



Article

# Measurement of Financial Competence—Designing a Complex Framework Model for a Complex Assessment Instrument

Andreas Kraitzek \* and Manuel Förster

School of Social Sciences and Technology, Technical University of Munich, 80333 München, Germany

\* Correspondence: andreas.kraitzek@tum.de; Tel.: +49-89-289-23211

**Abstract:** Financial competence is seen as a complex ability necessary for people to deal with personal financial issues on a daily basis. To foster young peoples' financial competence via sophisticated and tailored educational programs, the identification of "competence gaps" through complex and authentic assessments is required. While a large number of assessment tools in the field of personal finance already exist, many of them suffer from different shortcomings concerning a competence-oriented approach. Therefore, we present an innovative way to assess students' financial competence with a complex performance scenario about financial investment. The presented instrument is built on a specifically designed theoretical framework and addresses the need for holistic financial competence measurement. Results of pretesting trials indicate that the instrument is generally capable of measuring young learners' financial competence, but challenges in scoring remain. Against this background, implications for the instrument's iterative enhancement are presented and discussed with reference to validity and reliability properties, scoring issues, and statements about the overall feasibility of complex performance tasks in educational settings. The first draft of a scoring scheme is provided. The potential of the instrument in combination with modern technology-based measurement methods (eye tracking, emotion recognition) for competence assessment is described and suggestions for further research are outlined.

**Keywords:** financial literacy; financial competence; framework; competence model; problem-solving; assessment; complexity; performance scenario; test development



**Citation:** Kraitzek, Andreas, and Manuel Förster. 2023. Measurement of Financial Competence—Designing a Complex Framework Model for a Complex Assessment Instrument.

*Journal of Risk and Financial Management* 16: 223. <https://doi.org/10.3390/jrfm16040223>

Academic Editor: Thanasis Stengos

Received: 31 January 2023

Revised: 13 March 2023

Accepted: 14 March 2023

Published: 3 April 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Financial literacy (FL) is considered "a lifetime skill that everybody needs to have to be able to live and operate in today's complex economic environment" (Lusardi 2012, p. 10). Not only are economic and financial issues, products, and situations becoming increasingly complex and intertwined, but FL itself is seen as a complex, multi-dimensional construct that is the basic prerequisite for an individual's financial competence (Lusardi 2019), which by itself is also considered to contain a complex multi-faceted structure. There is broad consensus that FL and, subsequently, financial competence consists of cognitive and non-cognitive facets, as stated in Weinert's (2001b) widely recognized definition of competence. Numerous tests and instruments for the assessment of FL already exist (for overviews, see, e.g., Fürstenau et al. (2020); Ouachani et al. (2021); Rieger (2020); Stolper and Walter (2017)) and such instruments can surely act as a potential starting point for the investigation of the construct. However, concerning a holistic perspective on the construct of financial competence, many of these instruments and approaches are insufficient (Breuer 2016; Kaiser and Lutter 2015). They either focus on standardized tests for mere knowledge assessment (i.e., the cognitive side of the construct), or, if these approaches ever consider non-cognitive aspects (e.g., motivation, interest, attitudes, individual risk preferences, etc.), they are also operationalized with standardized questionnaires. Afterward, correlations between the constructs are calculated, from which claims about the levels of literacy are derived. To a certain extent, this is generally compatible with the understanding of (financial)

competence as a multi-faceted construct. However, in doing so, these approaches and instruments neither take situational aspects nor the complexity into account that both the domain of economics and personal finance, as well as the constructs of financial literacy and financial competence, require. In other words, these approaches do not model the complex interlinkages between construct, individual cognitive and non-cognitive dispositions, domain-specific situational aspects, and processes sufficiently. A competence-oriented approach to the assessment of personal financial literacy should thus integrate various aspects and principles that directly address the existing shortcomings described above (Aprea et al. 2015; Fürstenau et al. 2020):

First, concerning the individual, learning and assessment should not be limited to cognitive dispositions but rather consider the simultaneous investigation of the cognitive and affective-emotional side while dealing with the construct. Handling financially relevant situations is not only influenced by an individual's non-cognitive dispositions, it can also trigger emotional states and eventually affect (financial) well-being.

Second, the complexity of financially relevant situations and of the construct of competence itself has to be considered. Measuring and fostering various subdimensions of financial literacy separately and independently ignores the close interconnections and interdependencies between the necessary prerequisites for competent behavior and also falls short of the large number of variables that define financially relevant situations.

Third, the assessment of financial competence should also shed light on the actual process of handling financially relevant situations and not just focus on the results. It is necessary to investigate how and which cognitive and non-cognitive dispositions become relevant while solving financial problems and when transferring the solutions into observable behavior. As financial education is considered to be an essential part of young peoples' preparation for life, detailed insights into these issues will help to create more sophisticated educational programs along with specific and customized didactical concepts in order to foster financial competence systematically and more efficiently.

In this contribution, we present an innovative competence-oriented approach to determine financial literacy with a complex assessment scenario based on an elaborated model of financial competence and problem-solving. This complex tool with which we try to test students and young adults goes beyond the test procedures used so far. We try to implicitly assess cognitive and non-cognitive facets of financial competence and problem-solving simultaneously. Based on that model specifically designed for this purpose, this scenario has the potential to test an individual's level of financial competence while at the same time including domain-specific cognitive and non-cognitive dispositions, situational complexity, and the process perspective in one single assessment instrument. We aim to answer the question of how to conceptualize a complex performance instrument to meet the requirements for the assessment of the complex construct of financial competence effectively and validly. By following a systematic and structured process for the development of an alternative form of competence measurement, we also aim to make a substantiated contribution to theory formation in educational assessment for the field of personal finance.

Our contribution is thus structured as follows. Sections 2 and 3 outline the necessity for a new approach to modeling the complex construct of financial competence and describe the development of a holistic theoretical framework model for the assessment of financial competence. Section 4 illustrates an example of a complex performance scenario that was specifically designed upon this framework to assess financial competence in a holistic and innovative way. Sections 5 and 6 provide first insights into the pretesting phase of the scenario and discuss the implications of the initial results for education and test development. Furthermore, the limitations of such assessment instruments are outlined, and ideas for further research efforts are presented.

## 2. Theoretical and Conceptual Background

To establish a theoretical basis, this first section defines the relevant constructs for the subsequent discussion.

### 2.1. The Relationship between Economics and Financial Education

Educational efforts in economics and personal finance aim to foster peoples' ability to act as responsible 'economic citizens' (Billimoria et al. 2013; Davies 2006). This means that people should be able to show competent behavior in the face of economic and financially relevant situations. In general, the broad economic domain can be divided into a business perspective (for vocational matters) and a civic perspective, which includes mostly matters and content that are relevant for people as private consumers and income earners on a personal day-to-day basis (Dubs 2014; Eberle et al. 2016); for an in-depth discussion about the civic-economic domain, see also: Ackermann (2019). Considering such content, the economics domain can be split up into general economic topics (e.g., inflation, markets, prices) and specific financial topics (i.e., income, saving, investing, financial planning). In this instance, economic education is seen to be more comprehensive, whereas financial education is considered to be an integrated part of the wider area of economics (Retzmann and Seeber 2016; Wuttke et al. 2016). In the course of this contribution, it will be shown that our instrument aims to assess an individual's competence in the area of personal finance. From the viewpoint of financial education, assessment is an important part of learning as it allows the identification of existing educational gaps that can subsequently be addressed by more effective financial education interventions (Bongini et al. 2018).

### 2.2. Categorizing the Construct of Financial Competence for the Assessment Approach

Financial education (i.e., the process) should lead to peoples' financial literacy (i.e., the state or condition of being financially educated) and eventually to financial competence (i.e., proper behavior as a result of being financially literate, (Reifner 2011)). The terms financial literacy (FL) and financial competence (FC) by themselves are the subject of ongoing debate as there is a plethora of definitions and conceptual approaches. A definitive and exhaustive discussion about what the relevant constructs of FL or FC are or what they are not cannot be part of this contribution (for overviews, see, e.g., Cude (2022); Goyal and Kumar (2020); for discussions, see, e.g., Geiger et al. (2016); Seeber and Retzmann (2017)). Instead, we focus on the main aspects of the construct that are relevant to our approach to complex financial competence assessment.

First, FL is considered a *result of educational efforts* to enable individuals to act as their own personal financial manager (Aprea et al. 2015). This understanding is currently seen as the most circulated concept of FL, and it is also in line with the approach of the OECD's FL-PISA (Program for International Student Assessment) (Kraitzek et al. 2022). Basically, this understanding is oriented towards personal financial decision-making for people as subjects in their daily roles of consumers, earners, investors, and economic citizens. Individuals are entangled in complex economic and financially relevant situations every day, and decisions about what to buy, how much to spend, where to invest, and how much to save have to be made on a regular basis. Due to the principle of scarcity, personal financial resources are not limitless, and individuals need the ability to solve intricate financially relevant problems and situations on a daily basis. This comprehension can be further expanded to the understanding of FL as critical consumer education, which includes critical reflection about one's personal material needs, the control of impulsive buying behavior, or an awareness about an individual's risk priorities in financial decisions (Aprea et al. 2015).

Second, FL is considered a *multi-dimensional construct*. There is broad consensus that FL consists of various dispositions that can be categorized into a cognitive, a non-cognitive, and an application dimension (Fürstenau et al. 2020). One prior attempt at conceptualizing FL was undertaken by Huston (2010), as she originally segmented FL into a knowledge dimension and an application dimension. According to this idea, "financial knowledge is an integral dimension of, but not equivalent to, financial literacy" (Huston 2010, p. 307). Although this is in line with the idea that domain knowledge—i.e., the cognitive dimension—is a necessary prerequisite for mastering domain-specific problems, issues, or situations (Brückner et al. 2015; Weinert 2001b), considering the cognitive disposition

alone may be too short-sighted. Huston further claims that “FL has an additional application dimension which implies that an individual must have the ability and confidence to use his/her financial knowledge to make financial decisions” (Huston 2010, p. 307). In essence, this means that the application of knowledge in a behavioral sense is crucial, and that this usage of knowledge is influenced, among other things, by confidence—or, to put it more broadly, by an individual’s non-cognitive dispositions. Building on this theoretical concept, a vast number of definitions for the construct of FL have evolved, with the most famous one probably being the internationally renowned definition of the OECD for its FL-PISA assessment: “Financial literacy is knowledge and understanding of financial concepts and risks, and the skills, motivation, and confidence to apply such knowledge and understanding in order to make effective decisions across a range of financial contexts, to improve the financial well-being of individuals and society, and to enable participation in economic life” (OECD 2019, p. 128). This definition clearly states that a financially literate person should be able to show competent and sophisticated behavior in the face of financially relevant situations, issues, challenges, or problems.

This understanding is generally in line with our notion of financial competence, which acts as the underlying construct for our complex assessment scenario. The term *competence* traces back to Weinert, who defined it as “the cognitive abilities and skills available to or learnable by individuals to solve specific problems, as well as the associated motivational, volitional, social skills and willingness to use problem-solving successfully and responsibly in variable situations” (Weinert 2014, pp. 27–28, translated by the authors). Consequently, competence—such as literacy—implies complex action systems that integrate cognitive, motivational, social, and behavioral components rather than solely basic cognitive abilities or simple skills (Weinert 2001a). As such, we consider both constructs mostly congruent; however, with the slightly distinctive feature that financial literacy is seen as the general state of being financially educated, whereas financial competence implies the actual implementation process, namely the concrete execution of sound financial behavior (Reifner 2011). We understand financial competence as the sum of (domain-specific) financial knowledge and skills, abilities, motivation, attitudes, and the confidence to apply and execute such knowledge and skills in order to deal with financially relevant situations or issues appropriately.

Financial competence, such as financial literacy, is a latent construct that shows itself in observable performance (Erpenbeck et al. 2017). Consequently, as mentioned by Huston (2010), an instrument to measure financial literacy—or in our case, financial competence—has to determine not only if people know the information but also if they can apply it properly in the holistic sense of our understanding of the construct. As such, an instrument for the assessment of financial competence should not just measure financial knowledge, but it also must be able to depict a person’s domain-specific characteristics concerning, for example, the general motivation to deal with financial questions, the manifestation of people’s risk priorities or financial attitudes toward consumption, investing, or saving. Furthermore, the instrument must lead to the production of a “solution”, an action of some sort in the sense of an observable output, as a result of the application dimension from which conclusions about the aforementioned characteristics can be drawn.

Following this concept, it becomes clear that financial competence must be modeled as a complex construct and that this complexity has to be represented in the theoretical framework upon which instruments for competence assessment are built. Therefore, the next sections describe the development of such a theoretical framework that integrates the multidimensionality of the financial competence construct as the basis for an assessment.

### 2.3. Prior Approaches to Competence Assessment

Existing approaches to the assessment of competence can be divided into four categories (Sembill et al. 2013):

(1) The first approach, i.e., modeling and measuring competence as a unidimensional construct, is outdated and antique and represents an understanding of the construct which is clearly not complex enough.

(2) The second approach models competence as a fragmented construct without taking non-cognitive aspects into consideration. With this approach, the priority lies in the assessment of domain-specific knowledge since knowledge is often seen as the basic prerequisite for domain-specific competence. In practice, this means that knowledge and skills are often assessed in multiple-choice formats as isolated abilities and sub-abilities across different knowledge facets in a rather analytical way (Zlatkin-Troitschanskaia et al. 2019).

