Chapter 16

Modeling and Fostering Decision-Making Competencies Regarding Challenging Issues of Sustainable Development

Susanne Bögeholz, Sabina Eggert, Carolin Ziese, and Marcus Hasselhorn

Abstract A model of decision-making competence for secondary school students was developed and validated within the project "Decision-Making Competence Regarding Challenging Issues of Sustainable Development". The model rests on three pillars: Education for Sustainable Development, decision-making theory, and educational competence modeling. Three dimensions of decision-making competence were identified: (1) "Understanding values and norms" in the context of Sustainable Development (SD), (2) "Developing solutions", and (3) "Evaluating solutions" for SD problems. The two last-mentioned dimensions stem from decision-making theory, and were adapted to educational purposes. Related measurement instruments were developed according to Wilson's developmental cycle, using a between-item-multidimensionality approach. The test development procedures and results are described for the dimension "Developing solutions". Moreover, we started with an experimental validation of a theory of socioscientific decision making. More specifically, we used training-induced strategies to realize experimental variation to differentiate empirically between two decision-making dimensions and problem solving. The results of a pilot study addressing the validation of "Developing solutions" and "Evaluating solutions", vis-à-vis problem solving, are reported and discussed. We close with considerations of future research, to realign the boundaries of our research program.

Keywords Socioscientific decision making • Sustainable Development • Competence modeling • Problem solving • Experimental validation

S. Bögeholz (✉) • S. Eggert • C. Ziese
Georg-August-Universität Göttingen, Göttingen, Germany
e-mail: sboegeh@gwdg.de; seggert1@gwdg.de; cziese@gwdg.de

M. Hasselhorn
German Institute for International Educational Research (DIPF), Frankfurt/Main, Germany
e-mail: hasselhorn@dipf.de

16.1 Introduction

Worldwide biodiversity loss and climate change are challenging problems with respect to Sustainable Development (SD). These problems are tightly linked to political, economic, and societal concerns (Oulton et al. 2004). In the field of science education they are subsumed under the term socioscientific issues (e.g., Sadler et al. 2007). Typically, these issues are factually and ethically complex, ill-structured, subject to ongoing inquiry, and they lack an optimal solution (Bögeholz and Barkmann 2005; Ratcliffe and Grace 2003; Sadler et al. 2007). Rather, multiple solutions exist, all of which have their drawbacks. With respect to solving SD problems, decision-making competence is crucial to promote "technically and economically viable, environmentally sound, and morally just solutions" (Bögeholz et al. 2014, p. 237), and to foster student literacy as citizens (Ratcliffe and Grace 2003; Sadler et al. 2007).

Working with SD problems in the science classroom poses high processing demands on students (Eggert et al. 2013). Students do not only have to rely on a profound (scientific) knowledge base but also have to engage in various information search, argumentation, reasoning, and decision-making processes (Eggert et al. 2013; Jiménez-Aleixandre and Pereiro-Muñoz 2002; Ratcliffe and Grace 2003).

Socioscientific decision making was implemented in German science curricula (e.g., KMK 2005) as one reaction to German students' mediocre results in the PISA (Programme for International Student Assessment) studies. As one consequence, German educational authorities emphasized competence-oriented teaching (KMK 2005). In a similar vein, the priority program "Competence Models" was launched to overcome the lack of empirical support for basic assumptions of the competence approach.

According to Weinert (2001), the concept of competence is strongly linked to problem solving. It takes into account a "sufficient degree of complexity […] to meet demands and tasks", and includes "cognitive and (in many cases) motivational, ethical, volitional, and/or social components" (Weinert 2001, p. 62) in solving problems successfully. Referring to this definition, Klieme et al. (2008, p. 9) emphasize the cognitive facet and define competencies "as context-specific cognitive dispositions that are acquired by learning and needed to successfully cope with certain situations or tasks in specific domains". This definition was adopted for the present research on decision-making competencies with regard to the challenging issues of SD.

16.2 A Competence Model for Decision Making with Respect to Sustainable Development

Research on socioscientific reasoning and decision making as well as on argumentation in the area of science education draws on different theoretical models such as Toulmin's argumentation model (Toulmin 1958), Kuhn's developmental model of
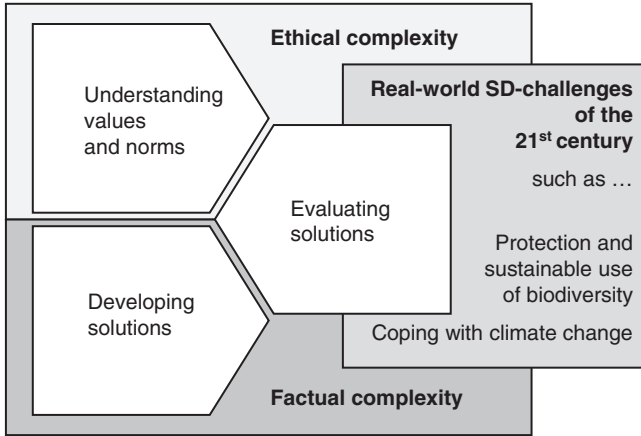
Fig. 16.1 Competence model for decision making with respect to challenging issues of Sustainable Development (SD)

critical thinking (Kuhn 1999), and models from descriptive decision theory (e.g., Betsch and Haberstroh 2005). All models highlight the need to compare and evaluate available options (i.e., solutions) by developing pro- and contra-arguments, and weighing these arguments or decision criteria in order to reach informed decisions (e.g., Eggert and Bögeholz 2010; Jiménez-Aleixandre and Pereiro-Muñoz 2002; Papadouris 2012; Ratcliffe and Grace 2003; Sadler et al. 2007). Being able to reach informed decisions is emphasized as a core competence in Education for Sustainable Development (ESD) as well as in citizenship education (Bögeholz and Barkmann 2005; Sadler et al. 2007). The competence model used in the present project is based on SD-related research as well as on a meta-model from descriptive decision theory (see Betsch and Haberstroh 2005; Bögeholz et al. 2014), and was adapted for educational purposes (Eggert and Bögeholz 2006; Bögeholz 2011). The model comprises three dimensions (see Fig. 16.1).

"Understanding values and norms": While working with SD problems, students need to consider and reflect on crucial normative guidelines, such as basic need orientation, intergenerational justice, international justice and simultaneous consideration of ecological, economic, and social objectives. This requires an understanding of the necessity of fulfilling human needs through a sustainable use of natural resources, and that satisfying needs in a sustainable manner eventually contributes to human well-being (MA 2005; cf. Bögeholz et al. 2014).

"Developing solutions": Students need to be able to comprehend and to describe multifaceted and complex SD problems, and to develop possible sustainable solutions. This implies taking into account various stakeholder perspectives with different ecological, economic, and social objectives. In addition, this dimension also includes the ability to reflect on developed solutions and the evidence that these solutions are based on (e.g., Gausmann et al. 2010).

