Oaks: Sage.
- Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement. Issues and Practice*, 25, 6–20.
- Mulder, M., Weigel, T., & Collins, K. (2006). The concept of competence concept in the development of vocational education and training in selected EU member states. A critical analysis. *Journal of Vocational Education and Training*, 59(1), 65–85.
- Nickolaus, R. (2011). Die Erfassung fachlicher Kompetenz und ihrer Entwicklungen in der beruflichen Bildung—Forschungsstand und Perspektiven. In O. Zlatkin-Troitschanskaia (Ed.), *Stationen empirischer Bildungsforschung: Traditionslinien und Perspektiven* (pp. 331–351). Wiesbaden: Springer.
- Nickolaus, R., Gschwendtner, T., & Geißel, B. (2008). Modellierung und Entwicklung beruflicher Fachkompetenz in der gewerblich-technischen Erstausbildung. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 104(1), 48–73.
- Oates, T. (2004). The Role of outcome-based national qualifications in the development of an effective vocational education and training system: The case of England and Wales. *Policy Futures in Education*, 2(1), 53–71.

- OECD. (2003). *The PISA 2003 assessment framework—mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD.
- Paek, I. (2002). *Investigation of differential item functioning: Comparisons among approaches, and extension to a multidimensional context*. Berkley: University of California, Berkeley.
- Pellegrino, J. W. (2012). The design of an assessment system focused on student achievement. A learning sciences perspective on issues of competence, growth and measurement. In S. Bernholt, K. Neumann, & P. Nentwig (Eds.), *Making it tangible—Learning outcomes in science education* (pp. 79–107). Waxmann: Münster.
- Pellegrino, J., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington: National Academy Press.
- Phillips, D. (1995). The good, the bad, and the ugly: The many faces of constructivism. *Educational Researcher*, 24(7), 5–12.
- Plummer, J. D., & Krajcik, J. (2010). Building a learning progression for celestial motion: Elementary levels from an earth-based perspective. *Journal of Research in Science Teaching*, 47(7), 768–787.
- Rasch, G. (1977). On specific objectivity: An attempt at formalizing the request for generality and validity of scientific statements. *Danish Yearbook of Philosophy*, 14, 58–93.
- Ryan, M., Fook, J., & Hawkins, L. (1995). From beginner to graduate social worker: Preliminary findings of an Australian longitudinal study. *British Journal of Social Work*, 25, 17–35.
- Rychen, D. S., & Salganik, L. H. (2003). *Key competencies for a successful life and a well-functioning society*. Göttingen: Hogrefe & Huber Publishers.
- Salthouse, T. A., & Tucker-Drob, E. M. (2008). Implications of short-term retest effects for the interpretation of longitudinal change. *Neuropsychology*, 22, 800–811.
- Seeber, S. (2008). Ansätze zur Modellierung beruflicher Fachkompetenz in kaufmännischen Ausbildungsberufen. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 104(1), 74–97.
- Shavelson, R. J. (2008). Reflections on quantitative reasoning: An assessment perspective. In B. L. Madison & L. A. Steen (Eds.), *Calculation vs. context: Quantitative literacy and its implications for teacher education* (pp. 27–47). New York: Mathematical Association of America.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428.
- Steedle, J. T., & Shavelson, R. J. (2009). Supporting valid interpretations of learning progression level diagnoses. *Journal of Research in Science Teaching*, 46(6), 699–715.
- Taylor, A. (2013). Learning progressions: An overview of current validation methods. *Pearson Assessment Bulletin*, 24, 1–4.
- Wilson, M. (2005). *Constructing measures: An item-response modeling approach*. Mahwah: Erlbaum.
- Wilson, M. (2008). Cognitive diagnosis using item response models. *Journal of Psychology*, 216(2), 74–88.
- Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching*, 46(6), 716–730.
- Winther, E., & Achtenhagen, F. (2009). Measurement of vocational competencies—A contribution to an international large-scale assessment on vocational education and training. *Empirical Research in Vocational Education and Training*, 1, 88–106.
- Winther, E., Sangmeister, J., & Schade, A. (2013). Zusammenhänge zwischen allgemeinen und beruflichen Kompetenzen in der kaufmännischen Erstausbildung. In R. Nickolaus, J. Retelsdorf, E. Winther & O. Köller, *Zeitschrift für Berufs- und Wirtschaftspädagogik* 26, 139–160.
- Worthy, C. (1996). Clinical ladders: Can we afford them? *Nursing Management*, 27, 33–34.
- Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: MESA Press.
- Wu, M. L., Adams, R. J., & Wilson, M. R. (1997). *ConQuest: Multi-aspect test software*. Camberwell: Australian Council for Educational Research.
- Zhang, J., & Stout, W. F. (1999). Conditional covariance structure of generalized compensatory multidimensional items. *Psychometrika*, 64, 129–152.

**Viola Katharina Klotz** 2006–2009 Bachelor in Economics at the Goethe-University Frankfurt. 2007–2009 Work at the KfW-Developmentbank. 2009–2011 Master in Educational Economics at the University of Paderborn. Since December 2011 PhD student at the University of Paderborn, Department for Educational Economics. Since January 2015 Research Assistant at the German Institute for Adult Education in Bonn.

**Esther Winther** 1999–2003 Diploma in Educational Economics at the Georg-August University, Göttingen 2005. Dissertation at the University of Göttingen. 2008–2009 Researcher at Stanford University 2009 Assistant Professor at the University of California Berkeley. 2010–2014 Professor at the University of Paderborn. Since October 2015 Scientific director at the German Institute for Adult Education in Bonn.

---

**Dagmar Festner** 1995–2002 Educational Studies at the University of Regensburg. 2007–2009 Research Assistant at the University of Regensburg. 2009–2011 Research Assistant at the FBB in Nuremberg 2012. Dissertation at the University of Regensburg. 2011–2014 Research Assistant at the University of Paderborn. Since 2015 Research Assistant at the PLAZ in Paderborn (Center for Educational Research and Teacher Education)