(3) The third approach considers competence as a multifaceted construct but with a separate measurement of cognitive and non-cognitive dispositions (Sembill et al. 2013). In this case, non-cognitive facets are measured alongside cognitive dispositions, for example, via standardized self-report questionnaires, and correlations between the facets are calculated. Modern approaches to FL assessments that fall into this category are, for example, the studies of Schürkmann (2017) (Financial Literacy Study “FILS”), Rudeloff (2019), Grohs-Müller (2020) or the FL-section of PISA (OECD 2013). Occasionally, also hybrid approaches, where standardized questionnaires are combined with situational prompts, are implemented (e.g., Financial Literacy Study of Adults “FILSA” (Schuhen et al. 2022)). While this approach is generally in line with our understanding of competence as a multi-dimensional construct, it clearly does not take the complex interconnections between cognitive and non-cognitive dispositions and the situational specificity into account as we outlined above.

(4) The fourth and yet rarely applied approach models competence as a multifaceted construct and takes on an integrated measure of cognitive and non-cognitive dispositions (Sembill et al. 2013). In other words, both cognitive and non-cognitive aspects are measured simultaneously with a single, specifically developed instrument or with a set of coordinated instruments. More generally, such a holistic, criterion-sampling approach recognizes the theoretically anticipated complexity of a construct and tries to examine it “as a whole” via an observable task performance in a close-to-real-life criterion situation (Zlatkin-Troitschanskaia et al. 2019). The idea of this fourth approach is the foundation on which the theoretical framework will be conceptualized, as described in the following section.

### 3. The Framework

#### 3.1. Competence as a Continuum: Person-Situation Interaction

Competence as a latent construct only delivers significant meaning within a specific framework (Erpenbeck and von Rosenstiel 2007). An underlying theoretical framework model is thus indispensable for empirical competence assessment since it allows propositions about the structure, levels, and context of the competence to be measured.

At the very core of our framework model, we follow the idea of Blömeke et al. (2015), who consider competence as a continuum rather than a fixed state. According to the authors, most of the previous approaches to competence assessment follow a dichotomic view, as they focus on either dispositional analytics (i.e., assessing assumed dispositions necessary for competence development) or the measurement of real-life performance (i.e., assessing the observable outcome without taking a closer look at how the outcome was delivered). From an educational point of view, both approaches are insufficient: the first ignores the situational context as well as the fact that even the right dispositions might not necessarily be transformed into the correct behavior; the latter does not enable effective interventions as this approach does not shed light on potential prerequisites that are needed for individual competence development. Blömeke et al. (2015) plead us to consider competence as a continuum between these two dichotomies and to view competence development and assessment as a person-situation interaction where individual dispositions interact with situation-specific aspects that lead to performance in the sense of observable behavior. For our framework, we take up this idea since it is in line with our general understanding of competence as a multi-dimensional construct, but we also take on a more differentiated view in some core aspects of our model.

### 3.2. Dispositions: Individual Preconditions

As stated in the definitions, a person's individual pre-conditions for competence can be categorized into cognitive and non-cognitive dispositions. In essence, cognitive processes refer to mental and intellectual procedures or states that are directly connected with the human brain. In our model, the most relevant cognitive features are knowledge, as "knowledge is the necessary foundation of competence" (Weinert 2001b, p. 45), along with other generic basic skills (numeracy or reading ability), intelligence, and working memory. For dealing with financially relevant issues or situations, a person needs basic knowledge in the domain of personal finance (e.g., how insurance works; knowledge about different buying or payment options) along with the abilities to read and comprehend (e.g., offers, advertising, bills, contracts) and the skills to do basic calculations (e.g., taxes, interest, payment rates).

While the cognitive aspects of our model are quite clear, things get trickier with the non-cognitive aspects of our framework, as various relevant terms (e.g., affect, emotion, motivation, volition, conation, etc.) are often used interchangeably or without real consensus (Rothermund and Eder 2011; Schmidt-Atzert et al. 2014). Basically, non-cognitive dispositions are characteristics that are not purely cognitive or that go beyond mere cognition<sup>1</sup>, and terms such as affect, motivation, or emotion are often used synonymously or as illustrative examples. The American Psychological Association (APA) defines affect as "any experience of feeling or emotion, [...] from the simplest to the most complex sensations of feeling, and from the most normal to the most pathological emotional reactions" (APA 2022). In our opinion, non-cognitive/affective dispositions need to be further categorized according to their duration and situational specificity (situation-specific state vs. situation-independent trait (Fridhandler 1986)), their intensity (low vs. high arousal), valence (positive vs. negative (Russel 1980)), or their reference to specific objects or domains (i.e., interest (Otto et al. 2000)). Consequently, non-cognitive dispositions can be broadly diversified. For our theoretical framework, we refer to them as those affectively shaped dispositions that are in any way relevant to the financial competence we aim to assess.

Weinert (2001b) and the FL-section of PISA (OECD 2013) explicitly mention *motivation* in their respective definitions. This means that a person should show at least a certain basic willingness to engage in financial topics, acquire domain-specific knowledge, and tackle financially relevant tasks. Closely connected to that is a certain level of *interest* in the topic and the *volition* to persistently deal with financial issues. So far, studies have shown that domain-specific motivation and interest in financial topics can yield a positive effect on test scores in a standardized set of economic and personal finance tests (Förster et al. 2018; Happ and Förster 2018). In other words, individuals lacking motivation, volition, and interest will probably have a hard(er) time when it is necessary for them to face financially relevant situations.

The domain of economics and personal finance itself is also characterized by specific features that become directly relevant to an individual's non-cognitive dispositions. One basic principle of economics is the tenet of scarcity. Usually, financial and other resources are limited, and individuals have to evaluate the best resource allocation according to specific *intentions, needs, or attitudes*. While these intentions, needs, or attitudes can be personal and socially shaped in nature, it depends on how much importance and relevance individuals attach to each component when acting. In general, individuals' *attitudes toward financial matters* and *money preferences* have the power to influence personal financial decision-making and are as such important elements of financial literacy (Atkinson and Messy 2012; Rai et al. 2019). Financial attitude shows itself in how people manage their financial resources, for instance, when it comes to the question of short-term consumption vs. long-term saving, or the decision between price and sustainable quality of a product. These aspects are closely linked to dispositions like an individual's *self-control over delayed gratification* (Aprea 2020; Greimel-Fuhrmann 2019). Furthermore, attitudes become assessable through the value people ascribe to money in general, for example, having a lot of money means power and prestige (Lay and Furnham 2019). In this respect, personal aspects of *morals, ethics, or*

preferences concerning *sustainability* can—or maybe even must—also be included in the considerations when it comes to financial decisions, and the question is if people reflect on such issues when they consume. In combination with the sheer complexity of modern financial products and assets, or the way money can be spent in general, it becomes clear that many financial decisions are made under uncertainty. As such, individuals' *risk preferences* suddenly are equally important (Nofsinger 2017). In financial investment, different financial products have different risk and return profiles, as do people when they become financial investors (Menkhoff and Sakha 2016). In this respect, studies have shown that risk preferences and risk perceptions tend to influence investment decisions significantly (Aren and Zengin 2016). In the most basic sense, in financial investment, people have to choose between financial products that are either rather safe but yield a low return (e.g., savings account at a bank) or products that promise much higher returns at the peril of an increased risk of loss (e.g., stocks). General awareness of such risks is an essential attribute for being financially literate. Therefore, risk perception and risk-taking behavior of individuals are integral in personal financial competence.<sup>2</sup> As such, dealing with questions of personal finance also tends to have a strong emotional component (Furnham et al. 2012). Financial behavior can be influenced by a person's emotions, meaning that financial decisions can be heavily biased depending on a person's emotions and feelings (Nofsinger 2017). At the same, personal financial issues have the potential to trigger emotional states (e.g., financial stress (Wrosch et al. 2000), money-related disorders (Canale et al. 2015)).

This brief overview illustrates the complex interlinkages between the personal dispositions of financial competence. We consider these cognitive and non-cognitive dispositions as important individual characteristics that will be relevant for the assessment of financial competence.

### 3.3. Domain: Situation-Specific Characteristics and Financial Problem-Solving

After having outlined the personal parameters of competence, we should take a closer look at the domain, i.e., the situational characteristics, when regarding competence as a person–situation interaction. However, not only the specific parameters of any situation are relevant for competence assessment. Rather, it is the process of how the aforementioned personal characteristics interact with those specific situational parameters. In that sense, Blömeke et al. (2015) speak of situation-specific skills, such as perception, interpretation, and decision-making, that mediate between personal dispositions and performance. In essence, this means that certain domain-specific characteristics define the respective situation. For example, financially relevant situations can be characterized by a particular dynamism, meaning that in an investment setting, the situation can change over time depending on the interplay between supply and demand, interest rates, or inflation. An individual must perceive and interpret such characteristics before an investment decision is made. The aspects of perception, interpretation, and decision-making show great similarities with features of (complex) problem-solving (Fischer et al. 2012; Funke 2019). As such, referencing Blömeke et al.'s understanding to the given definition of competence by Weinert (2014), we understand competence as an individual's problem-solving abilities in variable situations. By integrating the process of problem-solving into our framework model, we can take a more differentiated perspective concerning the actual process (Funke 2019) and find suitable connections for the influence of affectively-shaped and emotional dispositions on the various steps of the process (Springer et al. 2005). It also allows direct links to existing competence models that have already been developed for the domain of economics (Jung 2006). Therefore, we will briefly illustrate our concept of problem-solving as an exemplary description of the situation-specific parameters in the assessment of financial competence.

Competence shows itself in observable behavior in a specific situation. The situation as such must contain a problem—a “challenge” or “barrier” of some sort—that the individual must overcome by implementing cognitive efforts and behavioral activities (Funke 2011). In an idealized depiction, problem-solving is often displayed as a circular, sequential series

of various phases that primarily contain cognitive structures (e.g., recognition and mental representation of a problem, identifying potential solution strategies, goal elaboration, and decision-making, or reflection and re-evaluation of an intended outcome). Depending on the success or failure of a solution, these cognitive procedures can either be run through several times or lead to the generation of new problems (Funke 2019; Pretz et al. 2003). Following this view, we understand problem-solving as a circular sequence of five phases that become relevant for financial competence assessment:

1. Perception: The first step of each problem-solving activity is the identification or recognition of a problem (Pretz et al. 2003). If individuals do not identify a problem, they will most likely not make any effort to solve it. To put it in context, for instance, if a person is not aware of the necessity for retirement planning, strategies to find a pension scheme will probably not be implemented.

2. Appraisal: If a problem is identified, there is probably an increased need for a mental representation along with further categorization. The idea of appraisal processes originates from theoretical concepts in stress research and implies that information is classified according to relevance and quality (Lazarus and Folkman 1984). So, if a problem is identified, it must at least be categorized as relevant in order to be tackled. In other words, if individuals are aware of the need for retirement planning but do not consider it highly relevant, they will probably still not engage in strategies to find a pension scheme, and the process of problem-solving aborts at this stage. Such a possible lack of successful perception or appraisal of a problem as irrelevant is addressed by including an “opt-out” in our process model. In a positive sense, if perception is successful and appraisal is positive (i.e., “relevant”), the process of problem-solving continues.

3. Analysis of potential solutions and anticipation of desired outcomes: The evaluation of potential solutions to the problem follows afterward (Pretz et al. 2003). Not only do people think about strategies to overcome the problem, but they also anticipate how the desired solution or outcome should look (Sell and Schimweg 2002). During this phase, content knowledge about the domain or subject in question is essential. Ideally, along with the main consequences of the solution strategy itself, individuals also take potential side- and after-effects into account, as this is considered to be a mandatory part of competence (Sembill 1992). In context, this means that a person not only engages in actual strategies for retirement planning (for example, by taking out private pension insurance), but is also aware that this might lower the disposable income and that cost-cutting is necessary elsewhere in the monthly spending.

4. Decision (and subsequent action): At some point in the analysis phase, a decision for the best—or at least the most appropriate—solution strategy will be made, and mental and physical efforts will be undertaken to transfer the chosen option into practice (Betsch 2011; Pretz et al. 2003). Decision-making by itself is considered to be a multi-step process that consists of various phases (planning—executing—controlling (Aprea and Wuttke 2016); preselection—selection—post-selection of an option (Betsch et al. 2016)). Furthermore, decisions are a product of the interaction between an individual’s cognitive and affectively shaped dispositions (Betsch 2011). As such, a decision is seen as a mediator between a person and the situation. The mental model built during the analysis phase is ideally transferred into a specific action or behavior (Aprea and Wuttke 2016). In the example of retirement planning, this means that the chosen strategy and mental effort during the analysis phase (e.g., selecting and comparing different pension scheme offers, actively evaluating the fit between insurance conditions and personal need) result in the (physical) action of taking out a pension plan by signing the contract with an insurance company.

