This Accepted Author Manuscript was subsequently published as Nitz, S., Ainsworth, S. E., Nerdel, C., & Prechtl, H. (2014). Do student perceptions of teaching predict the development of representational competence and biological knowledge? *Learning and Instruction, 31*, 13-22.

Abstract

Dealing with representations is a crucial skill for students and such representational competence is essential for learning science. This study analysed the relationship between representational competence and content knowledge, student perceptions of teaching practices concerning the use of different representations, and their impact on students' outcome over a teaching unit. Participants were 931 students in 51 secondary school classes. Representational competence and content knowledge were interactively related. Representational aspects were only moderately included in teaching and students did not develop rich representational competence although content knowledge increased significantly. Multilevel regression showed that student perceptions of interpreting and constructing visual-graphical representations and active social construction of knowledge predicted students' outcome at class level, whereas the individually perceived amount of terms and use of symbolic representations influenced the students' achievement at individual level. Methodological and practical implications of these findings are discussed in relation to the development of representational competence in classrooms.

Keywords: representational competence, students' perceptions of teaching practices, instructional quality, assessment

1 Introduction

One goal of science education is to enable students to participate in decision-making and public debate regarding scientific issues. To be scientifically literate, students need to be supported in reading, writing, and communicating in science (Krajcik & Sutherland, 2010; Norris & Phillips, 2003; Yore, Pimm, & Tuan, 2007). Reading, writing, and communicating in science do not only rely on verbal discourse and written text. Science is instead a multimodal discourse utilising a variety of representations (e.g., graphs, diagrams, symbols, formulae) and so interpreting, constructing, transforming, and evaluating different scientific representations are crucial skills for students to build and communicate a conceptual understanding of science (Kress, Jewitt, Ogborn, & Tsatsarelis, 2001; Lemke, 2004; Yore & Hand, 2010). These skills have been referred to as representational competence (*RC*, Kozma, Chin, Russell, & Marx, 2000; Kozma & Russell, 1997, 2005) and contribute to being scientific literate. Scientific literacy thus comprises of the interacting dimensions of fundamental literacy, including the abilities to construct and interpret scientific discourses (*inter alia* RC), and the derived understanding about the principles and foundations of science (Norris & Phillips, 2003; Yore et al., 2007).

There is a growing body of studies that explored how RC is influenced by design factors of representations or by the strategies students use to reason with representations (e.g., Canham & Hegarty, 2010; Kozma, 2003; Kozma & Russell, 1997; Stieff, 2011; Stieff, Hegarty, & Deslongchamps, 2011). However, no study we know of has empirically analysed the interactive relationship between RC and content knowledge (CK) and only a few studies have analysed how teaching practices in authentic classroom settings affect RC (e.g., Hubber, Tytler, & Hastam, 2010; Kohl & Finkelstein, 2006; Prain & Waldrip, 2006; Tytler, Prain, & Peterson, 2007). This is surprising as there is strong evidence that students' cognitive and affective outcomes are generally influenced by teaching practices and instructional quality (e.g., Fraser, 2003; Kunter, Baumert, & Köller, 2007; Lipowsky et al., 2009; Schroeder et al., 2011). Therefore, we argue that it is crucial to incorporate an explicit representational focus in teaching and learning science as a part of domain-specific instructional quality in order to develop students' RC. The purpose of this study was twofold: First, we examine the relationship between students' RC and CK for one specific topic in biology education. Second, we analyse how scientific representations were used in biology classes using student perceptions of the classroom and explore how this impacted upon students' RC and CK. In view of the hierarchical nature of school settings with students being nested in classes, we chose a multilevel approach to disentangle student and class effects.

1.1 Representational Competence (RC)

There is consensus among researchers that teaching and learning science involves a variety of external scientific representations (Ainsworth, 2006; Kress et al., 2001; Lemke, 2004; Yore & Hand, 2010). To classify these different representations, various categories have been suggested but no consensus has been reached (Wu & Puntambekar, 2012). We refer to categories by Gilbert (2005) and Wu & Puntembekar (2012) that take the mode and code of representations into account. We focus on verbal-textual representations (spoken or written text, terms), visual-graphical representations with a distinction between realistic and logical pictures[1] (Schnotz, 2001), and symbolic representations in terms of chemical symbols and equations[2].

The development of abilities for dealing and reasoning with these scientific representations is crucial for learning science. RC addresses disciplinary skills for interpreting, constructing, translating, and evaluating representations that students should acquire to become proficient and literate in the domain. It includes the abilities to (a) use representations for describing scientific concepts, (b) identify, describe, and analyze features of representations, (c) construct and/or select a representation and (d) explain its appropriateness for a specific purpose, (e) compare and contrast different representations and their information content, (f) connect across different representations, map features of one type of representation onto those of another, and explain the relationship between them, (g) realize that representations correspond to phenomena but are distinct from them, and (h) use representations in discourse to support claims, draw inferences, and make predictions (Kozma & Russell, 2005).

Studies indicate that students' RC is closely related to students' conceptual understanding of the domain (Kozma & Russell, 1997; Stieff, 2011). However, these studies did not seek to empirically distinguish between content knowledge in the domain (CK) and RC. Kozma & Russell (2005) proposed five levels of RC ranging from a novice surface-based depictive use of representations—via symbolic, syntactic, and semantic use of representations—to an expert reflective and rhetorical use of representations. However, this developmental trajectory is likely to be neither stage-like nor automatic or uniform, and depends on the use of representations and the context of that use in learning environments. Kozma & Russell pointed out that "over time and given appropriate sets of physical, symbolic, and social situations, a student will increasingly display more advanced representational skills, come to internalize these, and integrate these into regular practice" (2005, p. 134). This pedagogical insight places special emphasis on teaching practices in instructional settings as these must provide appropriate opportunities for developing RC. In line with this assumption, we are interested in students' skills using scientific representations (RC), their understanding of the domain (CK), and the impact of teaching practices in biology education.