"Evaluating solutions": Students need to be able to compare and evaluate multiple possible solutions to a SD problem. This includes the ability to develop pro- and contra-arguments, and to weigh these arguments by making use of trade-offs and/or cut-offs to reach informed decisions. In addition, the dimension comprises the ability to reflect on and to monitor decision-making processes (Bernholt et al. 2012; Eggert and Bögeholz 2010; Eggert et al. 2010).

16.3   Measurement Instruments and Competence Modeling

All measurement instruments were developed on the basis of Wilson's developmental cycle (Wilson 2005), using a between-item-multidimensionality approach (Wu et al. 2007). With respect to the measurement instrument for "Evaluating solutions", the procedure and results are described in Eggert and Bögeholz (2010, 2014). In this Sec. 16.3., we focus on the measurement instrument for "Developing solutions". Both measures are used in Sec. 16.4. as dependent variables in a training study designed to examine the relationship between decision making and problem solving.

With respect to "Developing solutions", we assumed that the postulated unidimensionality could be empirically supported. Second, we assumed that items representing the description of a problem situation would be easiest, while items representing the development of solutions to SD problems should be of medium item difficulty. Finally, items representing a reflection of presented solutions were assumed to have the highest difficulty.

16.3.1   "Developing Solutions": Development of the Measurement Instrument

16.3.1.1   Sample

678 students were analyzed in two subsamples of eighth to ninth graders and tenth to twelfth graders. The subsample of eighth to ninth graders consisted of 319 students (157 females, 162 males; mean age: 14.32, $SD = 0.68$), and the subsample of tenth to twelfth graders consisted of 359 students (187 females, 172 males; mean age: 16.76, $SD = 0.90$). All students attended the German *Gymnasium*, which is the academic track that prepares students for studies in higher education.

16.3.1.2   Measures: Tasks and Items

To measure student competencies with respect to the dimension "Developing solutions", a questionnaire with open-ended as well as multiple-choice items was developed. Based on an extensive literature and curriculum review, preliminary test tasks

and items were developed, pre-piloted using think-aloud protocols, and optimized. Several complementary quantitative studies followed.

The contexts used in the questionnaire were overfishing of tuna in the South Pacific ("Tuna task"), soy production in the Paraguayan rainforest ("Soy task"), and the collection of hoodia plants in Africa for pharmaceuticals ("Hoodia task"). All these contexts are typical SD problems, also described as socio-ecological dilemmas (e.g., Ernst 1997).

With respect to the Soy task, for example, there is a growing worldwide demand for soy in meat production (economic aspect). This demand is met by installing more and more soy plantations in rainforest areas. As a consequence, rainforest areas decrease (ecological aspect). However, several social groups, such as local people, who depend on the rainforest as a resource (social aspect 1), are affected by rainforest conversion. Instead, soy plantation workers earn their living on the plantations (social aspect 2). Consequently, the soy industry influences the living conditions of the local farmers. In the long run, all involved social groups suffer from exploitation of the rainforest. In addition, institutions like governments and NGOs, but also consumers, play an important role in relation to such dilemmas.

With respect to the Tuna task and the Soy task, students were asked to describe the problem situation first, and then to develop a sustainable solution to the problem. With respect to the Hoodia task, students were given potential solutions to the SD problem, asked to reflect on these solutions in terms of their sustainability (Evaluate in Table 16.1), and to give suggestions for improvement (Improve in Table 16.1) to these solutions.

Student responses to the open-ended questions were analyzed with respect to the interrelated aspects of the socio-ecological dilemma (economic, ecological, and social aspects; see description above) as well as the institutions and consumers that influence the SD problem or may facilitate sustainable solutions (see Table 16.1).

In sum, eight items were used to analyze student answers to the description of the Tuna task and the Soy task (items 1–8 and 16–23). Seven items were used to analyze student answers on the development of solutions to each of these problems (items 9–15 and 24–30). Finally, for eighth to ninth graders, six items were used to analyze student answers to the Hoodia task with respect to the evaluation of Project A (items 31–36). For the older students (tenth to twelfth graders), the Hoodia task "Improve project B" was additionally used to depict student competencies at the upper end of the competency scale (items 37–42 added to items 31–36).

16.3.1.3 Instrument Functioning

Preliminary analyses showed that it is more appropriate to analyze eighth to ninth graders and tenth to twelfth graders separately, as several items exhibited medium to large differential item functioning (DIF) with respect to these two subsamples. Specifically, several items got disproportionally easier among the tenth to twelfth graders. Thus, in the following analyses, we analyzed both subsamples separately, using the unidimensional Rasch model (Rasch 1960). Item fit values as well as

Table 16.1 Tasks, items and item estimates for "Developing solutions" for eighth–ninth graders and tenth–twelfth graders, and their reliability indices

	Item no.	Item descriptions	Eighth to ninth graders Item estimates	Tenth to twelfth graders Item estimates
Tuna describe problem/Soy describe problem	1/16	Ecological-economic relation [R1]	−1.19/−1.69	−1.37/−2.17
	2/17	Social1-ecological relation [R2]	−2.85/−1.13	−3.47/−1.73
	3/18	Social1-social2 relation [R3]	−2.62/0.32	−3.02/0.36
	4/19	Social2-economic relation [R4]	−0.96/0.60	−1.21/0.28
	5/20	Social1-economic relation [R5]	−1.26/−0.92	−1.29/−1.44
	6/21	Social2-ecological relation [R6]	−2.50/0.11	−3.38/0.08
	7/22	Role of institutions [I1]	−0.44/0.21	−0.72/0.14
	8/23	Role of consumers [C2]	−0.22/−1.56	−0.17/−1.66
Tuna develop solution/Soy develop solution	9/24	Ecological-economic relation [R1]	1.43/0.34	0.90/0.45
	10/25	Social1-ecological relation [R2]	−0.41/0.55	−0.26/0.53
	11/26	Social1-social2 relation [R3]	−0.06/1.76	−0.02/1.59
	12/27	Social2-economic relation [R4]	1.76/2.38	1.42/1.98
	13/28	Social1-economic relation [R5]	1.87/1.56	1.26/1.38
	14/29	Social2-ecological relation [R6]	−0.62/1.20	−0.68/1.08
	15/30	Role of institutions [I1]	−1.99/−0.15	−2.70/−0.09
Hoodia reflect (Evaluate) project A	31	Ecological-economic relation [R1]	2.14	2.19
	32	Social1-ecological relation [R2]	−0.08	−0.79
	33	Social1-social2 relation [R3]	−0.08	−0.59
	34	Social2-economic relation [R4]	2.30	2.31
	35	Social1-economic relation [R5]	1.71	1.74
	36	Social2-ecological relation [R6]	0.50	−0.02

(continued)