5. Reflection and re-organizing: If a decision is made and, at best, transferred into actual behavior, it is necessary to reflect and evaluate it. Reflection is a mental retrospective process of (re-)structuring an experience, a problem, or existing knowledge in order to learn for future incidents (Korthagen 1999). In the example of retirement planning, it is difficult to see a rapid success or failure as retirement planning is a long-term process, and the success or failure of the strategy will only be seen at the end of an individual’s working

life. However, for short-term results, an individual can reflect on the engagement strategy in the problem-solution (e.g., collecting and selecting information on financial products that are suitable for retirement planning) and, by doing so, enhance knowledge and understanding of the domain of personal finance. Reflection is, along with re-structuring prior subject or domain knowledge and the evaluation of solution strategies, often seen as a mandatory phase in educational approaches (e.g., experiential learning (Kolb 1984)) and, as such, in the development of expertise and competence. According to this understanding, an individual's reflection about one's own financial knowledge, personal needs and values concerning money, risk attitudes and perception, and strategies to tackle financial challenges and situations is an essential part of developing financial competence. Thus, reflection should be integral to a framework model concerning competence assessment.

In summary, problem-solving is considered "an orchestration of cognitive, metacognitive [i.e., strategic], and non-cognitive processes [...] to find an initially unknown way of bridging the gap between an actual state and a desired state" (Rausch et al. 2019, p. 2). We anticipate that different affectively shaped and emotional dispositions become relevant at different stages during the cognitive process of (financial) problem-solving and that this might eventually influence competence assessment in the domain of personal finance.

Admittedly, the idea of putting the measurement of competence in relation to (complex) problem-solving abilities is not entirely new. For example, various forms of problem-solving were implemented in the OECD's international PISA studies, where problem-solving was defined as a cross-disciplinary skill to solve real-life problems (Leutner et al. 2004; OECD 2004, 2005). Although some similarities in the theoretical concept of the overall construct occur in comparison to the presented approach, PISA's construct operationalization was quite distinct from ours. The main differences were that the PISA studies implemented mostly or partly standardized item formats embedded in relatively short tasks with a limited degree of complexity (OECD 2013). Furthermore, to the best of our knowledge, financial literacy, and problem-solving have never been correlated in PISA so far.

In a more domain-specific approach, Ackermann (2019) also integrated basic ideas of problem-solving for the assessment of economic competence. In her study, she developed a three-part framework model for the measurement of civic-economic competence, which is based on a similar theoretical foundation as the model we aim to present in this contribution. Differentiated into four iterative steps (problem identification, multi-perspective analysis, comparison of solutions, and evaluation), she conceived problem-solving as a cognitive process that becomes relevant for an individual to handle economically shaped demands. However, in contrast to our draft, the focus of her study lies more on the domain of general economics and its adjoining areas (e.g., politics) in Switzerland.

### 3.4. Financial Competence and Performance Measurement

The theoretical considerations described in the previous sections are now integrated in a complex framework model that is depicted in Figure 1. Personal dispositions that are necessary prerequisites for financial competence are shown on the model's left side. In the center, the phases of problem-solving are outlined as they are necessary for an individual to deal with the specific characteristics of financially relevant situations. The model's right side depicts the instruments with which we want to measure an individual's performance.

We assess performance in a financial competence test with a set of instruments. The first—and optional—instrument is a standardized test for measuring knowledge and understanding in the field of personal finance. Depending on the target audience and its respective cognitive level (e.g., upper secondary education, high school learners, university students, etc.), this can be the Test of Financial Literacy (TFL) (Förster et al. 2017a; Walstad and Rebeck 2017), the Test of Financial Knowledge (TFK) (Kraitzek 2020; Walstad and Rebeck 2018) or any other financial literacy test available. The standardized test by itself, however, neither allows conclusions about non-cognitive aspects of personal finance nor does it shed light on the actual process of dealing with financial situations.

The second—and mandatory—instrument is a complex performance scenario that has been developed according to the seven facets that characterize competence assessments as suggested by Shavelson (2012), specifically to address the shortcomings that were outlined before. This scenario will be discussed in the next section.

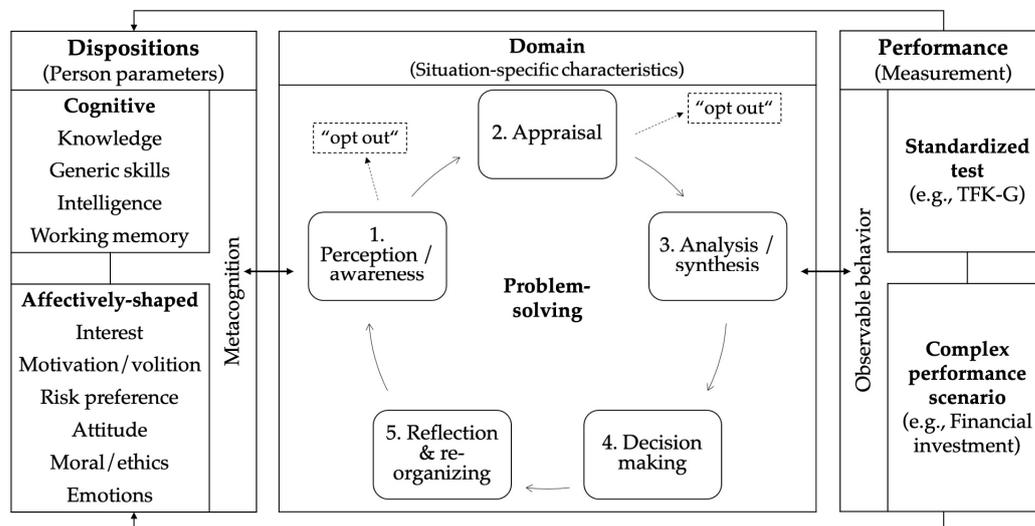


Figure 1. Integrated theoretical framework for financial competence assessment.

#### 4. A Complex Performance Scenario for the Assessment of Financial Investment

The complex scenario introduced in this section can be seen as an illustrative example of how an instrument for competence assessment integrates the aforementioned theories.

##### 4.1. Description of the Scenario Setting

The scenario’s story is about Sofie, a 16-year-old student who received 5000 euros as a gift from her grandparents for her high-school graduation. Overwhelmed by that large amount of money, she asks her best friend for help about financial investment.

Along with the setting, the reader is given information that Sofie also has some desires she wants to fulfill at different times in the near future (e.g., paying for her driver’s license, buying her brother’s car). Additionally, the information is given that she will continue in higher-secondary education for the next three years, that she will keep living with her parents at home, and that she will not have additional income as she does not want to get a job during that time.

Sofie asks various parties for advice: her father, who is portrayed as an “expert” in stock markets but actually gives her faulty and insufficient advice; her mother and her aunt, who suggest investing in a fixed-term deposit account with specific conditions (e.g., a rather low interest rate), and her grandmother who recommends saving a certain percentage of the money for emergencies. Furthermore, Sofie has picked out offers for a daily allowance account along with a few predefined options for stocks of different companies.

The test taker is supposed to take the role of Sofie’s friend. The task is to give her a well-founded recommendation for financial investment, taking into account her desires and the pros and cons of the various investment options. Along with this real-life task, real-world information and supplementary material are provided in a realistic way (e.g., diagrams with share prices of various companies, newspaper reports with comments about stock market development, chat messages with peers about financial products), as suggested by Shavelson (2012). Furthermore, explanations about the *magic triangle*<sup>3</sup> of financial investment are given. Important information about the respective financial assets, which is relevant to the solution, is also hidden within this material. For example, the offer about the daily allowance account contains a pitfall in the terms and conditions as the full interest rate is paid up to a certain amount of money and for a limited period of time

only. For the test taker to give Sofie a sound recommendation, it is important to become thoroughly acquainted with the material and to separate these bits of relevant information from those which are irrelevant or even harmful to the solution.

In total, the assessment scenario in its original form consists of 7 to 10 pages in a pen-and-paper format. This includes the description of the situation, the supplementary information material, and space for the actual written response. Depending on the target group and its performance level, the response length may be predefined (e.g., “Please write 250 to 350 words”). The administration of the scenario is designed for a timeframe of 45 to 60 min.

#### 4.2. Designing the Scenario

When designing an assessment for an educational context, it is important to consider that the content of the instrument is anchored in the school’s curriculum (content validation). Although there is no unified curriculum for economic and financial education in Germany (Kraitzek et al. 2022), the content of this specific scenario is linked to various curricula of vocational and commercial schools. The point of reference for the development of the financial investment scenario was the 10th-grade curriculum of “Wirtschaftsschulen” (a school type in Bavaria, Southern Germany, that focuses on economic and vocational education). The core content of the scenario is explicitly anchored in one of the curriculum’s main study areas called “Preparation for life” (ISB 2014). The basic skills necessary for the scenario (e.g., arguing, structuring texts, comprehension of extensive materials, calculations) should have already been achieved by the end of previous grades.

In designing the scenario, we followed the idea of *construct modeling* as outlined in the BEAR Assessment System (Berkeley Evaluation and Assessment Research) and as described by Wilson (2005, 2009). This system utilizes four *building blocks* (construct map, item design, outcome space, measurement model) that provide structure for the scenario’s iterative development (Wilson 2009).

The construct map as the starting point defines the construct to be measured. Essentially, the construct map is a graphical representation that outlines the continuum between (the) low(est) and high(est) possible manifestations of the construct. It provides a structure for the qualitatively distinct levels of a test taker’s result in order to classify it as “high”, “mediocre”, or “low” according to predefined criteria.

The item design is the actual construction of the scenario based on theoretical recommendations. At this stage, specific features of the scenario item (e.g., task format, response type, content, quality of the material, and its representation) were selected and implemented. For example, to ensure authenticity, a “real-world” problem in the domain of personal finance was chosen (Hedtke 2011; Shavelson 2012). Additionally, a set of difficulty-determining features (e.g., the syntactic level of the language, use of domain-specific terminology) was chosen a priori in order to achieve an appropriate level of difficulty for the target group; for an overview of such features, see: Schumann and Eberle (2011); for a discussion, see: Förster et al. (2017b).

The outcome space is one of the most important parts of the scenario since it provides a reference to which the test taker’s solution is compared. It is basically a horizon of expectations for each level of the construct’s manifestation. Providing a guideline for a “good”, “mediocre”, or “poor” solution to the scenario increases both raters’ objectivity and the reliability of the assessment.

Closely linked to the outcome space is the evaluation scheme, as it provides numerical values (“scores”) for the answers. The measurement model defines how inferences about the construct can be drawn from the scores, and this helps to determine a test subject’s location on the construct map (Wilson 2009). In other words, both the outcome space and measurement model lead to the development of a final scoring rubric with which the test participants can be evaluated. Table A1 (Appendix A) provides an exemplary draft of a preliminary outcome space with a potential scoring rubric for the financial investment scenario.

#### 4.3. Discussion of the Scenario

Working on the scenario and the solution is a multi-tier process: first, test takers must recognize that for each of Sofie's desires, some financial assets fit better than others, and some do not fit at all. This evaluation requires the identification of specific clues hidden in the supplementary material. Second, participants need to calculate the correct amount of money that must be assigned to each asset. Third, the remaining amount of money can be assigned by the test takers at will according to their preferences with respect to the magic triangle. For example, a participant with high risk aversion will tend to put the remaining amount of money in financial assets that are rather secure, whereas a person with tendencies for higher risk taking will potentially start to gamble and spend the amount on highly volatile shares. There is not one best solution to the task—although some solutions are better than others. Rather, it is important that the recommendation to Sofie is supported by a solid argumentative structure (e.g., "I would rather choose X because of Y;" "There are some trade-offs between Product A and Product B in terms of C, and that is why I would choose B for XYZ", etc.). The more elaborate a person's recommendation, the more sophisticated the level of financial competence appears to be. The written response must then be referenced to a predefined outcome space (see example in Table A1, Appendix A) to allow for a solid evaluation of the solution's quality.

In terms of problem-solving, the scenario addresses almost all steps in the circular approach: first, the test taker has to identify the problem (i.e., not every asset is suitable for each desire and due to the principle of scarcity, trade-offs are necessary according to one's personal financial type or character). If the problem is not perceived, problem-solving, and as such financial competence, is not possible. In the next step, the information material provided with the task has to be evaluated according to its usefulness, reliability, and relevance (e.g., does the information in the newspapers support the idea of investing in stocks, and is the information even reliable). Furthermore, it must be compared to one's anticipation (e.g., is a particular financial asset suitable for Sofie's needs). Eventually, a decision has to be made and put into practice (i.e., formulation of an investment recommendation for Sofie that addresses her desires). Ideally, this decision is reflected upon after working on the task and, depending on the outcome, a re-organization of content and process knowledge occurs (e.g., finding a better solution after the answers are compared in class or in any given educational setting).