1.2 Representations and Representational Competence (RC) in Science Class

To foster the development of RC and CK, science instruction should provide students with opportunities to engage actively in representational tasks that make the role and function of representations explicit (Greeno & Hall, 1997; Hubber et al., 2010; Prain & Waldrip, 2006; Stieff, 2011). However, only a few studies have explored how representations are used in science classrooms and how this is related to students' outcomes. Kohl & Finkelstein (2006) found that an increased use of representations in a large-lecture physics course led to increased representational skills. Qualitative analyses of physics teaching units provided evidence that explicit instruction and a representational focus in teaching with diverse opportunitiues for students to develop their RC can foster students' learning of scientific topics and deepen their conceptual understanding (Hubber et al., 2010; Prain & Waldrip, 2006; Tytler et al., 2007). However, these studies did not test specifically for RC or CK and focused on the description of the teaching practices rather than relating these to quantifiable learning outcomes. Hence, the relationship between existing teaching practices, students' RC, and CK needs to be further analysed.

1.3 Assessing Representational Practices in Science Class through Students' Perceptions

In research on instructional quality, there is much discussion concerning how to assess teaching practices. In general, students', teachers', and/or external observers' perceptions are used to this end. Empirical studies, however, found only low to moderate correlations between these different perspectives (e.g., Clausen, 2002; Kunter & Baumert, 2006). These authors pointed out that each perspective constitutes a *specific* perception of the classroom with *perspective-specific* validities depending on the research context and, thus, has a *specific* value for describing classroom-teaching practices (Clausen, 2002; Kunter & Baumert, 2006). Although it is desirable to use different perspectives for describing the complexity of the classroom environment and its impact on student learning (Seidel & Shavelson, 2007), many studies rely on students' perceptions and ratings (Fraser, 2003; Seidel & Shavelson, 2007). Students have encountered a variety of teachers and teaching practices and asking them to rate these is economically applicable in the classroom (Clausen, 2002; De Jong & Westerhof 2001). It has been argued that aggregated student ratings constitute a shared (and more objective) perception of teaching practices rather than representing individual perceptions (Kunter et al., 2007; Lüdtke, Robitzsch, Trautwein, & Kunter, 2009; Lüdtke, Trautwein, Kunter, & Baumert, 2006). In terms of construct validity, Wagner, Göllner, Helmke, Trautwein, and Lüdtke (2013) showed that students are able to differentiate between theoretical criteria of instructional quality and describe their teachers' teaching practices in this respect. Other studies revealed the predictive validity of (aggregated) student ratings for cognitive and affective outcomes (c.f. Fraser, 2003) that is higher than the observers' and teachers' perspectives with regard to student achievement in the domain (Clausen, 2002). Lüdtke et al. (2009, p. 120) assumed that a student is "more affected by his or her interpretation of the classroom context than by any objective indicator of that context". These findings are particularly important in the context of this project. Hence, asking students to rate specified features of their teachers' instruction is an efficient and valid way to assess teaching practices regarding the use of scientific representations in our project.

In an earlier stage of our research project, an instrument with six scales on which students rate representational features of their biology class was developed. They describe the use of visual-graphical, verbal-textual, and symbolical representations in class, namely the (1) interpretation of visual-graphical representations, (2) construction of visual-graphical representations, (3) use of scientific texts, (4) amount of (technical) terms used in class, (5) use of symbolic representations and (6) active social construction of knowledge. These scales were developed based on the assumption that science instruction should explicitly provide students with opportunities to engage in representational tasks (see 1.2) and correspond to theoretical criteria for instructional quality that are described in the following.

A growing body of literature emphasizes that—in addition to the interpretation of given representations—the active construction of representations is particularly important for learning (Ainsworth, Prain, & Tytler, 2011; Hand, Prain, Lawrence, & Yore, 1999; Yore & Hand, 2010). To construct a deeper understanding of science, students need to interpret and construct a variety of representations. Therefore, the instrument covers if and to what extent the interpretation *and* construction of visual-graphical (realistic and logical pictures, Schnotz, 2001), verbal-textual (scientific texts) and symbolic (chemical symbols and equations) representations by the students themselves is addressed and taught explicitly in class (scales 1-3, 5). With regard to verbal-textual representations, the instrument also assesses the amount of terms used in biology class (scale 4). Science, especially biology, is based on substantive terminology and studies showed that the use of too many terms in class can overwhelm students causing difficulties in understanding (Merzyn, 1996; Wellington & Osborne, 2001). Therefore, teachers should concentrate on few, important terms in class. In line with arguments from social semiotics and accepting a discursive and active view of teaching and learning, an important feature of teaching is to provide opportunities to negotiate meanings among students by active social construction of knowledge about representing science (Lemke, 1990; Mortimer & Scott, 2000; Scott, 1998). This aspect is therefore included in the instrument as the extent to which teachers enable students to work together autonomously in groups in which social construction of knowledge could potentially take place (scale 6).

The six-factorial structure of the instrument was established by exploratory factor analysis in a previous study, its construct validity was confirmed by multilevel confirmatory factor analysis at individual and classroom level, and the reliability of the six scales was satisfying at both levels (Nitz, Nerdel, & Prechtl, 2012, Nitz, Prechtl, & Nerdel, in press). In the current study, we applied this instrument to analyse the degree to which students' learning of representational and content aspects of biology was influenced by these representational teaching practices.