Table 16.1 (continued)

	Item no.	Item descriptions	Eighth to ninth graders	Tenth to twelfth graders
			Item estimates	Item estimates
Hoodia reflect (Improve) project B	37	Ecological-economic relation [R1]	–	2.44
	38	Social1-ecological relation [R2]	–	0.02
	39	Social1-social2 relation [R3]	–	0.25
	40	Social2-economic relation [R4]	–	2.78
	41	Social1-economic relation [R5]	–	2.61
	42	Social2-ecological relation [R6]	–	0.99
	Reliability indices			
	Item separation reliability		.99	.99
	WLE-Person separation reliability		.83	.87
	EAP/PV reliability		.75	.74
	Cronbach's alpha		.82	.85

traditional item discrimination values were analyzed. Items with discrimination values lower than .20 and weighted mean square (WMNSQ) values that were not within the range of 0.75 and 1.33 were eliminated (Wilson 2005). After deletion of non-functioning items, the final measurement instrument for eighth to ninth graders consisted of 36 items, and the instrument for tenth to twelfth graders consisted of 42 items respectively. Table 16.1 provides an overview of all final items, their item estimates and reliability indices, with respect to both subsamples.

16.3.2 Modeling of "Developing Solutions"

To investigate our assumptions with respect to a possible progression of item difficulty by task complexity, we classified the items into three different categories: "describing" (1), "developing" (2), and "reflecting" (3). With respect to eighth to ninth graders, average item difficulty for all "describing" items was −.96 logits, while item difficulty for all "developing" items was considerably higher (.53 logits). Average item difficulty for all "reflecting" items was highest, with 1.08 logits. An analysis of variance (ANOVA) of item difficulty, grouping items by item complexity, supports this assumption ($f(2, 33) = 8.69$, $p = .001$, $\eta2 = .35$). Post hoc Tukey tests revealed that "reflecting" items and "developing" items were harder than "describing" items ($p < .01$), while no significant difference could be found between "developing" items and "reflecting" items.

With respect to tenth to twelfth graders, average item difficulty for all "describing" items was -1.24 logits; for all "developing" items it was higher at .31 logits. Average item difficulty for all "reflecting" items was highest at 1.16 logits. In accordance with our assumptions, an ANOVA was again statistically significant ($f(2, 39)$ = 11.65, $p < .001$, $\eta^2 = .38$). Post hoc Tukey tests revealed that again "describing" items were easiest ($p < .01$), while the difference between "developing" items and "reflecting" items was again not significant. The Wright map for tenth to twelfth graders is depicted in Fig. 16.2.

In addition, we analyzed the influence of the different contexts on item difficulty. An ANOVA showed no significant differences between the Tuna task and the Soy task.

Analyzing the validity of the dimension "Developing solutions" for both groups, we found no relations with reading speed and reading comprehension ($p > .05$) or with different subject grades. Finally, no relation was found with strategy knowledge for solving problems (for measurement instrument see Scherer 2012).
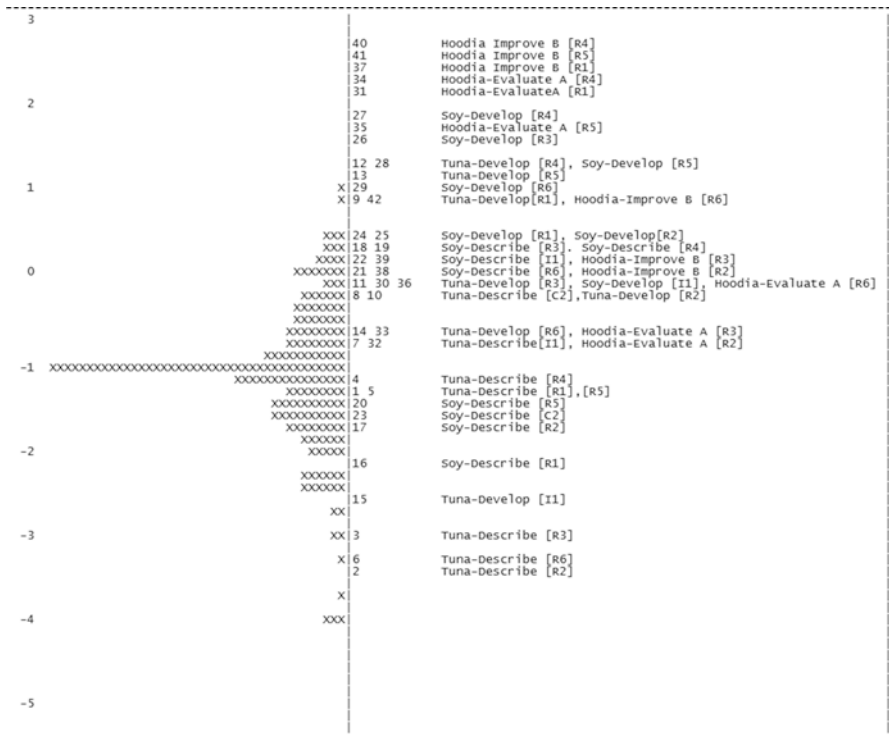

Fig. 16.2 Wright map for "Developing solutions" for tenth to twelfth graders
(*R* relation, *I* institutions, *C* consumers, *x* 2.2 cases)

16.3.3 Discussion

The purpose in this phase of the priority program was to develop new measurement instruments that could be used for analyzing student decision-making competencies with respect to complex issues of sustainable development. Analysis of item fit statistics as well as analysis of traditional item indices revealed that the instrument fits the requirements of the Rasch model. DIF analysis also showed that analyses should be conducted separately for eighth–ninth, and for tenth–twelfth graders. In addition, some items should be used specifically for measuring student competencies at the upper end of the competency continuum. In addition, we were able to show that the developed items can successfully differentiate between different cognitive processes (describing, developing, and reflecting).

Moreover, we could show that "Developing solutions", as part of socioscientific decision making, differs from reading comprehension, and from strategy knowledge in solving problems. This is quite important, as all items used in the measurement instrument ask students to read an information booklet on SD problems, to perform the tasks given, and to find solutions to SD problems.

16.4 Experimental Validation: A Comparison of Socioscientific Decision Making with Analytical Problem Solving

In the following we introduce a training-based experimental validation approach (cf. Mummendey and Grau 2008, p. 106) for the decision making part of our theoretical contribution. We argue that analytical problem solving is a good candidate for validation purposes, for studying decision making within an intervention study.

Even though decision making and problem solving are concepts from different theoretical research branches (Betsch and Haberstroh 2005; Pólya 1945; cf. Leutner et al. 2004), both refer to processes that deal with complex real-world problems. These processes include identifying the (decision-making) problem, identifying relevant information, developing solutions (solution paths), selecting solutions (solution strategies), solving the problem, and reflecting on the solution (Eggert and Bögeholz 2006; OECD 2004, p. 16).