The investment scenario also addresses the five facets of complex problems (Funke 2019). The scenario is *complex* in nature as it contains a large amount of information that has to be processed. For example, there are multiple options for investment products, and each product contains various advantages and disadvantages for the task at hand. To solve the problem, a selection of relevant information and, as such, a reduction of complexity is necessary. Additionally, these variables are *interconnected*. As defined by the magic triangle, there is a relationship between profitability (i.e., return), financial security (i.e., risk of loss), and liquidity (i.e., quick access to money) of the financial assets in question. Choosing one form of investment over another means that certain disadvantages have to be accepted. This is closely connected to the principle of *polytely*, which means that multiple conflicting options for the solution exist. In terms of *dynamics*, the assessment is rather static, and the provided charts of stock prices are just an extract of a specific and fixed timeframe. However, the test taker must anticipate the future development of the stocks—or of the financial markets in general—by analyzing the additional information given in the material. With this method, we address at least the basic idea of dynamic problems, although a full consideration of the financial domain as a dynamic environment would imply using adaptive testing methods via computer. The anticipation of the stock market's development is also a prime example of the *lack of transparency*. Not all of the relevant bits of information are either given or visible at first sight. Rather, some are hidden deep inside the supplementary material, and this requires a thorough analysis of the financial assets in question. Others are simply not given at all, as it is impossible to make a detailed forecast of economic development in the provided material.

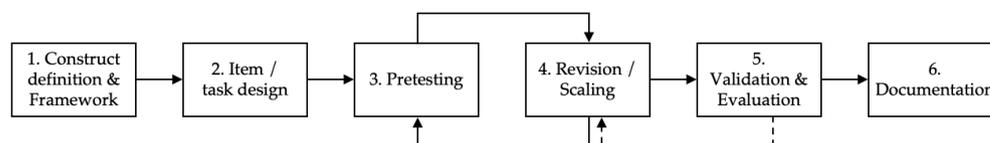
Concerning cognitive dispositions, a person needs at least some basic knowledge and understanding of the principles of financial investment (e.g., knowledge about interest rates; understanding of profit and how it is influenced). Additionally, generic skills (e.g., reading, calculations, argumentative writing) are necessary. As already outlined, problem-solving as a cognitive process is influenced by affectively-shaped and emotional dispositions (Güss et al. 2017). For example, a person who is highly motivated and interested in the topic of personal finance will probably give a more elaborate and comprehensive recommendation for an investment strategy than a person who lacks motivation for the domain or generally has a negative attitude towards financial matters. Risk preferences can be assessed via the magic triangle of investment. After the necessary calculations in the scenario, the test takers are left with a remaining amount of money they can invest freely at will. This decision may potentially allow insights into an individual's investment type, which can then be categorized according to the different objectives in the triangle (profit/return, risk/security, liquidity/accessibility; Bundesbank (2023)). Additionally, statements about long- or short-term investment can allow insights into one's financial attitude. Last but not least, a person's choice between the stocks of an oil company or a sustainable energy company can shed light on their ethical compass in financial investment. As stated in the construct definition, all these aspects become relevant to a person's financial competence.

In summary, the financial investment scenario can be seen as an illustrative example of an alternative form of competence assessment. The complex instrument has the potential to measure financial competence as a complex construct in a (more) holistic manner.

## 5. Pretesting and Findings

### 5.1. Methodology

In the development of assessment instruments, pretesting is an essential part of the process. To verify our theoretical assumptions, the financial investment scenario has been pretested over various trials to derive findings for its iterative development. The development process for the current project is outlined in Figure 2.



**Figure 2.** Process of the current project.

During the pretesting phase, studies with different versions of the scenario that existed at the respective time were conducted with varying intentions. For one thing, the aim of the pretests was to test the item's comprehensibility for the sample and to further improve the scenario with each iteration. Another goal was to gradually refine the scoring scheme on the basis of evidence for the various (sub-)constructs to be measured. Finally, eye tracking (ET) and emotion recognition (ER) technology were used to investigate whether the scenario—and the topic altogether—triggers emotional reactions in the test takers or whether there is no conspicuous emotional involvement at all.

For each study, drafts for scoring schemes were created to provide at least a minimal point of reference for the rating of the students' solutions. For this, a combination of deductive and inductive qualitative methods was used: prior to the investigation, the sub-constructs of financial competence were derived from literature; additionally, after the test taker's solutions were collected, they were also analyzed inductively to assign respective statements to each sub-construct. The frequencies of statements were counted and also qualitatively assessed to identify distinct levels (e.g., good vs. poor solution). Due to the limited sample sizes the evaluation of both the scenario and the scoring rubric often remained on a descriptive level. Based on these insights, the scenario's material and the scoring rubric was iteratively revised and enhanced.

To measure whether test takers' affectively shaped and emotional dispositions occur at all during work on the financial scenario, we implemented modern computer-based technology. One idea was to investigate in an explorative manner if specific emotional patterns of test takers processing the scenario could be identified or if certain steps in the process of financial problem-solving are interrelated with explicit emotional states.

Emotion recognition (ER), in this respect, is based on the idea that emotions contain an expression component that shows itself in facial expressions (Frenzel et al. 2015). Such expressions are defined by *action units*, i.e., significant points in the human face that form specific expression patterns through activity of facial muscles (e.g., the raising of the eyebrows and wide-open eyes indicate "surprise"). Various expressions are categorized according to the Facial Action Coding System (FACS) developed by Ekman and Friesen (1976). With ER technology, a test taker's face is recorded while working on a task, and computer algorithms continuously and instantaneously analyze the mimic musculature and, thus, the emotional states through facial expression. The results are presented in both numerical and graphic ways and allow insights into a person's valence and the arousal state of basic emotions as well as different affective states; for details about the Circumplex Model of Affect, see: Russel (1980). With this method, it is possible to generate a continuous sampling of emotional states throughout the processing of the complex performance scenario.

Additionally, a screen-based and computerized delivery of the investment scenario also allows the implementation of eye tracking (ET) as an additional research method. When applied to a screen-based assessment, this method can provide information about the visualized stimulus material and, as such, offer inferential evidence about the underlying cognitive processes (e.g., attentional control) during the processing of the assessment (Yaneva et al. 2021). Through the immediacy assumption and the eye-mind assumption (Just and Carpenter 1980), it is believed that a stimulus triggers cognitive processes at the exact moment this stimulus is visually perceived and fixated.

ER and ET recordings produce timestamps which can then be used to synchronize both data streams. As such, the combination of ER and ET technology may provide an innovative way to assess emotional and affective states that are directly linked to the processing of a complex financial performance scenario. In other words, through eye tracking and emotion recognition, we are able to connect an individual's emotional and affective state to the cognitive processing of financial content in the investment scenario. Thus, we can draw inferences about how domain content (personal finance), method (problem-solving) and affectively shaped dispositions (emotions) are interrelated. So far, ER and ET were used in only one pretest, as will be outlined below.

### 5.2. First Pretest: Overall Feasibility

After the initial design, an early version of the scenario was tested in the 9th grade of a *Wirtschaftsschule* in Southern Germany in the summer of 2019. The focus of the pretest was on general feasibility, item clarity (e.g., difficulty, comprehensibility), and the identification of potential pitfalls and problems. The sample consisted of  $n = 16$  students (10 male, 6 female) that worked on the scenario for a timeframe of 60 min. Afterward, a group discussion with the test takers was conducted about the pros and cons of the instrument, along with participants' anonymous feedback about the scenario via an audience response system. Additionally, various domain-related variables as well as questions about the perceived difficulty of the test, were assessed via Likert-scaled items (for descriptive statistics, see Table 1). The participants found the scenario to be of medium difficulty ( $M = 2.32$ ,  $SD = 0.557$ ; "0 = very difficult" to "4 = very easy"), but the overall solution rate was rather poor to average. A closer look at the given answers revealed that in terms of content and topic relevancy, the scenario seemed to be appropriate for the target group. However, based on the self-reported knowledge, students seemed to overestimate their skills and ability in dealing with financial topics, for example, by mixing-up technical terms. We also found that the students tended to have difficulties with the open-ended task

format and the response type: the written answers were rather short and mostly lacked an argumentative structure for the choice of the investment form. We also observed that some important steps in the solution process were missing (e.g., the calculation of the amount of money). However, at that time, it was not clear whether this was due to a lack of procedural knowledge (i.e., domain-specific problem-solving) or a lack of motivation in a low-stake assessment.

**Table 1.** Sample description for the first pretest.

	M	SD	Min	Max	n
Interest in financial topics *	3.13	0.500	2	4	16
Importance of financial investment *	4.06	0.680	3	5	16
Importance of money *	3.88	0.957	2	5	16
Self-reported knowledge about financial matters *	3.00	1.100	1	5	16
Word count in the written solution	117.69	62.567	36	266	16

n = 16; \* (1 = lowest; 5 = highest).

### 5.3. Second Pretest: Integration of a Decision Matrix

Based on the findings from the first field test, some adjustments were made to the task description and answer requirements (e.g., some additional information about the minimal length of the answer along with instructions concerning calculations were added). Furthermore, the informational material was slightly varied and updated (e.g., updated graphs about stock market development). The second pretest was carried out with n = 12 10th-grade students of another *Wirtschaftsschule* in Southern Germany in the winter of 2019/2020. For the work on the scenario, the students had a total timeframe of 60 min. According to our test protocols, the first students handed in the solution after 35 min. Other students complained that it took them almost 30 min to just read the whole material. Some students were not able to process the scenario in this timeframe at all and had to abort their work after 60 min. This is an indicator for the heterogeneity of our target group. For this trial, no other data on sociodemographic or domain-related variables were collected. However, structured interviews with two classroom teachers were conducted to investigate the instructors' perspectives and impressions of the scenario, the supplementary material, and the potential problems that may occur during classroom implementation. One finding from these interviews resulted in the integration of a *decision matrix* in which the students can collect their ideas and arguments before the final response is written. Essentially, the decision matrix is an additional document where the pros and cons for any given investment product can be accumulated before they are assigned to Sofie's desires; for other examples of decision matrices, see: Grünig and Kühn (2013). As such, the decision matrix can be seen as a scaffolding tool that helps to prestructure the task at hand. The integration of this decision matrix led to an improvement in the overall solution, although test protocols also revealed that students still had a lot of questions about the open task type per se (e.g., it was still unclear what to do with the remaining money). The students' solutions were further evaluated by two raters with a modified preliminary outcome space. Calculation of inter-rater reliability statistics revealed Cohen's Kappa values for three randomly selected answers of  $\kappa_1 = 0.727, p < 0.001$ ;  $\kappa_2 = 0.471, p = 0.01$ ; and  $\kappa_3 = 0.500, p < 0.013$  indicating a moderate to substantial agreement between the raters (Landis and Koch 1977). This can be seen as a good starting point for the evidence-based development of an outcome space.

### 5.4. Third Pretest: Combination of the TFK and the Scenario

The third pretest was conducted shortly afterward in the winter of 2019/2020 in the 10th grade of a *Berufsschule* (school for vocational education and training) in Southern Germany. The sample consisted of n = 21 trainees in retail (10 male, 11 female) who worked on the scenario for 45 min. For this trial, specific domain-related variables were collected and sample descriptions are given in Table 2. Additionally, for the first time after its adaption into a German context (Kraitzek 2020), the TFK (Walstad and Rebeck 2018)

was administered in combination with the investment scenario in a two-tiered assessment session. While scoring the TFK is relatively simple due to its “correct-or-wrong” type of single-choice items, the scoring of the scenario outcomes had been done with another preliminary version of an outcome space in combination with a rudimentary scoring rubric. The central findings were that the optional decision matrix was used by a majority of the students, and participants who used the matrix tended to produce better solutions in the scenario than those who did not use it. Based on this scoring scheme, the findings also revealed significant positive correlations between the scenario solution and the TFK’s overall scores ( $r = 0.574^{**}, p = 0.003$ ) as well as between the scenario solution and the scores of the TFK’s content area of financial investment ( $r = 0.505^{**}, p = 0.01$ ). With respect to the AERA standards for evidence-based validity (AERA, APA, and NCME 2014), this can be seen as an indicator that both the TFK and the investment scenario are comparable in terms of content, which speaks for the validity of the instruments. Another finding was that female participants produced better results in the scenario than male test takers. Although these differences were not significant, this is an interesting result given the outcomes from other studies where constructed response answers might have the potential to reduce the previously identified gender gap in economic and financial literacy assessments with multiple-choice items (Siegfried and Wuttke 2019). Despite these findings, one problem remains: a lack of motivation, given the fact that 10 out of 21 students wrote very short answers (less than 75 words) and only 6 students wrote responses of medium length (between 75 and 125 words). Participants’ statements collected after the test support this assumption. The students often complained about the high amount of text-based information that had to be processed. A lack of motivation to read longer texts eventually results in low scores for the solution process, as an elaborate argumentative structure in the written response is missing. This is in line with the impression that students tend to have problems with the open-ended task format and the response type. The integration of the decision matrix might seem like a potential solution to this issue, but it also has its disadvantages concerning the goal of competence assessment. This dilemma will be discussed in the limitations.

**Table 2.** Sample description for the third pretest.

		M	SD	Min	Max	n
Age (years)	15–18					12
	19–25					9
Years of economic education in school	none					6
	1 year					5
	2 years					1
	3 years					5
	4 years and more					4
Interest in financial topics *		2.62	1.117	1	5	21
Perceived difficulty of the assessment **		2.05	0.524	1	3	21
Word count in the written solution		91.90	68.847	7	272	21
Scenario scores ***	overall	17.33	8.193	4	36	21
	male	15.90	8.293	4	32	10
	female	18.64	8.274	9	36	11
TFK-G scores †	overall	20.33	6.529	11	32	21
	male	21.10	5.705	14	28	10
	female	19.64	7.406	11	32	11

n = 21; \* (1 = lowest; 5 = highest); \*\* (1 = easy; 2 = medium; 3 = hard); \*\*\* Based on the scoring scheme at that time, the max. score of the scenario was 39 points. † This early version of the adapted TFK consisted of 38 items, i.e., a max. score of 38 was possible. Note: due to data protection guidelines, some data could only be collected in a classified manner.