2 Current Study

How representations are dealt with in science classes and the impact of these representational practices on students' outcomes need to be further examined. We focused all of our work within the same teaching unit on photosynthesis in biology education (see also 4.2.2) as we also analyse students' CK in line with Kozma & Russell's (2005) argument that the development of RC is dependent on the context. First, we examine students' RC and CK on photosynthesis as well as the relationship between these two constructs. Previous research indicated that students' RC is closely related to students' conceptual understanding of the domain, but these studies did not empirically distinguish between CK and RC (Kozma & Russell, 1997; Stieff, 2011). We would like to answer the following research questions: (1) What role does students' CK play in relation to students' RC and vice versa? We hypothesize that both constructs are related but empirically distinguishable (Hypothesis 1a) and that RC influences CK and vice versa (Hypothesis 1b). (2) Do students' RC and CK vary systematically according to class? We think this is the case for both variables (Hypothesis 2). If so, analyses of teaching practices are a promising approach to account for this variation which is why we explore representational teaching practices as perceived by students to address the research questions: (3)

How often are visual-graphical, verbal-textual, and symbolic representations used in photosynthesis instruction? (4) Are there meaningful variations between classes in teaching practices concerning the use of these representations? Based on research on teaching practices in physics classrooms (see 1.2) we hypothesize that there are meaningful variations between classes (Hypothesis 3). We are then fundamentally interested in the relationship between these teaching practices and the students' RC and CK: i.e., (5) do teaching practices concerning the use of representations in the classroom impact on students' RC and CK? In line with previous research on the impact of teaching practices and on the importance of adequate learning opportunities, we hypothesize that there would be a positive effect of the frequency and extent to which representations are used in class on the students' RC and CK (Hypothesis 4).

3 Method

3.1 Sample

The overall project was conducted in two stages with in total 1253 students in 67 classes of German secondary school. To validate the representational competence measure in stage 1 (see also 3.2.1), we ran pilot studies 1 and 2 in five and three classes (grade 11 and 12) with 80 and 67 students (44% female, 33% male), respectively. The factorial structure of the instrument to assess representational practices (see 3.2.3) was established using exploratory factor analysis in stage 1 with eight classes and 175 students (51% female, 49% male) aged 15 to 19 (M = 16.7, SD = 0.6) and its construct validity was confirmed by multilevel confirmatory factor analysis using data from 626 students in 45 classes (Nitz et al., 2012; in press).

In the current study, we report data from 931 students (61% female, 39% male) aged 16 to 21 years (M = 17.8, SD = 0.7) in 51 grade 12 classes of 20 different schools. Nearly all participating classes had been taught by their respective biology teacher for at least half a year

(97%) at the time of data collection. Therefore, both teachers and students were familiar with the classroom environments. The mean number of students per class was 18.

3.2 Measures

3.2.1 Representational competence (RC).

RC is a complicated and broad construct; therefore we decided to focus on a single aspect when developing a measure of students' skills in using representations. We focused on the transformation of scientific information between at least two different modes of representation. We set this focus because both interpretation of a given representation and the construction and selection of another representation are necessary (but not sufficient) to transform information. The measure incorporated the following skills described by Kozma & Russell (2005): use representations for describing scientific concepts (*interpretation*); generate and/or select a representation (*construction*); identify, describe, and analyse features of representations (*interpretation*); and connect across different representations and explain the relationship between them (*translation*). Most items relate to the syntactic or semantic level of RC (Kozma & Russell, 2005).

For item development in stage 1 of this project, teaching materials were analysed to identify representations that are used in instruction on photosynthesis based on the state-prescribed curriculum. We found—beside the text—visual-graphical (realistic pictures, e.g., photomicrographs and schematic illustrations of chloroplasts; logical pictures, e.g., graphs of absorption spectra, flowcharts of the photosynthetic z-scheme) and symbolic representations (e.g., chemical equations, chemical formulas, and structural formulas). In pilot study 1, these representations were incorporated in 27 tasks with 45 open-ended items to explore students' ideas of how to represent a specific topic. Students' responses were analysed to identify common difficulties and errors. Based on these analyses, distractors for multiple choice items were created and tested in a pilot study 2. In the final measure, 12 tasks with 18 multiple choice items that showed satisfactory difficulties in prior analyses were chosen for the current study (see supplementary material for an example). These items included scientific texts, realistic pictures, logical pictures, chemical symbols, and chemical equations and asked students to translate a given representation into another one. Content-related information relevant to solve the task was given in the item stem to ensure that the items could potentially be solved independently from knowledge about photosynthesis. All items were scored dichotomously.

In the current study, we used confirmatory factor analysis (CFA, using the WLMSV estimator in Mplus 6) to establish the factorial structure of the instrument and to test for measurement invariance over time. Three items were excluded from final analyses due to insignificant factor loadings. A CFA with factor loadings and thresholds constrained to be equal over the two measurement points showed satisfying fit indices for a unifactorial solution with auto-correlated residuals [$\chi^2(417, N = 926) = 474.01$, $p = 0.03$, CFI = 0.920, RMSEA = 0.012; see 3.4 regarding the evaluation of fit indices]. The assumption of strong factorial invariance is therefore supported. The reliability for the post-test was estimated as internal consistency at the individual level (Cronbach's $\alpha = 0.68$) and as ICC(2) at the classroom level [ICC(2) = 0.90]. When interpreting the reliability at the individual level, which is not yet satisfactory, it is necessary to keep in mind that RC is a heterogeneous construct, which necessarily lowers the internal consistency. We therefore think that the reliability is acceptable in the context of our study.