However, decision making and problem solving differ from each other in some aspects. Problem-solving tasks primarily require one correct solution, even if, theoretically speaking, there should be different solution paths. In contrast, decision making focuses on argumentation and reasoning while taking a decision; consequently, there might be several legitimate decisions.

Problem solving as conceptualized in PISA 2003 covers the "overall capability to solve problems in real-life situations beyond the specific context of school subject areas" (OECD 2004, p. 16). In contrast, the relationships between "Developing solutions", "Evaluating solutions", and problem solving are not yet completely

understood. Moreover, analytical problem solving seems to be a good candidate for validating the new concept of socioscientific decision making because of its structural similarity. In analytical problem-solving tasks, all information is given simultaneously or can be inferred, and individual competence is measured via paper-pencil tests. Both features are parallel to the assessment of decision making (e.g., Eggert and Bögeholz 2010). This allows us to concentrate on comparison of the two constructs, instead of dealing with changing conditions during problem solving, and divergent computer-based assessment, which are features of dynamic problem solving (cf. Leutner et al. 2004).

16.4.1   Objectives and Research Design

The purpose of the validation study was to analyze whether "context-specific" decision making—with its two dimensions—can be empirically differentiated from problem solving as a "cross-curricular" competence (cf. OECD 2004). We conducted a pre-posttest control group training study. The design included three training groups, each of which focused on specific processes of decision making or problem solving: Training Group 1 was trained in "Developing solutions" (TG1), Training Group 2 in "Evaluating solutions" (TG2), and Training Group 3 in "Problem solving" (TG3). In addition, a control group (CG) was tested, in which students attended regular biology courses without any explicit training in decision making or problem solving. However, the CG studied the same content as TG1, TG2, and TG3 (see below). The following hypotheses (dependent variables mentioned in first place) were derived:

- *"Developing solutions":* Students of Training Group 1, "Developing solutions", outperform students of all other groups (TG1 > TG2, TG3, CG).
- *"Evaluating solutions":* Students of Training Group 2, "Evaluating solutions", outperform students of the remaining groups (TG2 > TG1, TG3, CG).
- *"Problem solving":* Students of the "Problem solving" group (TG3) outperform students of the remaining groups (TG3 > TG1, TG2, CG).

16.4.2   Methods

In the pre- and posttest, paper-and-pencil tests for "Developing solutions" (see Sect. 16.3.1.2.), "Evaluating solutions" (cf. Eggert et al. 2010), and "Problem solving" (cf. OECD 2004) were used. Testing time for the pretests was 120 minutes, and for the posttests 90 minutes.

16.4.2.1   Participating Students and Teachers

Participants included four eighth grade classes from one high school in North Rhine-Westphalia, Germany (63 females and 54 males; mean age: 13.59, *SD* = 0.62). The study was conducted from January to March 2014, and supported by the school's vice-director. Three biology teachers participating in the study had no specific prior training in socioscientific decision making or problem solving.

All three teachers received introductory one-to-one coaching with respect to their specific treatment condition. The teaching units for TG1, TG2, and TG3 were developed by the researchers. The teaching approach, the materials and the methods of the corresponding teaching unit were discussed during the coaching sessions.

Teacher A (4 years teaching experience) taught TG1 (*n* = 28 students) and TG3 (*n* = 26). This teacher had a weak commitment to the study, was challenged in having to teach two different training groups, and underestimated student abilities with respect to the content of the teaching units. In addition, he was more used to teacher-centered instruction and had a more transmissive orientation towards teaching and learning. He spent the least amount of time in preparing for and reflecting on his teaching.

Teacher B (33 years teaching experience; vice-director) taught TG2 (*n* = 28), while teacher C (8 years teaching experience) taught in the control group and, thus, followed his own teaching approach (CG; *n* = 28). Teacher C used materials provided by the researchers but was free to restructure it, searched with enthusiasm for additional information, and developed the material to his own needs. Teachers B and C were highly committed to our study; they were self-confident and showed high identification with their teaching units.

All lessons were documented by a researcher who wrote a chronological protocol (Böhmann and Schäfer-Munro 2008). Observations revealed that students in the TG3 and the CG were interested in the teaching units and actively participated in the course. In contrast, students of the TG1 and the TG2 were more heterogeneous in terms of interest and motivation to participate.

16.4.2.2   Trainings and Learning Material

All trainings (TG1–3) and the regular CG instructions comprised 6 teaching units of 45 miuntes each, and taught in 90 minutes double periods. In the first two double periods students worked on palm oil production in Indonesian rainforest areas. In the final double period they worked on cotton production in Uzbekistan, and its consequences for the drying Aral Sea (see Table 16.2). All four conditions used cooperative learning methods such as gallery walk, jigsaw puzzle, fishbowl, and pair/team discussions. The three treatment conditions only differed with respect to the teaching of specific strategies for socioscientific decision making and problem solving.

Students in TG1 ("Developing solutions") focused on the analytical and comprehensive description of the SD problems, as well as development of solutions and

Table 16.2 Unit objectives for the three training groups of the experimental validation study (90 minutes: a double period)

		"Developing solutions"	"Evaluating solutions"	"Problem solving"
		Students	Students	Students
Palm oil production	90 min.	… understand the problem of palm oil production and their role as consumers by discovering palm oil substances in everyday products	… understand the problem in its factual complexity by considering the ecological, economic, and social aspects	…explore the SD problem associated with palm oil from Indonesia
		… apply an analytical framework to understand and describe the factual complexity of the problem as well as possible solutions	… apply a decision matrix to collect and collate necessary information for three given, real-world solutions	… understand the provided analytical problem-solving framework and perform step 1: "read and understand" to cope with the factual complexity
	90 min.	… develop solutions to the problem that integrate different stakeholder perspectives	… use the decision matrix to evaluate the three given solutions and their underlying value considerations, applying different decision-making strategies	… use problem-solving strategies to develop solutions by considering different stakeholder perspectives and perform steps 2-5: develop a plan, choose a plan, apply it and evaluate the solutions
		… use the analytical framework to reflect on the developed solutions and on one specific given real-world solution	… use the decision matrix to identify and reflect on the factual and ethical complexity in their own decision processes and decision processes of others	… perform the problem- solving steps to reflect on their own solution from a certain stakeholder perspective and on given solutions
Cotton production (Transfer)	90 min.	… use the analytical framework to develop solutions to the problem, acknowledging the factual complexity	… use the decision matrix to evaluate solutions to the problem, acknowledging the factual and ethical complexity	… use the steps of the analytical problem-solving framework to develop solutions to the problem

reflection on solutions. To help students understand the complex relations between the different ecological, economic and social aspects of the SD issue, the teacher used a specific analytical framework (see Bögeholz 2011; Gausmann et al. 2010; Ostermeyer et al. 2012; see Table 16.2).