5.5. Fourth Pretest: Assessment of Affectively-Shaped Dispositions and Emotional States

The first attempt to investigate the scenario’s general capabilities to measure domain-specific affectively shaped dispositions and emotional states was made in the summer of 2022. For this trial, the scenario was transferred into a computer-based format.

To prove the applicability and feasibility of the methodological combination of ER and ET, we conducted an experiment in a student research seminar. In this trial, n = 9 randomly selected student test takers (4 male, 5 female) had to work on a computer-based version of the financial investment scenario while being recorded for the purpose of eye tracking and emotion recognition. Descriptions of this sample are presented in Table 3. Based on prior findings during the initial pretesting phase of the scenario, it was assumed that the topic of personal finance by itself would interfere with the emotional state of the test takers. This was due to our impression during the school pretesting phase as the subject of personal financial investment did not appear to be one of the student’s favorite topics. This is also in line with the assumption that personal finance and dealing with financial topics are likely to contain a strong emotional component, as described earlier.

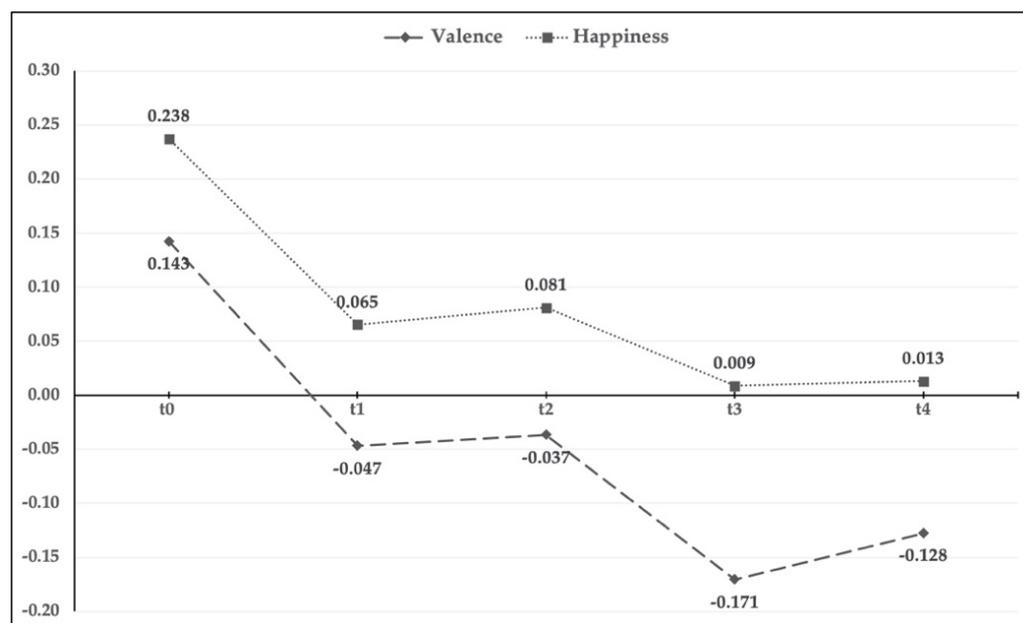
**Table 3.** Sample description for the fourth pretest.

	M	SD	Min	Max	n
Age	26.44	5.790	19	39	9
Highest degree of education					1
					2
					5
					1
Interest in financial topics *	5.22	1.563	2	7	9
Importance of financial topics *	5.00	2.121	2	7	9
Word count in the written solution	92.56	102.905	17	349	9

n = 9; \* (1 = lowest; 7 = highest).

The study was designed in a way that the test takers at first did not know that an assessment in the field of personal financial competence would follow. Five times of measurement were set during the course of one testing cycle: after a short introduction, the students were told to answer general questions not directly related to personal finance (t0). This initial period served as a baseline for the ER measurements and was used to calibrate the ER and ET tools for the subsequent measures. After completion of the initial questionnaire, the test takers were told that they would then encounter a task about personal financial investment (t1). After that, the test takers answered the “Big Three” (Lusardi 2019) financial literacy questions (t2) before they started with the investment scenario. These short questions served as a benchmark for the students’ financial knowledge as well as a brief introduction to the topic. As work on the scenario is not necessarily linear, the next measurements were linked to specific actions within the scenario: (t3) was set when the test takers looked at one of the stock charts that were provided with the information material, (t4) was set when the test takers clicked on the “submit solution” button after they had written their final investment recommendation.

After data collection was completed, the times of measurement were used to synchronize the data from the ER and ET measures according to the respective timestamps. This allowed us to calculate the mean values of the emotional states for the periods between measurements to explore the changes. Figure 3 shows the descriptive development of the mean values for the n = 9 test takers’ “emotional valence” and their “happiness” over the course of the work on the scenario.



**Figure 3.** Means for *valence* and *happiness* during the scenario processing;  $n = 9$ .

These basic descriptive results indicate that at the beginning of the study, the test takers experienced, on average, positive emotions as the valence value for  $t_{\text{Valence}}(0) = 0.143$  is higher than zero. After the test takers had learned that they had to work on a task concerning financial issues, this value dropped to  $t_{\text{Valence}}(1) = -0.047$ . The negative value indicates that negative emotions were experienced more strongly than before. Negatively-shaped feelings remained for the rest of the work on the scenario as the negative values for  $t_{\text{Valence}}(2) = -0.037$ ,  $t_{\text{Valence}}(3) = -0.171$ , and  $t_{\text{Valence}}(4) = -0.128$  indicate.

A similar trend is observable in the measured values for “happiness”. At the beginning of the assessment, test takers showed, on average, a relatively high happiness value with  $t_{\text{Happiness}}(0) = 0.238$ . This value dropped to  $t_{\text{Happiness}}(1) = 0.065$  after they learned about the subsequent financial task. Except for a small rise in  $t_{\text{Happiness}}(2) = 0.081$ , this negative trend seemed to last, as the low values for  $t_{\text{Happiness}}(3) = 0.009$  and  $t_{\text{Happiness}}(4) = 0.013$  indicate. Whereas the values of the overall valence, i.e., the quality of the emotional state, can range from  $-1$  (overall negative emotions) to  $+1$  (overall positive emotions), the value range from “happiness”, i.e., a per se positive emotion, only comprises values from  $0$  (not happy at all) to  $+1$  (very happy).

The observed development of these values indicates that just the announcement of a financial task seems to compromise the emotional condition of people. In a *ceteris paribus* experimental setting, the simple knowledge about a pending financial task might potentially influence test takers’ affective states. Of course, in further investigations, additional exploration will be needed to elaborate if, for example, the test setting by itself has an effect on people’s emotional states as well. Nevertheless, it is suspected that the domain component has a considerable influence on the way people feel about a task and how they tackle challenges that require competent behavior, like in the domain of personal finance.

At this point, the investigations of this trial remain on a purely descriptive level due to the small sample size. Consequently, this little study does not allow for generalizable implications. However, it does illustrate the potential of the combination of ER and ET as modern technology-based methods for holistic competence assessment, and it provides first evidence that emotional facets can be assessed during financial problem solving and that there is variance throughout the process. More extensive investigations with larger and more representative samples along with different foci (e.g., the inclusion of the overall quality of the scenario’s solution) are scheduled for the near future.

## 6. Discussion: Limitations and Outlook

All pretests have been highly beneficial for the development of the complex performance scenario, and the instrument has already shown its potential.

A first key finding was that the pretests and the initial impressions on the produced solutions provided a suitable starting point for the iterative development of the outcome space and the scoring rubric for the investment scenario. It became clear that for scoring, the overall multi-faceted construct will most likely have to be differentiated into its different components according to the construct definition, thus creating a measurement model with various (sub-)dimensions that are rated separately (Table A1 in Appendix A). For each (sub-)facet, anchor examples are needed to enhance the scoring rubric for raters in order to give them a qualitative guideline for quantitative scoring. An illustrative example for this technique in the criterion-sampling approach can be found in Zlatkin-Troitschanskaia et al. (2019). Still, although this procedure enables a general evaluation of the scenario outcomes, scoring remains a challenge due to limitations which will be discussed below.

### 6.1. Single-Item Testing

Although assessment via complex performance scenarios is an innovative strategy that addresses many shortcomings of previous competence diagnostics, the approach is not without limitations. One major critique is that the concept is basically “single-item testing” (Allen et al. 2022; Funke 2010; Greiff and Fischer 2013). The scenario is essentially one extensive item that aims to represent the underlying construct in a holistic manner. As such, it draws inferences about this construct from the test taker’s results from just one item. This naturally has the potential to cause constraints concerning construct validity, as the multidimensionality of the underlying complex construct may distort the clarity of the results. In other words, there is a risk that it is not fully clear what the item measures when just looking at the outcomes on a superficial level. For example, when assessing financial competence as defined above, the construct consists of a content component (i.e., the actual content of the personal finance domain) and a process component (i.e., the problem-solving). If the assessment eventually shows poor results for a participant, this could be due to a lack of content knowledge or due to a lack of problem-solving ability. Of course, with respect to the construct definition—i.e., competence as the combination of person parameters (cognitive and non-cognitive dispositions) and situational parameters (the process of problem-solving)—the assessment by itself seems valid, but the implications drawn from the results could be blurred. One way to counteract this problem is to aim for convergent validity, which means to control the content aspect of the construct with comparable assessment instruments as an additional source of validity evidence (AERA, APA, and NCME 2014). This is why we additionally plan to implement the standardized test in the framework of our assessment approach.

The problem of single-item testing leads to the fact that the scenario is only able to display a fraction of the domain in which it is embedded. We aim to assess competence in the domain of personal finance with only one single item about financial investment, but we are tempted to draw inferences from this assessment about the domain as a whole. For example, some content areas of personal finance (e.g., insuring) are not covered in the investment scenario. Yet, inferences drawn from the relatively narrow scenario about investment might be carelessly transferred to the whole domain, which would include the content of insuring, although this was not assessed intentionally. This means that while the scenario provides an in-depth assessment of this specific area, the results concerning the domain altogether should only be interpreted with caution. In the context of test development, this problem is often referred to as a *validity-reliability dilemma*, as reliability usually increases with a small number of more or less homogeneous items, whereas validity increases the broader—and as such, the more heterogeneous—the assessment is designed in terms of content (Seifried et al. 2020).

### 6.2. Risk of Subjectivity

Another challenge with performance scenarios is directly related to the open-ended response type of the task itself. Since the outcomes of the (problem-solving) process are not standardized—or even in any way fully predictable—the scoring of the answers might contain subjective components. As such, the objectivity of the analysis cannot entirely be guaranteed, in contrast to a standardized multiple-choice assessment that leaves no room for a rater's interpretation. This underlines the necessity for a well-founded scoring rubric that integrates the outcome space and the measurement model as described in the BEAR Assessment Process (Wilson 2005, 2009). The scoring scheme should include prototype answers which act as a guideline for raters when it comes to scoring. The prototype answers should be sampled from real-life pretesting results with the target group since a literature-only derivation of potential answers or solutions might not be exhaustive enough to establish a sound scoring guide. The answers could also be evaluated by more raters to identify whether significant deviations between the interpretation of some aspects of the problem scenario and the respective answers occur. From this, inter-rater reliability values could be calculated so that the essential prerequisites for the development of a scoring scheme are met. However, due to the high expenditure of rating open-ended answers or short essays, especially when multiple raters are involved, trade-offs between efficiency and validity or objectivity will be likely. One possible solution to this dilemma could be the rise of automated essay scoring based on artificial intelligence, but this opens up a whole new level of reliability issues and questions about legal challenges in high-stake educational testing (Ferrara and Qunbar 2022). Closely related to the problem of subjectivity is the fact that some construct-sensitive contents in the scenario cannot be uniquely differentiated and thus, not be clearly quantified for scoring. The clear allocation of the test takers' answers to the respective (sub-)constructs—and eventually the scoring—gets difficult due to the “fuzziness” of the categories and linguistic vagueness. One solution to this challenge might be the future implementation of fuzzy set theory and its methodological approaches as alternative forms of scoring multi-criteria decisions like they occur in the presented scenario; for an illustrative example, see: Hussain and Merigo (2022).

### 6.3. Reduced Content

Another key finding was that both the language level regarding technical terms as well as the content of the scenario appeared to be generally appropriate for the target group. However, the complexity of the instrument's structure requires a narrowing of the scope of financial content to be measured. In 2021, the U.S. Council for Economic Education (CEE) published the revised *National Standards for Personal Financial Education*, which categorize the domain of personal finance into six major content areas: (1) earning income, (2) spending, (3) saving, (4) investing, (5) managing credit, and (6) managing risk (CEE 2021). In this respect, the content of the presented assessment scenario is mainly about financial investment as described in content area (4) of the 2021 Standards. Nevertheless, the scenario partly addresses other content areas as well (e.g., saving or spending). This is a connection to the demand that FL requires knowledge, skills, and motivation in more than one key content area (Remund 2010). Thus, one challenge for future research efforts might be the development of one big scenario that integrates more—or maybe all—of the CEE content areas into one instrument, although this might result in a trade-off between complexity and diagnostic validity, as discussed earlier. Another possibility is to combine additional scenarios with different content areas into a larger assessment battery, following the content structure of the TFL or TFK that also combines all of the CEE's content areas in one instrument.