3.2.2 Content knowledge (CK) on photosynthesis

The test on CK consisted of six multiple choice items and covered central concepts within the topic photosynthesis which are prescribed in the curriculum for that teaching unit (e.g., light-dependent and light-independent reactions, influence of ecofactors). All items were scored dichotomously. A CFA (using the WLMSV estimator in Mplus 6) with factor loadings constrained to be equal over the two measurement points showed good fit indices for a unifactorial solution with auto-correlated residuals [$\chi^2(52, N = 929) = 66.80$, $p = 0.08$, CFI = 0.952, RMSEA = 0.018; see 3.4 regarding the evaluation of fit indices]. Implementing the assumption of strong factorial invariance [$\chi^2(57, N = 929) = 93.31$, $p = 0.002$, CFI = 0.880, RMSEA = 0.026] lowered the fit of the model significantly [$\Delta\chi^2(5, N = 929) = 29.63$, $p < 0.001$] and is not supported. A model with three thresholds estimated freely (according to modification indices) reached satisfying fit indices [$\chi^2(54, N = 929) = 74.24$, $p = 0.035$, CFI = 0.935, RMSEA = 0.020]. We therefore assume partial strong measurement invariance.

The reliability at the individual level (Cronbach's α) was 0.62 and at the classroom level [ICC(2)] 0.90 for the post-test. The reliability at the individual level is not satisfactory. However, due to pragmatic reasons we wanted to cover different concepts within the topic to assess CK in breadth with as few items as possible, which lowers the internal consistency.

3.2.3 Student perception of teaching practices regarding the use of scientific representations.

To assess the teachings practices regarding the use of scientific representations in class we applied the instrument that had been developed during an earlier stage of this research project. Students were requested to rate 21 items on six scales on a 4-point scale ranging from 1 = "never" to 4 = "regularly" to document their perceptions of how frequently and to what extent the following aspects occurred in teaching practices (see 1.3): (1) *interpretation of visual-*

graphical representations, (2) *construction of visual-graphical representations*, (3) *use of scientific texts*, (4) *amount of (technical) terms used in class, (5) use of symbolic representations* and (6) *active social construction of knowledge*. Items for the scale active social construction of knowledge were adapted from the COACTIV-study (Baumert et al., 2009). To account for the role inference plays in student ratings (see Wagner et al., 2013) and to make the rating easier for the students, the items describe specific, observable practices in the classroom (e.g., constructing a graph) and therefore constitute rather low-inferent ratings. Furthermore, short definitions of and examples for each representation were included in the instruments' instruction to make sure that students interpret the items accordingly.

In order to obtain classroom level variables in this study, class means of student ratings on each scale were used. All scales used in this study showed satisfying reliability at both individual level (measured as Cronbach's α, see Table 1) and classroom level [measured as ICC(2)] as well as sufficient agreement among students within a classroom (measured as index of within-group agreement r_{WG}, see Lüdtke et al., 2006).

The convergent validity of the scales with regard to the teachers' perspective was analysed in a subsample of 25 classes. A teacher version of the instrument with parallel items was answered by the respective teachers after the teaching unit. Correlations between student and teacher scales ranged from r = -0.46 to 0.57 with an average effect size of r = 0.25 (see Table 1), which are similar to values reported by Clausen (2002). The negative correlation points to an idiosyncratic perception of the amount of terms in class by students and teachers. However, this is in line with the argument that different perspectives constitute specific values for describing teaching practices (see 1.3).

Table 1:

Employed scales with sample items, reliability coefficients [Cronbach's α, ICC(2)], intra-class correlations, means of within-class agreement for each scale ($r_{WG(J)}$), descriptive statistics for individual and class level, and convergent validity coefficients (with teachers' perspective)

Scale	Sample item	α	ICC(1)	ICC(2)	M $r_{WG(J)}$	M (SD), range, valid n		$r_{teacher}$
						Ind. level	Class level	
IntVis (4 items)	We practice how to obtain information from graphs.	0.71	0.18	0.76	0.85	3.38 (0.54) 1-4, 676	3.40 (0.27) 2.83-3.90, 45	0.33*
ConstVis (4 items)	We create graphs on our own in biology class.	0.83	0.30	0.86	0.78	2.25 (0.74) 1-4, 668	2.29 (0.45) 1.05-3.34, 45	0.25
Texts (4 items)	We write our own short scientific texts on biological topics in class.	0.63	0.35	0.89	0.78	2.61 (0.63) 1-4, 675	2.65 (0.40) 1.58-3.45, 45	0.23
Terms+ (3 items)	Too many technical terms are used in one lesson.	0.83	0.14	0.70	0.57	2.78 (0.83) 1-4, 664	2.76 (0.38) 1.86-3.36, 45	-0.46**
Sym (3 items)	Our biology teacher explicitly addresses the handling of reaction equations in class.	0.70	0.27	0.84	0.69	2.60 (0.73) 1-4, 667	2.62 (0.43) 1.72-3.51, 45	0.49**
Asc (3 items)	We present results found in team work, individual work and project work in biology class.	0.79	0.50	0.93	0.72	2.69 (0.82) 1-4, 674	2.71 (0.60), 1.02-3.70, 45	0.57**

Note: N = 806 students, 45 classes, N ($r_{teacher}$) = 25 classes

IntVis: Interpretation of visual-graphical representations, *ConstVis*: Construction of visual-graphical representations, *Texts*: Use of scientific texts, *Terms*: Amount of terms, *Sym*: Use of symbolic representations, *Asc*: Active social construction of knowledge

$^{+}$ Items were recoded, thus high values represent the use of less terms in class.