Students in TG2 concentrated on the comparison of different, equally legitimate solutions to solve the presented SD problems. This also included the development of pro- and contra-arguments and the weighing of arguments or decision criteria in order to reach informed decisions. To help students compare the different possible options and their criteria in a systematic manner, a decision matrix was used. This decision matrix was also used to make value decisions transparent, and therefore allowed for discussing, reflecting on, and respecting different (legitimate) solutions and decision-making processes (e.g., Bögeholz 2006; Eggert and Bögeholz 2006; see Table 16.2).

TG3 worked on the presented SD problems by following the problem-solving steps (Buchwald et al. 2017, in this volume; see Table 16.2). While the students worked on the problem-solving steps "developing problem solving ideas" and "choosing a problem solving plan", they got to know a set of six problem-solving strategies (see Blum et al. 2006, p. 39), namely: principle of analogy, principle of decomposition, principle of illustration, working forward, working backward, and systematic trying. Our training builds on experiences from Buchwald et al. 2017, in this volume).

16.4.2.3 Measures

Socioscientific decision making and analytical problem solving were assessed as dependent variables. With respect to socioscientific decision making, both measures were used in an abridged version. The pretest for "Developing solutions" consisted of three tasks: (1) Rattan from Indonesia (see Eggert et al. 2013), (2) Oil and gas extraction in Siberia ("describing" and "developing" items), and (3) Shrimps from South-East Asia ("reflecting" items; cf. Eggert et al. 2013; Table 16.3). The final scale included 24 items (α = .75). With respect to "developing" solutions, the scoring procedure was altered (comparing Table 16.3 with Table 16.1). Within the new scoring each single aspect (see [A] in Table 16.3) was scored instead of related aspects (see [R] in Table 16.1). The new scoring better aligns with student responses, due to the degree of item complexity. Compared to our measure in Table 16.1, we presented only one project per reflection task, and we reduced the number of items as a consequence of limited testing time.

The corresponding posttest integrated the Rattan and Soy tasks (see Table 16.1) with "describing" items, and "developing" items, as well as the Hoodia task (see Table 16.1) with "reflecting" items. The final scale included 24 items (α = .74). All items for "describing" the SD problems, "developing" solutions and "reflecting" solutions were dichotomous.

The pretest for "Evaluating solutions" again comprised three different tasks: (1) the problem of cabbage white butterfly larvae in vegetable gardens, (2) a problem-

Table 16.3 Abridged measure of "Developing solutions"

		Item descriptions	Item no. Pre- and posttest
Rattan/Oil and gas (t1); Rattan/Soy (t2)	Describe problem	Ecological-economic relation [R1]	1/10
		Social1-ecological relation [R2]	2/11
		Social1-social2 relation [R3]	3/12
		Social2-economic relation [R4]	4/13
		Social1-economic relation [R5]	5/14
		Social2-ecological relation [R6]	6/15
	Develop solution	Economical aspect [A1]	−/16
		Ecological aspect [A2]	7/17
		Social2 aspect [A3]	8/18
		Social1 aspect [A4]	−/19
		Institution aspect [I2]	9/20
Shrimps (t1); Hoodia (t2)	Reflect project A	Ecological-economic relation [R1]	21
		Social1-ecological relation [R2]	22
		Social1-economic relation [R5]	23
		Social2-ecological relation [R6]	24

atic neophyte for riverbanks (both decision tasks), and (3) a reflection task on the means of transportation for holidays. The final scale included 11 items ($\alpha = .78$). In the posttest, we used the Neophyte task again and varied the other tasks ("overfishing of codfish", and a consumer choice task; cf. Eggert et al. 2010). The final scale included 11 items ($\alpha = .88$).

To assess analytical problem solving we applied items from PISA 2003. Thereby, we used a selected set of items and the corresponding scoring guide provided by a collaborating working group (Buchwald et al. 2017, in this volume). Specifically, we analyzed problem solving via a scale of three dichotomous items (cinema 1, watergate, design) as well as three trichotomous items (train, holiday camp, vacation). For the pretest ($\alpha = .51$) and posttest ($\alpha = .52$) we used the identical problem-solving tasks.

For all three measures, half of the items were double coded (Cohen's Kappa: .93–.99). As expected, validation analyses revealed very weak to weak correlations between "Developing solutions" and "Problem solving" ($r = .31$, $p < .01$), between "Evaluating solutions" and "Problem solving" ($r = .27$, $p < .01$) and also between "Developing solutions" and "Evaluating solutions" ($r = .20$, $p < .05$).

16.4.3 Results of the Pilot Study

As a first step, we conducted one-way ANOVAs to check for possible group differences on the pretest scores. Post hoc Tukey tests showed significant differences between the four treatment conditions with respect to all three dependent variables: "Developing solutions", "Evaluating solutions", and "Problem solving". The training group "Evaluating solutions" (TG2) always displayed the lowest test performances (except for the measure on "Problem solving"), and differed from the Control Group in always having the best test performances ($p < .05$).

As a consequence of the identified pretest differences, we conducted multiple regression analyses using the pretest scores (prior knowledge) and the treatment conditions as independent variables. Concerning treatment conditions, contrasts were coded. The mean and standard deviations of the dependent variables by time and treatment are displayed in Table 16.4.

"Developing solutions" at posttest were predicted by prior knowledge as well as by both contrasts (see Table 16.5). The final statistical model accounts for 30 % of the variance with prior knowledge accounting for 14 %, the second contrast variable for 6 %, and the third contrast variable for 10 %. Remarkably, the third contrast reveals a negative relationship with posttest learning outcomes, that is TG3 outperforms TG1.

For "Evaluating solutions", prior knowledge and the first contrast variable predict students' learning outcomes in the posttest (see Table 16.5). The final statistical model accounts for 40 % of the variance, with prior knowledge accounting for

32 % and the contrast variable accounting for 8 %. The analyses reveal that the CG shows better posttest performances than the training groups.