### 6.4. Dilemma: Prestructuring and Nuancing the Level of Complexity

One key finding of the early pretests was that many students of the target group were overstrained with the scenario's level of complexity and its minimalistic structure. Although the introduction of a decision matrix in one of the pretests helped to improve

the assessment solutions, it was also a source for discussion about the general goal of the assessment in itself. The decision matrix was inspired by the idea of instructional scaffolding and the learning approach of Cognitive Apprenticeship (Brandt et al. 1993). Scaffolding, in an educational sense, refers to the initial support teachers give to learners until they can carry on alone. Strictly speaking, the decision matrix as a scaffold reduces the originally intended openness of the task format by providing additional structure. Since openness and minimalistic structure is said to be one of the prerequisites for the measurement of competence as defined (Shavelson 2012), additional structure counteracts the intention of assessment in itself. Transferring this to the topic at hand, many complex real-world problems that require an individual's competent behavior to find a solution (such as complex financial challenges) do not come in a prestructured manner either. It is an integral part of one's competence development to train the way complex, unstructured problems are tackled and as such, handling complexity should also be an essential part of competence assessment. This is especially true for a complex topic like financial literacy. Reducing complexity might lead to a potentially dangerous over-simplification of financial issues. This might result in students' overconfidence and perhaps over-estimation of their financial competence—which in turn might have potentially harmful effects on their real-world financial behavior (Aprea et al. 2015). On the other hand, too much complexity not only leads to problems and additional challenges in diagnostics and assessment, but also can lower students' motivation to deal with the important topic of financial literacy (as was observable in the pretests). Eventually, this could have a negative effect on their financial education per se.

Against the background of this dilemma, we developed ideas to modify the scenario according to a continuum ranging from "fully structured" to "fully open/unstructured." We created different versions of the scenario which are equal in terms of content, but vary in terms of structure (i.e., different degrees of educational scaffolding). In future experimental studies, these different versions could be randomly distributed in a classroom to evaluate which type of educational scaffolding fits best for the target group. One advantage of this approach would be a better match of the scenario to any target group in terms of cognitive level. Furthermore, the transition from assessment scenario to learning instrument for educational purposes and vice versa could be much more fluent, for example, by starting educational efforts with a fully structured instrument and then gradually moving to the completely unstructured open version.

#### *6.5. Alternatives for the Assessment of Affectively-Shaped Dispositions and Emotional States*

One major key finding of the pretest involving ER and ET technology was that the performance scenario about financial investment appears to interfere with the test takers' emotional states even in a low-stake setting. As such, on a purely descriptive level, this finding supports our assumption that dealing with personal financial challenges contains a considerable emotional component. However, an in-depth assessment of affectively shaped and domain-specific emotions still remains a challenge for a couple of reasons.

First, the measurement of affective dispositions and specific emotional states depends heavily on the exact operationalization of these factors. While this is rather easy for some affective dispositions (e.g., financial risk preference), it is quite difficult for others (e.g., financial ethics and morality; financial attitudes and values). Essentially, the challenge is to assess such constructs without literally "assessing" them in the sense of a separate implementation of standardized questionnaires. It is thus necessary to identify suitable proxies for the previously described non-cognitive dispositions within the test takers' written answers. These proxies must allow raters to draw conclusions about the underlying affective constructs.

Second, while ET and ER appear to be promising concerning the assessment of emotional states, using these methods is quite cumbersome as they require a solid technical infrastructure along with the availability of and access to these technologies. Furthermore, both the studies and the data processing can be very time-consuming as a test taker's work on the scenario is not necessarily linear, and each trial and each recording has to be processed individually.

An alternative way is a pen-and-paper delivery of the scenario which could be enhanced with a low-threshold method of assessing affectively shaped and emotional dispositions. In this respect, following the approach of [Rausch et al. \(2019\)](#), we developed one version of the financial investment scenario, which implemented the method of Embedded Experience Sampling (EES). EES aims to collect self-report data on various non-cognitive facets in the way that the prompts are directly embedded in the actual situation and, therefore, in the task itself ([Rausch et al. 2019](#)). These short prompts are an integral part of the actual storyline of the problem situation. For example, throughout the financial investment scenario, the test person receives (fictitious) messages from Sofie, who asks how the test person feels about the task at this exact moment, how the coping with the task at hand is going, or how confident the person is of solving the problem. These prompts are (repeatedly) embedded in the printed test booklet at different places, for example, right after the initial problem situation is presented or right before the answer has to be written. To date, a draft of the scenario containing EES exists, but it has not yet been pretested. Nevertheless, [Rausch et al. \(2019\)](#) have shown that EES can generally produce reliable and valid data on non-cognitive facets of problem-solving competence without causing too much distraction from the actual work on the task itself.

## 7. Conclusions

In this contribution, we described an alternative approach to financial competence assessment by pointing out three major aspects. First, we suggested modeling the construct of financial competence in a new, holistic, and complex way, as the construct by itself is regarded to be complex, multi-layered, and multi-dimensional. Second, we outlined the development of a theoretical framework model that takes the construct's complexity into account and further addresses the limitations of prior attempts in financial competence assessment. Third, we illustrated an innovative performance scenario for the topic of financial investment, which is built directly onto this framework and allows the assessment of financial competence as a complex construct in a holistic and integrative way. To the best of our knowledge, this featured scenario is the most complex assessment tool for financial competence and financial problem-solving for the target group. Such a complex measurement item comes with a multitude of challenges that require an evidence-based development process including mixed-method research approaches. To date, the first steps for validating the scenario have been made, and the results of the pretests have contributed to the iterative improvement of the instrument. Additionally, we were able to show that affective and emotional states are related to test takers' work on the scenario and that these states varied during the course of solving the presented problem. However, to produce generalizable results as a reliable basis for educational interventions, the assessment must be implemented in a larger-scale study with a more representative sample of the target group. The basis for valid and reliable scoring requires sophisticated qualitative research approaches prior to explanative, quantitative hypotheses testing. Consequently, additional in-depth interviews with teachers and instructional designers about the scenario may also produce useful insights and contributions from an educator's perspective. The implementation of technology-based research methods such as modern eye tracking and emotion recognition in the future may further open up completely new ways to investigate the process component of financial problem-solving as well as its interconnections of emotions and affective dispositions within financial competence.

The described performance assessment is an alternative way to assess competences not only in the field of personal finance but can also be transferred into other domains as well. We would like to invite researchers to see this as a potential starting point for a holistic approach to assessment as a part of learning. The instrument is capable of revealing in which areas of the domain there are still difficulties for learners (e.g., in knowledge, in the process of problem-solving or in the affective and emotional aspects of learning). In the future, educators and policy makers can use these findings to tailor educational concepts to specific target groups in different educational settings. As such, the presented scenario

about financial investment has the potential to be linked to or even transferred directly into classroom practices and instructional activities (Wilson 2009). So far, the instrument was designed primarily as a closed assessment tool where test participants are asked to work only with the information material provided with the task. It would, however, require only some minor modifications to use the scenario as a learning assignment for the classroom. For example, opening up the task so that the learners can look up additional possible investment products for themselves and evaluate whether these could fit Sofie’s needs. Or students can use this scenario to build their own investment portfolio based on the research and the discoveries they made while working on the scenario. Such learning assignments can further be stretched to multiple learning sequences and eventually be interconnected with other topics in the field of personal finance, as well as with other relevant subjects.

**Author Contributions:** Conceptualization, A.K. and M.F.; methodology, A.K.; validation, A.K.; formal analysis, A.K.; investigation, A.K.; resources, M.F.; data curation, A.K. and M.F.; writing—original draft preparation, A.K.; writing—review and editing, A.K.; visualization, A.K.; supervision, M.F. and A.K.; project administration, A.K. and M.F.; funding acquisition, M.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** The APC was funded by the Technical University of Munich.

**Institutional Review Board Statement:** The TUM Ethics Committee provided guidelines for subject information and participant education that were used for this study.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Due to data protection rights the research data cannot be provided.

**Acknowledgments:** We would like to thank the principals and teachers of the participating schools, as well as the respective school authorities, for their support and administrative approval. We would also like to thank all students who contributed directly or indirectly to this project, whether it was through fruitful discussions, the sharing of conceptual ideas, or the collecting of data during research seminars.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Exemplary outcome space with a potential scoring rubric for a complex financial performance scenario (extract); sample provided for the “financial investment” scenario.

Construct	Sub-Construct	Definition	Operationalization and Attributes	Levels	Numerical Values	Weighting (Optional)	Reference Example(s)
Financial Competence	Financial content knowledge	...	E.g., use of technical terms, diversification of the investment portfolio, correct allocation of desires to financial products, etc.	High	3	...	...
				Medium	2		...
				Low	1		...
				None/not evident	0		...
	Problem solving abilities	...	E.g., identification of conflicting goals/trade-offs between investment products, anticipation of potential side effects, reflection about the final solution, etc.	High	3	...	...
				Medium	2		...
				Low	1		...
				None/not evident	0		...
	Calculations (Numeracy)	...	E.g., remaining amount of money, interest rates in %, etc.	Fully correct	2	...	...
				Partly correct	1		...
				Wrong/none	0		...
	Risk preferences/financial attitudes	...	E.g., use of the remaining amount of money according to the magic triangle	Security	0/1 *	-	...
Liquidity				0/1 *	...		
Profitability				0/1 *	...		
Sum	...	...	...	...	...	...	Σ

\* Due to the nominal character of these sub-constructs, they may be treated as categorical variables with a dichotomous scoring (e.g., 0 = non-existent/requirements not fulfilled; 1 = existent/requirements fulfilled).

## Notes

- <sup>1</sup> Non-cognitive/affective dispositions also contain cognitive components. A clear separation between the two is only expedient for the framework's theoretical illustration. For example, emotions as affectively-shaped dispositions originate in the limbic system in the human brain and are as such cognitive in nature. Modern psychological theories assume a bidirectional relationship between cognition and emotion as both constructs are thought most likely to be interdependent (Pekrun 2000).
- <sup>2</sup> Research on factors influencing risk-taking and risk perception is not limited to the area of individual finance and personal money management. Similar investigations can be found in the field of corporate finance, banking or financial management in businesses, e.g., Nguyen (2022).
- <sup>3</sup> In a German context, the relationship between security, profitability, and liquidity in financial investment is called the “magic triangle” as it is not possible for an investment to be safe, instantly accessible, and offer high returns all at the same time (Bundesbank 2023; Kühn and Kühn 2020).