ICC(1): Proportion of variance explained by class membership

ICC(2) was calculated based on ICC(1) and the average class size (14.31) applying Spearman-Brown formula.

*$p \leq 0.05$, **$p < 0.01$ (one-tailed)

3.3 Procedure

The overall project took place within the teaching unit of photosynthesis because an analysis of teaching materials had revealed that this topic provided potentially rich opportunities for teachers to incorporate representational aspects (see 3.2.1). For the current study, students' CK and RC was assessed using the measures described in 3.2.1 and 3.2.2 in separate test booklets before the teaching unit on photosynthesis took place in 51 participating classes. The measures were administered during the first lesson of the teaching unit and students had 45 minutes to answer it. Then a photosynthesis unit which was based on the state-prescribed curriculum and lasted 15 lessons on average (*SD* = 6) was taught by the respective teacher. Teachers were neither advised on how to teach photosynthesis nor were given any instructional materials. They were also encouraged to teach the photosynthesis unit per their usual procedures. Afterwards, students' CK and RC were assessed using the same measures during the last lesson of the teaching unit. The students were also asked to rate specific representational features of their teachers' instruction using the instrument developed in an earlier stage of this project (Nitz et al., 2012; in press, see 3.2.2) in another test booklet. Students had 45 minutes to answer all three test booklets.

3.4 Statistical Analyses

To establish the factorial structure of the RC and CK tests, we ran confirmatory factor analyses for dichotomous indicators in Mplus 6 (see also 3.2.1, 3.2.2). The WLMSV estimator was used to account for the nested structure of our data. Multiple fit indices [χ^2 likelihood ratio statistic, comparative fit index (CFI), root-mean-square-error of approximation (RMSEA), and the standard root mean residual (SRMR)] were used to assess the goodness of fit (Hu & Bentler, 1999).

Multilevel analyses were used to analyse the relationship among students' RC, CK, and the ratings on teaching practices concerning the use of representations. Multilevel analyses allows considering predictors at both individual (level 1) and classroom levels (level 2) simultaneously and correcting standard errors (Hox, 2002; Snijders & Bosker, 1999). We accounted for both levels because Fauth, Decristan, Rieser, Klieme, and Büttner (2014, p. 6) argued that constructs such as teaching practices are "conceptually a classroom level construct that is individually perceived by students". All models reported are random-intercept models estimated by full information maximum likelihood (FIML). Individual student variables were introduced as predictors at level 1 and aggregated variables (mean class scores) at level 2. All predictors were standardized to a common metric (z-scores) and individual student variables were centered at the group mean as recommended by Lüdtke et al. (2009). Within-class and between-class effects on students' outcome could thus be clearly disentangled. In Model 2, we controlled for prior RC and CK, respectively, at both levels. Therefore, the coefficients of all other predictors represent the effect on the development of RC and CK over the time of the teaching unit.

4 Results

First, we examine the relationship of students' RC and CK and present results regarding the change over the time of the teaching unit and the variation among classes. Second, we describe students' perception of the frequency with which different representations were interpreted and constructed and test whether there were meaningful variations among different classes. Third, we analyse to what extent the perceived teaching practices concerning the use of representations in biology class impacted upon students' RC and CK.

4.1 Students' Representational Competence (RC) and Content Knowledge (CK)

To answer the question what role students' CK plays in relation to students' RC and vice versa, we ran confirmatory factor analyses to test whether a two-dimensional model with separate factors for CK and RC or a one-dimensional model fits better to the data. The assumptions of partial strong measurement invariance over time (factor loadings for all items and thresholds for all but three CK items were constrained to be equal over time, see 3.2) and autocorrelated residuals were included in the models. Confirmatory factor analyses showed a better fit to the data for a two-dimensional model [$\chi^2(827, N = 931) = 902.77$, $p = 0.034$, CFI = 0.916, RMSEA = 0.010] than for an one-dimensional model [$\chi^2(835, N = 931) = 977.04$, $p = 0.0005$, CFI = 0.842, RMSEA = 0.014]. This difference was statistically significant [$\Delta\chi^2(8, N = 931) = 120.78$, $p < 0.001$]. CK and RC are correlated in pre- and post-test (see Table 2).

Table 2

Correlations (from CFA) of representational competence (RC) and content knowledge (CK) in pre- and post-test

		Pre-test	Post-test	
		CK	RC	CK
Pre-test	RC	0.52	0.78	0.44
	CK		0.59	0.57
Post-test	RC			0.69

Note: $N = 931$

All correlations significant at $p < 0.001$

Mean scores of the RC measure showed a small increase from Time 1 to 2 [pre-test: M (SD) = 0.41 (0.18), post-test: M (SD) = 0.46 (0.20); $t(663) = 6.37$, $p < 0.001$, $d = 0.27$]. In contrast, students' CK increased significantly over the time of the teaching unit [pre-test: M (SD) = 0.24 (0.18), post-test: M (SD) = 0.69 (0.27); $t(663) = 41.18$, $p < 0.001$, $d = 1.61$]. To find out if there were systematic variations between different classes, we calculated intra-class correlation coefficients. The results of 0.26 for pre-test and 0.37 for post-test indicate that a substantial amount of variance in RC could be explained by class membership. This also holds true for the students' CK [$\text{ICC}(1)_{\text{pre-test}} = 0.18$, $\text{ICC}(1)_{\text{post-test}} = 0.34$].