Table 16.4 Mean scores and standard deviations for "Developing solutions", "Evaluating solutions", and "Problem solving" by time and treatment (TG: training group; CG: control group)

		TG1	TG2	TG3	CG
"Developing solutions"					
Pretest	*M* (*SD*)	13.95 (4.47)	10.52 (5.57)	13.00 (4.65)	15.81 (4.30)
Posttest	*M* (*SD*)	12.81 (5.39)	11.04 (5.24)	17.22 (4.03)	14.93 (3.52)
"Evaluating solutions"					
Pretest	*M* (*SD*)	12.10 (3.89)	9.72 (4.23)	14.04 (3.65)	14.11 (2.33)
Posttest	*M* (*SD*)	7.00 (5.29)	5.52 (4.94)	10.16 (5.60)	12.74 (3.05)
"Problem solving"					
Pretest	*M* (*SD*)	3.81 (1.75)	3.93 (1.69)	4.22 (2.17)	5.89 (1.93)
Posttest	*M* (*SD*)	3.81 (2.25)	3.59 (2.48)	4.87 (2.16)	6.11 (1.83)

Table 16.5 Multiple regression predicting posttest performance on "Developing solutions" and "Evaluating solutions" by prior knowledge and treatment

"Developing solutions"			
	B	*SE*	β
Step 1			
Prior knowledge (pretest score)	0.37	.09	.37***
Step 2			
Prior knowledge (pretest score)	0.31	.09	.31**
Contrast 2 (TG1, TG3 vs. TG2)	1.06	.39	.26**
Step 3			
Prior knowledge (pretest score)	0.33	.09	.33***
Contrast 2 (TG1, TG3 vs. TG2)	1.01	.36	.25**
Contrast 3 (TG1 vs. TG3)	−2.36	.65	−.31***
Note: $R^2 = .14$ for Step 1, $\Delta R^2 = .06$ for Step 2, $\Delta R^2 = .10$ for Step 3, ** $p < .01$, *** $p < .001$			
"Evaluating solutions"			
	B	SE	β
Step 1			
Prior knowledge (pretest score)	0.79	.12	.57***
Step 2			
Prior knowledge (pretest scores)	0.69	.11	.50***
Contrast 1 (TG2, TG1, TG3 vs. CG)	−0.91	.25	−.30***
Note: $R^2 = .32$ for Step 1, $\Delta R^2 = .08$ for Step 2, *** $p < .001$			

In addition, "Problem solving" was revealed to be exclusively predicted by prior knowledge, which accounts for 33 % of the variance (Prior knowledge [pretest score]: $B = 0.66$, $SE = .10$, β = .57, $p < .001$). Thus, the investigated contrast variables did not contribute to explaining the variance of the posttest scores.

16.4.4 Discussion

The aims of this pilot study were (1) to further improve the training procedures, (2) to further develop abridged measures, and (3) to initiate a training-based experimental validation of our approach in conceptualizing socioscientific decision making. A number of crucial factors have to be taken into account:

Educational and experimental setting: Even though the school administration showed an extraordinary commitment, our study was affected by the perils of field research. Specifically, our study was influenced, for example, by differences in teacher enthusiasm, different amounts of time spent in preparing teaching, and reflecting on the lessons taught. Working with just one school eased project management demands and tended to ensure a socially more homogeneous student population. Teachers were recruited by the school administration in a top-down approach. The school administration created special timetables so that all classes had an inter-

vention of three double periods. However, the latter had the side effect that TG1 had to sacrifice their physical education lessons for the posttest, and a considerable decline in interest for TG1 was documented in the chronological protocols. With respect to the main study, we will follow a more bottom-up approach for recruiting teachers.

Measures: All SD problems addressed in the measurement instruments differ from the SD problems addressed in the treatments. The instruments applied for decision making were used in abridged versions. The abridged version of "Developing solutions" still covers all crucial features of our competence dimension. With respect to the abridged version of "Evaluating solutions", here again, all core characteristics of the competence dimension are considered in the measure (cf. Eggert and Bögeholz 2010). In sum, the reliabilities of our decision-making measures are promising and we succeeded in having widely varying pre- and posttest measures. However, it still remains a challenge (1) to model "Developing solutions" with polytomous items, and (2) to analyze the pre- and posttest design with IRT (cf. procedure in Eggert et al. 2010 for "Evaluating solutions").

Training outcomes: With respect to the dependent variable "Developing solutions", students of TG3 benefited from the well-designed teaching material with challenging tasks as well as from the participative teaching methods (e.g., fishbowl). In contrast, TG1—even though they had the same teacher—did benefit less. This might at least partly be due to the fact that the students had to cope with the disappointment that they missed their physical education lessons in favor of the posttest. With respect to "Evaluating solutions", the students of the CG performed best. The latter can partly be traced back to the enthusiasm of teacher C (cf. Kunter 2011). Teacher C used a constructivist approach of teaching, which might have produced higher levels of motivation and performance among the students of the CG compared to the students of the training groups. This can be explained by research on teacher beliefs and their impact on learning outcomes (constructivist beliefs > transmissive beliefs see Voss et al. 2011, p. 250). Teacher beliefs might change with teaching experience over the career span, for example, a portion of experienced teachers overcame their teacher-centered metaphors and proceeded with student-centered metaphors (Alger 2009). The three teachers in our pilot study varied strongly in their teaching experiences (4–33 years). In the main study, the (average) teacher experience of the different treatment groups will be more balanced.

Beyond these explanations, more general phenomena might also have influenced the results: The acquisition of complex strategies is accompanied by a stage of so-called inefficient utilization ("*Nutzungsineffizienz*"; see Hasselhorn and Gold 2013, p. 100). If students are confronted with a new, complex strategy, a hugh additional strain is placed on their working memory. As a result, learning outcomes may be worse after a training than before. Lower achievement at posttest measures can also be traced back to a "motivational valley", and can be overcome by strategy automating (Hasselhorn and Gold 2013, pp. 100, 101). The latter may finally result in higher learning outcomes in time-delayed measuring.

In sum, we could successfully advance the training procedures as well as their corresponding measures. We are in a good position now to optimize the realization

of our main study. Even though we did not obtain much support for the hypotheses, the results can be plausibly explained by the circumstances, while validation of our theoretical contribution still remains a challenging endeavor.

16.5 Conclusions and Outlook

Though our research program on decision making regarding SD issues is far from being finalized, we provide several measures that stem from the Göttingen competence model. All in all, they allow for the adequate assessment of student competencies with respect to socioscientific decision making (Eggert et al. 2010, 2013; Gresch et al. 2013; Sakschewski et al. 2014). In addition, our approach has already inspired other working groups within the research community (Böttcher and Meisert 2013; Heitmann 2012; Hostenbach et al. 2011; Papadouris 2012).

To finish our experimental validation approach, we are currently conducting our main study, which includes six classes for each of the three training groups and the control group. The participating schools were recruited from four German federal states, and the composition of the treatment conditions (e.g., with respect to teacher experiences, teacher beliefs, teacher enthusiasm, student motivation, students social backgrounds) was as balanced as possible. To better cope with any potential "motivational valley" in acquiring complex strategies, we are carrying out the posttest of the main study six to eight weeks after the trainings.

The present contribution refers to an instructional approach to test whether the theoretical assumptions of the socioscientific decision-making theory addressed in our project are valid. The approach has been used in several fields of psychological research in order to test assumptions whether specific processes or strategies are responsible for the quality of specific individual behavior. It is the idea of the instructional approach to manipulate the relevant strategies and see whether the instruction has an effect on the target behavior. To be clear, although our project started to validate the socioscientific decision-making theory by means of an instructional approach, much further research seems to be necessary in order to come to a final assessment of the validity of the theory.