## References

- Ackermann, Nicole. 2019. *Wirtschaftsbürgerliche Kompetenz Deutschschweizer Gymnasiastinnen und Gymnasiasten: Kompetenzmodellierung, Testentwicklung und evidenzbasierte Validierung*. Zurich: Universität Zürich, Philosophische Fakultät.
- AERA, APA, and NCME. 2014. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Allen, Mark S., Dragos Iliescu, and Samuel Greiff. 2022. Single Item Measures in Psychological Science. *European Journal of Psychological Assessment* 38: 1–5. [CrossRef]
- APA. 2022. Affect. Available online: <https://dictionary.apa.org/affect> (accessed on 19 August 2022).
- Aprèa, Carmela. 2020. Financial Literacy in der beruflichen Bildung: Ein kompetenzorientierter Zugang. [Financial literacy in Vocational Education: A Competency-Based Approach.] *bwp@Profil 6: Digitale Festschrift für Eveline Wuttke*: 1–18. Available online: [https://www.bwpat.de/profil6\\_wuttke/aprea\\_profil6.pdf](https://www.bwpat.de/profil6_wuttke/aprea_profil6.pdf) (accessed on 29 March 2023).
- Aprèa, Carmela, and Eveline Wuttke. 2016. Financial Literacy of Adolescents and Young Adults: Setting the Course for a Competence-Oriented Assessment Instrument. In *International Handbook of Financial Literacy*. Edited by Carmela Aprèa, Eveline Wuttke, Klaus Breuer, Noi Keng Koh, Peter Davies, Bettina Greimel-Fuhrmann and Jane S. Lopus. Singapore: Springer Singapore, pp. 397–414.
- Aprèa, Carmela, Eveline Wuttke, and Michaela Stock. 2015. Exploring the possibilities for a bildungs-oriented conceptualisation of financial literacy education. In *Instrumentalism in Education—Where Is Bildung Left?* Edited by Stefanie Hillen and Carmela Aprèa. Münster: Waxmann, pp. 89–104.
- Aren, Selim, and Asiye Nur Zengin. 2016. Influence of Financial Literacy and Risk Perception on Choice of Investment. *Procedia Social and Behavioral Sciences* 235: 656–63. [CrossRef]
- Atkinson, Adele, and Flore-Anne Messy. 2012. Measuring Financial Literacy: Results of the OECD/International Network on Financial Education (INFE) Pilot Study. In *OECD Working Papers on Finance, Insurance and Private Pensions*. Paris: OECD Publishing.
- Betsch, Tilmann. 2011. Entscheiden. [Decision making]. In *Denken—Urteilen, Entscheiden—Problemlösen. [Thinking—Judging, Deciding—Problem-Solving]*. Edited by Tilmann Betsch, Joachim Funke and Henning Plessner. Berlin/Heidelberg: Springer, pp. 66–77.
- Betsch, Tilmann, Johannes Ritter, Anna Lang, and Stefanie Lindow. 2016. Thinking beyond Boundaries. In *Cognitive Unconscious and Human Rationality*. Edited by Laura Macchi, Maria Bagassi and Riccardo Viale. Cambridge, MA: The MIT Press, pp. 241–55.
- Billimoria, Jeroo, Monique Cohen, Jared Penner, and Floor Knoote. 2013. Developing the next generation of economic citizens: Financial inclusion and education for children and youth. *Enterprise Development and Microfinance* 24: 204–17. [CrossRef]
- Blömeke, Sigrid, Jan-Eric Gustafsson, and Richard J. Shavelson. 2015. Beyond Dichotomies. *Zeitschrift für Psychologie* 223: 3–13. [CrossRef]
- Bongini, Paola, Paola Iannello, Emanuela E. Rinaldi, Mariangela Zenga, and Alessandro Antonietti. 2018. The challenge of assessing financial literacy: Alternative data analysis methods within the Italian context. *Empirical Research in Vocational Education and Training* 10: 1–22. [CrossRef]
- Brandt, Barbara L., James A. Farmer, and Annette Buckmaster. 1993. Cognitive apprenticeship approach to helping adults learn. *New Directions for Adult and Continuing Education* 1993: 69–78. [CrossRef]
- Breuer, Klaus. 2016. Assessment of Financial Literacy. In *International Handbook of Financial Literacy*. Edited by Carmela Aprèa, Eveline Wuttke, Klaus Breuer, Noi Keng Koh, Peter Davies, Bettina Greimel-Fuhrmann and Jane S. Lopus. Singapore: Springer Science, pp. 381–82.
- Brückner, Sebastian, Manuel Förster, Olga Zlatkin-Troitschanskaia, and William B. Walstad. 2015. Effects of prior economic education, native language, and gender on economic knowledge of first-year students in higher education. A comparative study between Germany and the USA. *Studies in Higher Education* 40: 437–53. [CrossRef]
- Bundesbank. 2023. Glossary—The “Magic Triangle” of Financial Investment. Available online: <https://www.bundesbank.de/action/en/729724/bbksearch?firstLetter=T> (accessed on 15 January 2023).
- Canale, Anthony, Kristy L. Archuleta, and Bradley T. Klontz. 2015. Money Disorders. In *Financial Therapy*. Edited by Bradley T. Klontz, Sonya L. Britt and Kristy L. Archuleta. Cham: Springer, pp. 35–67.

- CEE. 2021. *National Standards for Personal Financial Education*. New York: CEE Council for Economic Education; Jump\$tart Coalition for Personal Financial Literacy. Available online: <https://www.councilforeconed.org/wp-content/uploads/2021/10/2021-National-Standards-for-Personal-Financial-Education.pdf> (accessed on 29 March 2023).
- Cude, Brenda J. 2022. Defining Financial Literacy. In *The Routledge Handbook of Financial Literacy*. Edited by Gianni Nicolini and Brenda J. Cude. London and New York: Routledge T&F Group, pp. 5–17.
- Davies, Peter. 2006. Educating citizens for changing economies. *Journal of Curriculum Studies* 38: 15–30. [CrossRef]
- Dubs, Rolf. 2014. *Unterrichtsplanung in der Praxis: Ein Handbuch für den Lernbereich Wirtschaft*. [Lesson Planning in Practice: A Handbook for the Learning Area of Economics]. Stuttgart: Franz Steiner.
- Eberle, Franz, Stephan Schumann, Esther Kaufmann, Andreas Jüttler, and Nicole Ackermann. 2016. Modellierung und Messung wirtschaftsbürgerlicher Kompetenz von kaufmännischen Auszubildenden in der Schweiz und in Deutschland (CoBALIT). [Modeling and Measurement of Business Civic Competence of Commercial Apprentices in Switzerland and Germany (CoBALIT)]. In *Technologiebasierte Kompetenzmessung in der beruflichen Bildung: Ergebnisse aus der BMBF-Förderinitiative ASCOT*. [Technology-Based Competence Measurement in Vocational Education and Training: Results from the BMBF Funding Initiative ASCOT]. Edited by Klaus Beck, Margarete Landenberger and Fritz Oser. Bielefeld: Wbv, pp. 93–117.
- Ekman, Paul, and Wallace V. Friesen. 1976. Measuring facial movement. *Environmental Psychology and Nonverbal Behavior* 1: 56–75. [CrossRef]
- Erpenbeck, John, and Lutz von Rosenstiel. 2007. Einführung. [Introduction]. In *Handbuch Kompetenzmessung: Erkennen, verstehen und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis* [Handbook of Competence Measurement: Recognizing, Understanding and Assessing Competences in Operational, Pedagogical and Psychological Practice]. Edited by John Erpenbeck and Lutz von Rosenstiel. Stuttgart: Schaeffer-Poeschel, pp. XVII–XXIX.
- Erpenbeck, John, Sven Grote, and Werner Sauter. 2017. Einführung. In *Handbuch Kompetenzmessung: Erkennen, Verstehen und bewerten von Kompetenzen in der Betrieblichen, Pädagogischen und Psychologischen Praxis*. [Handbook of Competence Measurement: Recognizing, Understanding and Assessing Competences in Operational, Pedagogical and Psychological Practice]. Edited by John Erpenbeck, Lutz von Rosenstiel, Sven Grote and Werner Sauter. Stuttgart: Schäffer-Poeschel, pp. IX–XXIX.
- Ferrara, Steve, and Saed Qunbar. 2022. Validity Arguments for AI-Based Automated Scores: Essay Scoring as an Illustration. *Journal of Educational Measurement* 59: 288–313. [CrossRef]
- Fischer, Andreas, Samuel Greiff, and Joachim Funke. 2012. The Process of Solving Complex Problems. *The Journal of Problem Solving* 4: 19–42. [CrossRef]
- Förster, Manuel, Roland Happ, and Andreas Maur. 2018. The relationship among gender, interest in financial topics and understanding of personal finance. *Empirische Pädagogik* 32: 293–309.
- Förster, Manuel, Roland Happ, and Dimitar Molerov. 2017a. Using the U.S. Test of Financial Literacy in Germany—Adaptation and validation. *The Journal of Economic Education* 48: 123–35.
- Förster, Manuel, Sebastian Brückner, Roland Happ, Klaus Beck, and Olga Zlatkin-Troitschanskaia. 2017b. Strukturanalyse eines kognitiven Messinstruments im Multiple Choice-Format: Das Beispiel des Test of Economic Literacy (TEL4-G). [Structural Analysis of a Cognitive Measurement Instrument in Multiple Choice Format: The Example of the Test of Economic Literacy (TEL4-G)]. *Zeitschrift für Berufs- und Wirtschaftspädagogik* 113: 366–96. [CrossRef]
- Frenzel, Anne C., Thomas Götz, and Reinhard Pekrun. 2015. Emotionen. [Emotions]. In *Pädagogische Psychologie*. [Educational Psychology]. Edited by Elke Wild and Jens Möller. Berlin, Heidelberg: Springer, pp. 201–24.
- Fridhandler, Bram M. 1986. Conceptual note on state, trait, and the state–trait distinction. *Journal of Personality and Social Psychology* 50: 169–74. [CrossRef]
- Funke, Joachim. 2010. Complex problem solving: A case for complex cognition? *Cognitive Processing* 11: 133–42. [CrossRef]
- Funke, Joachim. 2011. Problemlösen. [Problem-solving]. In *Denken—Urteilen, Entscheiden, Problemlösen*. [Thinking—Judging, Deciding—Problem-Solving]. Edited by Tilmann Betsch, Joachim Funke and Henning Plessner. Berlin/Heidelberg: Springer Medizin, pp. 136–59.
- Funke, Joachim. 2019. Problem Solving. In *The Psychology of human Thought. An Introduction*. Edited by Robert J. Sternberg and Joachim Funke. Heidelberg: Heidelberg University Publishing, pp. 155–76.
- Furnham, Adrian, Emma Wilson, and Kate Telford. 2012. The meaning of money: The validation of a short money-types measure. *Personality and Individual Differences* 52: 707–11. [CrossRef]
- Fürstenau, Bärbel, Mandy Hommel, Manuel Förster, Andreas Kraitzek, Eveline Wuttke, Carmela Aprea, Michelle Rudeloff, and Christin Siegfried. 2020. Messung von Finanzkompetenz—Ergebnisse eines Symposiums. [Measuring financial literacy—Results of a symposium]. In *Jahrbuch der Berufs- und Wirtschaftspädagogischen Forschung 2020* [Yearbook of Vocational and Business Education Research 2020]. Edited by Eveline Wittmann, Dietmar Frommberger and Ulrike Weyland. Opladen, Berlin and Toronto: Barbara Budrich, pp. 33–64.
- Geiger, Jan-Martin, Ute Meretz, and Andreas Liening. 2016. Systematisierung deutschsprachiger Studien zur Kompetenzerfassung von financial literacy. [Systematization of German-language studies on the assessment of financial literacy competencies]. *Zeitschrift für ökonomische Bildung* 5: 72–93.
- Goyal, Kirti, and Satish Kumar. 2020. Financial literacy: A systematic review and bibliometric analysis. *International Journal of Consumer Studies* 45: 80–105. [CrossRef]

- Greiff, Samuel, and Andreas Fischer. 2013. Measuring complex problem solving: An educational application of psychological theories. *Journal for Educational Research Online* 5: 38–58. [CrossRef]
- Greimel-Fuhrmann, Bettina. 2019. Why wait? Über die Bedeutung des Belohnungsaufschubs bei ökonomischen Entscheidungen. [Why wait? On the Importance of Reward Deferral in Economic Decision Making]. *Berufs- und Wirtschaftspädagogik online bwp@Österreich Spezial* 2: 1–16.
- Grohs-Müller, Stefan. 2020. *Jugendliche und ihr Umgang mit Geld. Eine empirische Studie zu Erfahrungen, Einstellungen und Verhaltensweisen von Schüler/innen am Ende der Sekundarstufe I. [Adolescents and Their Approach to Money. An Empirical Study on the Experiences, Attitudes, and Behaviors of Students at the End of Lower Secondary School]*. Wien: WU Universitätsverlag/Facultas.
- Grünig, Rudolf, and Richard Kühn. 2013. *Prozess zur Lösung komplexer Entscheidungsprobleme. [Process for Solving Complex Decision Problems]*. Berlin: Springer.
- Güss, C. Dominik, Madison Lee Burger, and Dietrich Dörner. 2017. The Role of Motivation in Complex Problem Solving. *Frontiers in Psychology* 8: 1–5. [CrossRef] [PubMed]
- Happ, Roland, and Manuel Förster. 2018. The correlation between vocational school students' test motivation and the performance in a standardized test of economic knowledge: Using direct and indirect indicators of test motivation. *Empirical Research in Vocational Education and Training* 10: 1–18. [CrossRef]
- Hedtke, Reinhold. 2011. *Konzepte ökonomischer Bildung. [Concepts of Economic Education]*. Schwalbach: Wochenschau Verlag.
- Hussain, Walayat, and Jose M. Merigo. 2022. Onsite/offsite social commerce adoption for SMEs using fuzzy linguistic decision making in complex framework. *Journal of Ambient Intelligence and Humanized Computing*. [CrossRef]
- Huston, Sandra J. 2010. Measuring Financial Literacy. *Journal of Consumer Affairs* 44: 296–316. [CrossRef]
- ISB. 2014. *LehrplanPLUS Wirtschaftsschule. Lehrplan für die bayerische Wirtschaftsschule, Vierstufig, Dreistufig und Zweistufig. [CurriculumPLUS Business School. Curriculum for the Bavarian Business School, Four-Level, Three-Level and Two-Level]*. Available online: <https://www.lehrplanplus.bayern.de/sixcms/media.php/107/LehrplanPLUS%20Wirtschaftsschule%20-%20August%202014.pdf> (accessed on 5 September 2022).
- Jung, Eberhard. 2006. Möglichkeiten zur Überprüfung von Kompetenzmodellen in der Ökonomischen Bildung. [Ways to test competency models in economics education]. In *Kompetenzentwicklung, -förderung und -prüfung in der ökonomischen Bildung. [Competence Development, Promotion and Testing in Economic Education]*. Edited by Bernd O. Weitz. Bergisch-Gladbach: Verlag Thomas Hobein, pp. 33–60.
- Just, Marcel A., and Patricia A. Carpenter. 1980. A theory of reading: From eye fixations to comprehension. *Psychological Review* 87: 329–54. [CrossRef]
- Kaiser, Tim, and Andreas Lutter. 2015. Empirische Forschung zu Financial Literacy. *Zugänge—Befunde—Desiderata. [Empirical research on financial literacy. Approaches—Findings—Desiderata]*. *Zeitschrift für Didaktik der Gesellschaftswissenschaften* 6: 77–95.
- Kolb, David A. 1984. *Experiential learning: Experience as the Source of Learning and Development*. Englewood Cliffs: Prentice-Hall.
- Korthagen, Fred A. J. 1999. Linking Reflection and Technical Competence: The logbook as an instrument in teacher education. *European Journal of Teacher Education* 22: 191–207. [CrossRef]
- Kraitzek, Andreas. 2020. Utilization of the U.S. Test of Financial Knowledge in Germany—Adaption and Validation. Poster Presentation. Paper Presented at Annual Conference of the American Educational Research Association (AERA) 2020, San Francisco, CA, USA, April 19.
- Kraitzek, Andreas, Manuel Förster, and William B. Walstad. 2022. Comparison of financial education and knowledge in the United States and Germany: Curriculum and assessment. *Research in Comparative and International Education* 17: 153–73. [CrossRef]
- Kühn, Stefanie, and Markus Kühn. 2020. *Handbuch Geldanlage. [Handbook of Investment]*. Berlin: Stiftung Warentest, Finanztest.
- Landis, J. Richard, and Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33: 159–74. [CrossRef]
- Lay, Alixe, and Adrian Furnham. 2019. A New Money Attitudes Questionnaire. *European Journal of Psychological Assessment* 35: 813–22. [CrossRef]
- Lazarus, Richard S., and Susan Folkman. 1984. *Stress, Appraisal, And Coping*. New York: Springer Publishing Company.
- Leutner, Detlev, Eckhard Klieme, Katja Meyer, and Joachim Wirth. 2004. Problemlösen. [Problem-solving]. In *PISA 2003—Der Bildungsstand der Jugendlichen in Deutschland—Ergebnisse des zweiten int. Vergleichs. [PISA 2003—The Educational Level of Adolescents in Germany—Results of the Second Int. Comparison]*. Edited by Manfred Prenzel, Jürgen Baumert, Werner Blum, Rainer Lehmann, Detlev Leutner, Michael Neubrand, Rainer Pekrun, Hans-Günter Rolff, Jürgen Rost and Ulrich Schiefele. Münster, New York, München and Berlin: Waxmann, pp. 147–75.
- Lusardi, Annamaria. 2012. Numeracy, Financial Literacy, and Financial Decision-Making. *NBER Working Paper* 17821: 1–14.
- Lusardi, Annamaria. 2019. Financial literacy and the need for financial education: Evidence and implications. *Swiss Journal of Economics and Statistics* 155: 1. [CrossRef]
- Menkhoff, Lukas, and Sahra Sakha. 2016. Wie die Kombination von Messungen helfen kann, die Risikobereitschaft besser einzuschätzen. [How combining measurements can help better assess risk tolerance]. *DIW Wochenbericht* 83: 1008–17.
- Nguyen, Quang Khai. 2022. Audit committee effectiveness, bank efficiency and risk-taking: Evidence in ASEAN countries. *Cogent Business & Management* 9: 1–12. [CrossRef]
- Nofsinger, John R. 2017. *The Psychology of Investing*, 6th ed. New York: Routledge.