To investigate the relation of students' RC and CK over the time of the teaching unit, a multilevel model (Model 1) was specified that included the stabilities and cross-lagged relationships between these two variables (see Figure 1). Both variables showed stability as performance at Time 1 predicted performance at Time 2. RC was more stable than CK and both variables were positively related with each other at the same measurement points. The cross-lagged analysis revealed that variance in RC at Time 2 can be explained by individual differences in CK at Time 1 and vice versa. At the between class level, prior CK predicted variance in RC at Time 2, but prior RC did not predict CK. The correlation of the residuals at Time 2 indicated that there are additional, common factors—such as teaching practises, for example—that influence both RC and CK. Furthermore, the explained variance for both variables at both levels was rather small. The standardized solution for this model is presented in Figure 1.

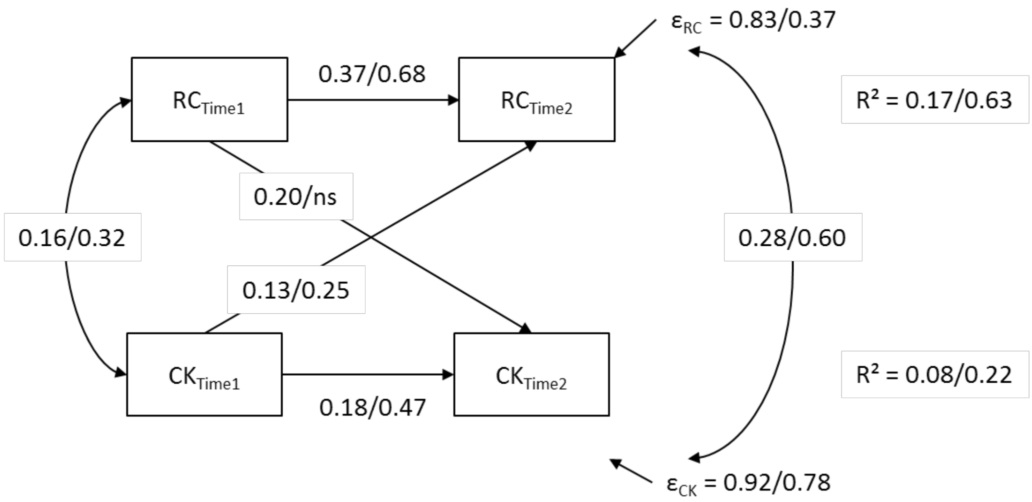

Figure 1. Two-level path model (Model 1) for the relationship of representational competence (RC) and content knowledge (CK) with standardized estimates (within-class/between-class) (N = 931 students, 51 Classes)

Due to the results of the Model 1 and the systematic variation of RC and CK across different classrooms as indicated by ICC(1), the influence of additional, common variables on both variables—such as teaching practices—should be explored.

4.2 Perceived Teaching Practices Concerning the Use of Representations

Research questions 3 and 4 concerned how often visual-graphical, verbal-textual and symbolic representations are perceived to be used in teaching photosynthesis and if there are meaningful variations among classes. Overall mean ratings on teaching practices scales (see Table 2) differed significantly [$F(4.01, 2658.84) = 226.65$, $p < 0.001$, $\eta^2 = 0.26$; we used the Huynh-Feldt correction ($\varepsilon = 0.80$) because Mauchly's test indicated that the assumption of sphericity was violated ($\chi^2(14) = 395.38$, $p < 0.00$)]. Fairly equal and moderate frequencies of representational aspects of teaching were found with the exception of the interpretation of visual-graphical representations, which students rated to be most commonly integrated in teaching practices. In contrast, the construction of visual-graphical representations by the students themselves was rated to occur least often and only rarely in classes.

There was also systematic variation in the perceived teaching practices concerning the use of representations across different classrooms. The intra-class correlations (see Table 2) indicate that 14% to 50% of total variance was located between classes. This variation makes it important—similarly to the systematic variation of RC and CK between classes—to account for the multilevel structure of the data.

4.3 Influence of Perceived Teaching Practices on Students' Representational Competence (RC) and Content Knowledge (CK)

To determine the effect of student ratings on teaching practices concerning the use of representations in biology classes on students' RC and CK, we incorporated the teaching practices in the multilevel model. Table 3 presents standardized regression coefficients of individual and class predictors as well as the proportion of explained variance (Model 2). In this model, the unique contribution of each teaching practice is evaluated with regard to the prediction of the development of students' outcome. Based on previous research, we hypothesized a positive effect of the frequency and extent to which representations are used in class on students' RC and CK. The results did not confirm this assumption.

Turning first to the class level, the frequency and extent of interpreting visual-graphical representations in class had a positive effect, the frequency and extent of constructing visual-graphical representations a negative effect on the development of students' RC (controlling for each of the other teaching practices). These teaching practices constituted a similar effect on students' CK. Additionally, we found a positive effect of the active social construction in classes on students' CK.

At the individual level, only the students' individual perception of the amount of terms used in class significantly predicted RC in the post-test when controlling for all of the other teaching practices (Items of this scale were recoded, thus high values on this scale refer to less use of terms in class). The development of students' CK was predicted by the individually perceived amount of terms and frequency of using symbolic representations.