Beyond the above-mentioned open questions related to the validity of the theory, upcoming research on student competencies should go mainly into three directions: First of all, the model should be elaborated in more detail, since the evaluation of SD-problem solutions additionally requires considering quantitative impacts. Here, decision making profits from the use of simplified methods of economic validation, such as cost benefit analysis, cost effective analysis or profitability analysis. Thus, mathematical-economic modeling will complement the current research. A promising fourth competence dimension, "Evaluating solutions quantitatively-economically", is described in Bögeholz et al. (2014) as well as in Bögeholz and Barkmann (2014).

Second, the developed measures on "Evaluating solutions" of decision-making competence have been successfully applied to analyzing gains in learning outcomes in a pre- and posttest study via IRT-modeling (Eggert et al. 2010). For the current main study aiming at experimental validation, the abridged measure for "Developing solutions" has to be further strengthened so that decision making can be modeled with IRT in at least two measurement points, for both assessed decision-making dimensions. Besides having sensitive measures for decision making in intervention studies, we aim to further develop our measures for longitudinal studies with IRT-modeling, as well as for computer-based adaptive testing in the long run.

Third, our previous research has addressed the cognitive components of decision making. For the future, studies to foster decision making and studies on competence development should consider motivational factors. Studies on decision making with respect to biodiversity challenges should also integrate measures of interest in biodiversity issues. Because motivational factors impact learning outcomes (cf. Weinert 2001; cf. Rotgans and Schmidt 2011), linking research on motivation and cognitive competence is of practical relevance for real-world learning settings.

Beside these recent and future endeavors regarding student competencies, we aim at modeling and fostering teacher PCK for teaching socioscientific decision making. The latter benefits from the knowledge gained in the priority program—that is, knowledge on student decision-making competencies and on strategies to improve them.

In sum, our competence research on SD issues is a promising approach not only for ESD, but also for science teaching, and for citizenship education (e.g., Sadler et al. 2007; Eggert et al. 2013; Sakschewski et al. 2014; Bögeholz et al. 2014).

References

Alger, C. L. (2009). Secondary teachers' conceptual metaphors of teaching and learning: Changes over the career span. *Teaching and Teacher Education, 25*, 743–751. doi:10.1016/j.tate.2008.10.004.

Bernholt, S., Eggert, S., & Kulgemeyer, C. (2012). Capturing the diversity of students' competences in the science classroom: Differences and commonalities of three complementary approaches. In S. Bernholt, K. Neumann, & P. Nentwig (Eds.), *Making it tangible: Learning outcomes in science education* (pp. 173–199). Münster: Waxmann.

Betsch, T., & Haberstroh, S. (Eds.). (2005). *The routines of decision making*. Mahwah: Erlbaum.

Blum, W., Drüke-Noe, C., Hartung, R., & Köller, O. (2006). Bildungsstandards Mathematik: Konkret. In *Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen [German educational standards in mathematics for lower secondary degree: Tasks, teaching suggestions, thoughts with respect to further education]*. Berlin: Cornelsen Scriptor.

Bögeholz, S. (2006). Explizit Bewerten und Urteilen: Beispielkontext Streuobstwiese [Explicit reasoning and decision making: The topic of orchards]. *Praxis der Naturwissenschaften, Biologie in der Schule, 55*, 17–24.

Bögeholz, S. (2011). Bewertungskompetenz im Kontext nachhaltiger Entwicklung: Ein Forschungsprogramm [Reasoning and decision-making competence in the context of sustainable development: A research program]. In D. Höttecke (Ed.), *Naturwissenschaftliche Bildung als Beitrag zur Gestaltung partizipativer Demokratie* (pp. 32–46). Münster: LIT.

Bögeholz, S., & Barkmann, J. (2005). Rational choice and beyond: Handlungsorientierende Kompetenzen für den Umgang mit faktischer und ethischer Komplexität [Rational choice and beyond: Action-oriented competencies for dealing with factual and ethical complexity]. In R. Klee, A. Sandmann, & H. Vogt (Eds.), *Lehr- und Lernforschung in der Biologiedidaktik* (pp. 211–224). Innsbruck: Studien-Verlag.

Bögeholz, S., & Barkmann, J. (2014). "… to help make decisions": A challenge to science education research in the 21st century. In I. Eilks, S. Markic, & B. Ralle (Eds.), *Science education research and education for sustainable development* (pp. 25–35). Aachen: Shaker.

Bögeholz, S., Böhm, M., Eggert, S., & Barkmann, J. (2014). Education for sustainable development in German science education: Past-Present-Future. *EURASIA Journal of Mathematics, Science & Technology Education, 10*, 219–236. doi:10.12973/eurasia.2014.1079a.

Böhmann, M., & Schäfer-Munro, R. (2008). *Kursbuch Schulpraktikum: Unterrichtspraxis und didaktisches Grundwissen [Course book teaching internship. Teaching practice and basic didactical knowledge]* (2nd ed.). Weinheim: Beltz.

Böttcher, F., & Meisert, A. (2013). Effects of direct and indirect instruction on fostering decision-making competence in socioscientific issues. *Science Education, 43*, 479–506. doi:10.1007/s11165-011-9271-0.

Buchwald, F., Fleischer, J., Rumann, S., Wirth, J., & Leutner, D. (2017). Training in components of problem-solving competence: An experimental study of aspects of the cognitive potential exploitation hypothesis. In D. Leutner, J. Fleischer, J. Grünkorn, & E. Klieme (Eds.), *Competence assessment in education: Research, models and instruments* (pp. 315–331). Berlin: Springer.

Eggert, S., & Bögeholz, S. (2006). Göttinger Modell der Bewertungskompetenz: Teilkompetenz "Bewerten, Entscheiden und Reflektieren" für Gestaltungsaufgaben nachhaltiger Entwicklung [Göttingen model for socioscientific decision making—Subdimension: "Evaluating solutions" for tasks concerning sustainable development]. *Zeitschrift für Didaktik der Naturwissenschaften, 12*, 177–199.

Eggert, S., & Bögeholz, S. (2010). Students' use of decision-making strategies with regard to socioscientific issues: An application of the Rasch partial credit model. *Science Education, 94*, 230–258. doi:10.1002/sce.20358.

Eggert, S., & Bögeholz, S. (2014). Entwicklung eines Testinstruments zur Messung von Schülerkompetenzen [Development of a measurement instrument for student competencies]. In D. Krüger, I. Parchmann, & H. Schecker (Eds.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (pp. 371–384). Berlin: Springer.