- OECD. 2004. *The PISA 2003 Assessment Framework*. Paris: OECD Publishing. Available online: <https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/33694881.pdf> (accessed on 29 March 2023).
- OECD. 2005. *Problem Solving for Tomorrow's World*. Paris: OECD Publishing. Available online: <https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/34009000.pdf> (accessed on 29 March 2023).
- OECD. 2013. *PISA 2012 Assessment and Analytical Framework*. Paris: OECD Publishing. [CrossRef]
- OECD. 2019. *PISA 2018 Assessment and Analytical Framework*. Paris: OECD Publishing. [CrossRef]
- Otto, Jürgen H., Harald A. Euler, and Heinz Mandl. 2000. Begriffsbestimmung. [Definitions]. In *Emotionspsychologie—Ein Handbuch [Emotional Psychology—A Handbook]*. Edited by Jürgen H. Otto, Harald A. Euler and Heinz Mandl. Weinheim: Beltz Psychologie Verlagsunion, pp. 11–18.
- Ouachani, Sonia, Olfa Belhassine, and Aïda Kammoun. 2021. Measuring financial literacy: A literature review. *Managerial Finance* 47: 266–81. [CrossRef]
- Pekrun, Rainer. 2000. Persönlichkeit und Emotion. [Personality and Emotions]. In *Emotionspsychologie. Ein Handbuch [Emotional Psychology—A Handbook]*. Edited by Jürgen H. Otto, Harald A. Euler and Heinz Mandl. Weinheim: Beltz Psychologie Verlagsunion, pp. 334–48.
- Pretz, Jean E., Adam J. Naples, and Robert J. Sternberg. 2003. Recognizing, Defining, and Representing Problems. In *The Psychology of Problem-Solving*. Edited by Janet E. Davidson and Robert J. Sternberg. Cambridge: Cambridge University Press, pp. 3–30.
- Rai, Kamini, Shikha Dua, and Miklesh Yadav. 2019. Association of Financial Attitude, Financial Behaviour and Financial Knowledge Towards Financial Literacy: A Structural Equation Modeling Approach. *FIIB Business Review* 8: 51–60. [CrossRef]
- Rausch, Andreas, Kristina Kögler, and Jürgen Seifried. 2019. Validation of Embedded Experience Sampling (EES) for Measuring Non-cognitive Facets of Problem-Solving Competence in Scenario-Based Assessments. *Frontiers in Psychology* 10: 1–16. [CrossRef]
- Reifner, Udo. 2011. Finanzielle Allgemeinbildung und ökonomische Bildung. [Financial literacy and economic education]. In *Finanzielle Bildung in der Schule. Mündige Verbraucher durch Konsumentenbildung. [Financial Literacy in Schools. Mature Consumers through Consumer Education]*. Edited by Thomas Retzmann. Schwalbach: Wochenschau Wissenschaft, pp. 9–30.
- Remund, David L. 2010. Financial Literacy Explicated: The Case for a Clearer Definition in an Increasingly Complex Economy. *Journal of Consumer Affairs* 44: 276–95. [CrossRef]
- Retzmann, Thomas, and Günther Seeber. 2016. Financial Education in General Education Schools: A Competence Model. In *International Handbook of Financial Literacy*. Edited by Carmela Aprea, Eveline Wuttke, Klaus Breuer, Noi Keng Koh, Peter Davies, Bettina Greimel-Fuhrmann and Jane S. Lopus. Singapore: Springer Singapore, pp. 9–23.
- Rieger, Marc O. 2020. How to Measure Financial Literacy? *Journal of Risk and Financial Management* 13: 324. [CrossRef]
- Rothermund, Klaus, and Andreas Eder. 2011. *Allgemeine Psychologie: Motivation und Emotion. [General Psychology: Motivation and Emotion]*. Wiesbaden: Springer.
- Rudeloff, Michelle. 2019. *Der Einfluss informeller Lerngelegenheiten auf die Finanzkompetenz von Lernenden am Ende der Sekundarstufe I. [The Impact of Informal Learning Opportunities On Learners' Financial Literacy at the End of Lower Secondary School]*. Wiesbaden: Springer.
- Russel, James A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39: 1161–78. [CrossRef]
- Schmidt-Atzert, Lothar, Martin Peper, and Gerhard Stemmler. 2014. *Emotionspsychologie. [Emotional Psychology]*. Stuttgart: Verlag W. Kohlhammer.
- Schuhen, Michael, Susanne Kollmann, Minou Seitz, Gunnar Mau, and Manuel Froitzheim. 2022. Financial Literacy of Adults in Germany FILSA Study Results. *Journal of Risk and Financial Management* 15: 488. [CrossRef]
- Schumann, Stephan, and Franz Eberle. 2011. Bedeutung und Verwendung schwierigkeitsbestimmender Aufgabenmerkmale für die Erfassung ökonomischer und beruflicher Kompetenzen. [Significance and use of difficulty-determining task features for the assessment of economic and vocational competencies]. In *Grundlagenforschung zum Dualen System und Kompetenzentwicklung in der Lehrerbildung. [Basic Research on the Dual System and Competence Development in Teacher Education]*. Edited by Uwe Faßhauer, Bärbel Fürstenau and Wuttke Eveline. Opladen: Barbara Budrich, pp. 77–89.
- Schürkmann, Susanne. 2017. *FILS: Financial Literacy Study. Validierung und Analyse einer schülerorientierten Financial Literacy. [FILS: Financial Literacy Study. Validation and Analysis of Student-Oriented Financial Literacy]*. Berlin and Boston: De Gruyter.
- Seeber, Günther, and Thomas Retzmann. 2017. Financial Literacy—Finanzielle (Grund-)Bildung—Ökonomische Bildung. [Financial Literacy—Financial Education—Economic Education]. *Vierteljahreshefte zur Wirtschaftsforschung* 86: 69–80. [CrossRef]
- Seifried, Jürgen, Steffen Brandt, Kristina Kögler, and Andreas Rausch. 2020. The computer-based assessment of domain-specific problem-solving competence—A three-step scoring procedure. *Cogent Education* 7: 1719571. [CrossRef]
- Sell, Robert, and Ralf Schimweg. 2002. Probleme lösen. In *komplexen Zusammenhängen denken*, 6th ed. [Solve Problems. Thinking in Complex Contexts, 6th ed]. Berlin/Heidelberg and New York: Springer.
- Sembill, Detlef. 1992. *Problemlösefähigkeit, Handlungskompetenz und Emotionale Befindlichkeit. Zielgrößen Forschenden Lernens. [Problem-Solving Ability, Action Competence, and Emotional State. Target Values of Research-Based Learning]*. Göttingen, Toronto and Zürich: Hogrefe Verlag für Psychologie.
- Sembill, Detlef, Andreas Rausch, and Kristina Kögler. 2013. Non-Cognitive Facets of Competence. In *From Diagnostics to Learning Success: Proceedings in Vocational Education and Training*. Edited by Klaus Beck and Olga Zlatkin-Troitschanskaia. Rotterdam: SensePublishers, pp. 199–211.
- Shavelson, Richard J. 2012. Assessing business-planning competence using the Collegiate Learning Assessment as a prototype. *Empirical Research in Vocational Education and Training* 4: 77–90. [CrossRef]

- Siegfried, Christin, and Eveline Wuttke. 2019. Are multiple-choice items unfair? *And if so, for whom? Citizenship, Social and Economics Education* 18: 198–217. [[CrossRef](#)]
- Spering, Miriam, Dietrich Wagener, and Joachim Funke. 2005. Brief Report. *Cognition & Emotion* 19: 1252–61. [[CrossRef](#)]
- Stolper, Oscar A., and Andreas Walter. 2017. Financial Literacy, Financial Advice, and Financial Behavior. *Journal of Business Economics* 87: 581–643. [[CrossRef](#)]
- Walstad, William B., and Ken Rebeck. 2017. The Test of Financial Literacy: Development and measurement characteristics. *The Journal of Economic Education* 48: 113–22. [[CrossRef](#)]
- Walstad, William B., and Ken Rebeck. 2018. The measurement properties of the Basic Finance Test for children and the Test of Financial Knowledge for youth. In *Empirische Pädagogik*. Edited by M. Förster, R. Happ, W. B. Walstad and C. J. Asarta. Landau: Verlag Empirische Pädagogik, pp. 248–71.
- Weinert, Franz E. 2001a. Competencies and Key Competencies: Educational Perspective. In *International Encyclopedia of the Social & Behavioral Sciences*. Edited by Neil J. Smelser and Paul B. Baltes. Oxford: Pergamon, pp. 2433–36.
- Weinert, Franz E. 2001b. Concept of competence: A conceptual clarification. In *Defining and Selecting Key Competencies*. Edited by Dominique S. Rychen and Laura H. Salganik. Seattle, Toronto: Hogrefe & Huber Publishers, pp. 45–65.
- Weinert, Franz E. 2014. *Leistungsmessung in Schulen*, 3rd ed. [Performance Measurement In Schools, 3rd ed.]. Weinheim: Beltz.
- Wilson, Mark. 2005. *Constructing Measures: An Item Response Modeling Approach*. Mahwah: Routledge.
- Wilson, Mark. 2009. Assessment from the Ground up. *Phi Delta Kappan* 91: 68–71. [[CrossRef](#)]
- Wrosch, Carsten, Jutta Heckhausen, and Margie E. Lachman. 2000. Primary and secondary control strategies for managing health and financial stress across adulthood. *Psychology and Aging* 15: 387–99. [[CrossRef](#)] [[PubMed](#)]
- Wuttke, Eveline, Stephan Schumann, and Jürgen Seifried. 2016. Editorial: Economic Competence and Financial Literacy of Young Adults—Status and Challenges. In *Economic Competence and Financial Literacy of Young Adults. Status and Challenges*. Edited by Eveline Wuttke, Jürgen Seifried and Stephan Schumann. Berlin and Toronto: Barbara Budrich.
- Yaneva, Victoria, Brian E. Clauser, Amy Morales, and Miguel Paniagua. 2021. Using Eye-Tracking Data as Part of the Validity Argument for Multiple-Choice Questions: A Demonstration. *Journal of Educational Measurement* 58: 515–37. [[CrossRef](#)]
- Zlatkin-Troitschanskaia, Olga, Richard J. Shavelson, Susanne Schmidt, and Klaus Beck. 2019. On the complementarity of holistic and analytic approaches to performance assessment scoring. *British Journal of Educational Psychology* 89: 468–84. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.