Table 3

Multilevel regression of individual and class variables on students' representational competence (RC) and content knowledge (CK) after the teaching unit (Model 2)

	Individual level				Class level			
	RC		CK		RC		CK	
	β	*SE*	β	SE	β	SE	β	SE
Prior RC (Time 1)	0.33***	0.04	0.13**	0.05	0.67***	0.10	0.10	0.09
Prior CK (Time 1)	0.11**	0.04	0.17***	0.04	0.08	0.13	0.27	0.15
Teaching practices								
Interpretation of visual representations	0.03	0.03	0.01	0.03	0.20*	0.09	0.34**	0.11
Construction of visual representations	0.05	0.04	-0.01	0.05	-0.51***	0.12	-0.67***	0.14
Use of scientific texts	-0.03	0.04	0.05	0.05	0.02	0.13	0.08	0.14
Amount of terms	0.15**	0.04	0.18***	0.04	0.08	0.09	0.20	0.11
Use of symbolic representation	0.04	0.03	-0.08*	0.04	0.12	0.12	0.08	0.13
Active social construction of knowledge	-0.03	0.03	0.01	0.04	0.17	0.11	0.53***	0.13
Explained variance (%)	19		12		80		78	

Note: $N = 664$ students, 45 classes

β = standardized regression coefficients

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

5 Discussion

One purpose of this study was to analyse students' representational competence (RC) and content knowledge (CK) and their development over one teaching unit. We found that the constructs were interactively related, but empirically distinguishable (Hypothesis 1a) which is in line with Kozma & Russell (2005) and Stieff (2011). CK improved robustly over the course of the lessons (45%). However, although it was significant, the improvement in RC was slight (5%). The results also showed a directional effect at class level as pre-test class means of CK positively predicted RC after the teaching unit but not vice versa (Hypothesis 1b). Although there is little directly related research to compare our findings to they are consistent with arguments in the representational competence literature. Kozma & Russell (2005) argue the slow developmental trajectory of RC means that it takes considerable time and practice to achieve expertise. Thus, one teaching unit (an average of 15 lessons) in this study is probably too short a period of time to see large effects on representational competence. Moreover, the representational aspects of photosynthesis were generally rated by students to be only moderately included in teaching practices (except for the interpretation of visual-graphical representations). Consequently, it would appear that either little time is spent in these classes on these representational practices or students do not perceive them as opportunities to learn. Moreover, representational competence does not appear to develop easily without explicit instruction. Accordingly, one implication of our study is that representational activities should be more frequently incorporated into the biology classroom (although not necessarily in the same way they are at present). We also need to understand why teachers choose or not to include them (e.g. perceptions of the value of these activities, lack of time, assessment practices, etc.).

In line with our hypotheses, students' RC and CK (Hypothesis 2) and the perception of teaching practices (Hypothesis 3) varied systematically across classes, i.e. the classroom environment influenced student variables and ratings and differences between classes were substantial. Clearly, the chosen teaching unit does not necessitate one specific approach and so does provide opportunities for teachers to incorporate a range of representational teaching practices (see 3.3). Given these differences, we examined the data to explore how the students' perception of representational teaching practices in different biology classes impacts on their RC and CK. We had hypothesised a positive relationship between the perceived frequency and extent of representational teaching practices and the development of students' understanding (Hypothesis 4). We took both levels into account in our analyses because the representational teaching practices are classroom level constructs which are individually perceived by students (c.f. Fauth et al., 2014). However, there was striking variation between the six assessed teaching practices for students' learning outcomes.

We found three substantial between-class effects of the assessed teaching practices (when controlling for prior RC and CK). These aggregated student ratings represent the mean (or shared) perception of all students in a class and tend to be more reliable and objective indicators for teaching practices than individual student ratings. In total, 80% and 78% of variance in students' RC and CK, respectively, could be explained by these aggregated student ratings alongside prior RC and CK. As these effects were limited to the class level, student outcomes were not dependent on a student's idiosyncratic perception of these practices.

The commonly perceived frequency and extent of *interpreting visual-graphical representations* (which was also commonly rated as most frequent occurring in participating classes) had a clear positive association with both students' RC and CK. Accordingly, it seems that students develop a deeper conceptual understanding of both representations and content in classes in which they are more supported by their teachers' instruction in interpreting visual-graphical representations which is in line with previous research (Hubber et al., 2010; Kohl & Finkelstein, 2006; Prain & Waldrip, 2006; Tytler et al., 2007).

In contrast, we found a substantial negative between-class effect for the mean perception of *constructing visual-graphical representations* in class on the development of students' RC and CK. Again, there was no effect at individual level, indicating that students' conceptual understanding decreases the more time is spent (from students' perspective) on constructing visual-graphical representations in class. This finding is controversial as recently many have argued for increasing focus on constructing visual-graphical representations in science education (e.g. Ainsworth et al. 2011; Hubber et al 2010). However, before concluding that we should now caution teachers against encouraging construction of visual-graphical representations in the class, two aspects of this study should be noted. First, the outcomes measures used in this study, did not ask students to actively construct visual-graphical representations; in the RC measure instead they selected the representation they thought to be the most adequate one among different representations. Accordingly, this may have resulted in a mismatch between the classroom practices with representations and the nature of the assessment. Thus, future research should consider adding a visual-graphical construction task to learning outcomes to align these more explicitly to see if construction of visual representations is beneficial (see Lowe, Schnotz, & Rasch, 2011). Secondly, we must consider the role that support and scaffolding might play. Studies showed that support is particularly crucial if constructing visual-graphical representations is to be as an effective strategy for learning (e.g., Van Meter, Aleksic, Schwartz, & Garner, 2006; Van Meter & Garner, 2005). Although the teaching practices assessed here have included opportunities for students to construct representations by themselves, we cannot be sure about the presence of adequate support in the class. From this point of view, we can conclude that not only providing opportunities but also support and scaffolding may be needed to foster students' RC and CK.