Eggert, S., Bögeholz, S., Watermann, R., & Hasselhorn, M. (2010). Förderung von Bewertungskompetenz im Biologieunterricht durch zusätzliche metakognitive Strukturierungshilfen beim kooperativen Lernen: Ein Beispiel für Veränderungsmessung [The effects of metacognitive instruction on students' socioscientific decision making: An exemplary procedure for measurement of change]. *Zeitschrift für Didaktik der Naturwissenschaften, 16*, 299–314.

Eggert, S., Ostermeyer, F., Hasselhorn, M., & Bögeholz, S. (2013). Socioscientific decision making in the science classroom: The effect of embedded metacognitive instructions on students' learning outcomes. *Education Research International.* doi:0.1155/2012/309894.

Ernst, A. (1997). *Ökologisch-soziale Dilemmata [Socio-ecological dilemmas]*. Weinheim: PVU.

Gausmann, E., Eggert, S., Hasselhorn, M., Watermann, R., & Bögeholz, S. (2010). Wie verarbeiten Schülerinnen und Schüler Sachinformationen in Problem- und Entscheidungssituationen

nachhaltiger Entwicklung [How do students process information in problem and decision-making situations]? *Zeitschrift für Pädagogik, Beiheft, 56*, 204–215.

Gresch, H., Hasselhorn, M., & Bögeholz, S. (2013). Training in decision-making strategies: An approach to enhance students' competence to deal with socio-scientific issues. *International Journal of Science Education, 35*, 2587–2607. doi:10.1080/09500693.2011.617789.

Hasselhorn, M., & Gold, A. (2013). *Pädagogische Psychologie: Erfolgreiches Lernen und Lehren [Pedagogical psychology: Successful learning and teaching]*. Stuttgart: Kohlhammer.

Heitmann, P. (2012). *Bewertungskompetenz im Rahmen naturwissenschaftlicher Problemlöseprozesse: Modellierung und Diagnose der Kompetenzen Bewertung und analytisches Problemlösen für das Fach Chemie* [Decision making in the context of problem solving regarding natural sciences: Modeling and diagnosis of decision-making and analytical problem-solving competencies with respect to chemistry]. Berlin: Logos.

Hostenbach, J., Fischer, H. E., Kauertz, A., Mayer, J., Sumfleth, E., & Walpuski, M. (2011). Modellierung der Bewertungskompetenz in den Naturwissenschaften zur Evaluation der nationalen Bildungsstandards [Modeling the evaluation and judgement of competence in science to evaluate national educational standards]. *Zeitschrift für Didaktik der Naturwissenschaften, 17*, 261–288.

Jiménez-Aleixandre, M.-P., & Pereiro-Muñoz, C. (2002). Knowledge producers or knowledge consumers? Argumentation and decision making about environmental management. *International Journal of Science Education, 24*, 1171–1190. doi:10.1080/09500690210134857.

Klieme, E., Hartig, J., & Rauch, D. (2008). The concept of competence in educational contexts. In J. Hartig, E. Klieme, & D. Leutner (Eds.), *Assessment of competencies in educational contexts* (pp. 3–22). Göttingen: Hogrefe.

KMK (Standing Conference of the Ministers of Education and Cultural Affairs of the States in the Federal Republic of Germany). (Ed.). (2005). *Bildungsstandards im Fach Biologie für den mittleren Schulabschluss (Jahrgangsstufe 10): Beschluss vom 16.12.2004* [German educational standards in biology for the lower secondary degree (Grade 10): Resolution approved by the standing conference on 16 December 2004]. München: Luchterhand.

Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher, 28*, 16–25. doi:10.3102/0013189X028002016.

Kunter, M. (2011). Motivation als Teil der professionellen Kompetenz: Forschungsbefunde zum Enthusiasmus von Lehrkräften [Motivation as part of professional competence: Research results regarding enthusiasm of teachers]. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (pp. 259–275). Münster: Waxmann.

Leutner, D., Klieme, E., Meyer, K., & Wirth, J. (2004). Problemlösen [Problem solving.]. In P. I. S. A.-K. Deutschland (Ed.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland — Ergebnisse des zweiten internationalen Vergleichs* (pp. 147–175). Münster: Waxmann.

MA (Millennium Ecosystem Assessment). (Ed.). (2005). *Ecosystems and human well-being: A framework for assessment*. http://www.maweb.org. Accessed 10 Oct 2014.

Mummendey, H. D., & Grau, I. (2008). *Die Fragebogenmethode [Method: Questionnaires]* (5th ed.). Göttingen: Hogrefe.

OECD (Organisation for Economic Co-operation and Development). (2004). *Problem solving for tomorrow's world: First measures of cross-curricular competencies from PISA 2003*. Paris: Author.

Ostermeyer, F., Eggert, S., & Bögeholz, S. (2012). Rein pflanzlich, dennoch schädlich [Exclusively vegetable, however dangerous]? *Unterricht Biologie, 36*(377/378), 43–50.

Oulton, C., Dillon, J., & Grace, M. M. (2004). Reconceptualizing the teaching of controversial issues. *International Journal of Science Education, 26*, 411–423. doi:10.1080/0950069032000072746.

Papadouris, N. (2012). Optimization as a reasoning strategy for dealing with socioscientific decision-making situations. *Science Education, 96*, 600–630. doi:10.1002/sce.21016.

Pólya, G. (1945). *How to solve it: A new aspect of mathematical method*. Princeton: University Press.

Ratcliffe, M., & Grace, M. (2003). *Science education for citizenship: Teaching socio-scientific issues*. Maidenhead: Oxford University Press.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danmarks Padagogiske Institut.

Rotgans, J. I., & Schmidt, H. G. (2011). Situational interest and academic achievement in the active-learning classroom. *Learning and Instruction, 21*, 58–67. doi:10.1016/j.learninstruc.2009.11.001.

Sadler, T., Barab, S., & Scott, B. (2007). What do students gain by engaging in socio-scientific inquiry? *Research in Science Education, 37*, 371–391. doi:10.1007/s11165-011-9260-3.

Sakschewski, M., Eggert, S., Schneider, S., & Bögeholz, S. (2014). Students' socioscientific reasoning and decision making on energy-related issues: Development of a measurement instrument. *International Journal of Science Education, 36*, 2291–2313. doi:10.1080/09500693.2014.920550.

Scherer, R. (2012). *Analyse der Struktur, Messinvarianz und Ausprägung komplexer Problemlösekompetenz im Fach Chemie [Analyzing the structure, measurement invariance, and performance of complex problem-solving competency in chemistry]*. Berlin: Logos.

Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.

Voss, T., Kleickmann, T., Kunter, M., & Hachfeld, A. (2011). Überzeugungen von Mathematiklehrkräften [Beliefs of mathematics teachers]. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (pp. 235–257). Münster: Waxmann.

Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45–65). Seattle: Hogrefe.

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah: Erlbaum.

Wu, M. L., Adams, R. J., Wilson, M. R., & Haldane, S. A. (2007). *Acer conquest version 2.0*. Victoria: Acer.