We also found a substantial positive effect of the perceived opportunities to negotiate meanings among students by *active social construction* of knowledge in the class on students' CK that was limited to the class level, indicating that students have higher CK post-test scores in classes in which group work is more commonly perceived than in classes in which it is less perceived. However, there was no effect on students' RC. When working together autonomously in groups, students seem more likely to pay attention to the content rather than the representational aspects of the domain. This might indicate that a teacher-mediated explicit negotiation of representational aspects in the class matters for supporting students to develop RC (see approaches such as Tytler et al., 2007; Waldrip, Prain, & Carolan, 2010).

The class-level variables in our study were good predictors of students' development of RC and CK. The commonly perceived teaching practices conceptually represent the students' perspective on instructional quality which is also closely related to the teachers' professional competence (Baumert et al., 2010). The large between-class effects therefor imply that more attention should be paid to the role teachers play in implementing representational aspects in teaching and learning.

Interestingly, two of the six assessed teaching practices were we found to be predictive for students' outcome only at the individual level. Only small amounts of variance between students' RC and CK, respectively, could be explained by individual variables and the effects were rather small compared to the class level. Nevertheless, these effects are interesting because they underline the importance of a student's specific individual perception of the class-level construct teaching practices and therefore the need to account for both levels in research utilizing student ratings of the classroom (c.f. Fauth et al., 2014).

Firstly, it was found that students who perceive a greater *use of terms* in class compared with other students in their class tend to have lower RC and CK in post-test. This finding is consistent with other research in science education showing that the use of too many terms in class can overwhelm students and hinder the acquisition of conceptual knowledge (Merzyn, 1996; Wellington & Osborne, 2001). However, the lack of class effects indicates that not necessarily the objective number of terms in class is the decisive factor but rather a student's individual experience. Students may only notice the use of scientific terminology when they are unfamiliar with its meaning (and this varies across the individuals within a class). In this regard, it is also important to note that generally students perceived the amount of scientific terms in class differently to their teachers' (see Table 1). Due to their domain expertise, teachers may not notice how many such terms they actually use. Consequently, in addition to the established educational implication that teachers should focus on a limited number of central terms (e.g., Kattmann, 2008), a further implication follows from our result: Teachers could be advised to monitor their students' understanding of (central) terms, and then—in terms of differentiated instruction within classes—look for opportunities to specifically support individual students' understanding of these terms.

We also found a slightly negative influence of the individually perceived *use of symbolic representations* on students' CK: Students who perceived a more frequent use of symbolic representations compared to their classmates achieved lower post-test scores for CK. There is no doubt that interpreting and constructing symbolic representations like chemical formulae and reaction equations are necessary to understand biological-physiological processes. However, research has shown that many students have considerable difficulties with symbolic representations which leads to a misunderstanding of related scientific concepts (Taskin & Bernholt, 2012). Our finding suggests that, similar to the impact of the individually perceived amount of terms, the students' individual perception of these practices matter and simply spending time on symbolic representations in class does not ensure that students will learn how to use them and that this will enhance their CK. Consequently, more differentiated learning opportunities are needed within classes.

The final perceived aspect of teaching we explored (*use of scientific texts)* was not found to influence the development of students' RC and CK either positively or negatively.

To summarize briefly, the teaching practices found in our study did not, in general, help students to build rich RC although CK did increase substantially. In the final section therefore we will explore whether there are significant limitations to our study that would make these findings unreliable before turning to the implication of these results.

6 Limitations and Future Research

The most obvious question when assessing the limitation of this study is whether students rating are a valid and reliable source of evidence for assessing teaching practices. Previous research has found only a low to medium correlation between student and others' ratings (Clausen, 2002; Kunter & Baumert, 2006). However, this does not necessarily mean that students are unable to accurately report teaching practices. Students are after all the addressees of teaching and, thus, provide specific insights into classroom practices which previous work has found can be more predictive of learning outcomes than others' views (Clausen, 2002; Kunter & Baumert, 2006). We collected data from only a subsample of teachers and so are not in a position to know whether the ratings of teacher, student or external observers would have differed in their predictions concerning the relation of teaching practices and the development of students' understanding. Clearly, this is an important question for future research.

Secondly, this study focussed on the frequency of teaching practices and not their perceived quality. Future work could explore the nature of effective support and scaffolding strategies to adequately foster students' RC and CK. This could then be included within experimental designs to provide robust evidence for what works and what does not work. Preferably, the implementation of these strategies and the analysis of their efficacy in classroom settings should be addressed simultaneously (c.f. Van Meter & Garner, 2005).

Another limitation of our study is the restricted generalizability of the results. As we were interested in students' domain-specific RC and CK, we only surveyed one specific teaching unit in biology education. The next step in terms of generalizability would be to survey different topics to detect similarities and differences in teaching practices and the interactive development of RC and CK. Another important question is if RC acquired in one topic might be transferable to another one and future work could assess the extent to which it transferred to subsequent science topics in different domains.

Nonetheless, we suggest that this line of research is important. Few studies have focused on domain-specific teaching practices and examined pre and post-tests to establish a causal direction to interpret the effects of teaching practices on students' outcomes (see also Kunter et al., 2007; Lipowsky, et al., 2009). None (to our knowledge) have looked specifically at representational competence although there is now widespread acceptance of its importance with science practice and education (Yore & Hand, 2010). Our findings suggest that this oversight is unfortunate. There was only limited improvement in students' representational competence and perceived frequencies of teaching practices were both positively and negatively related to its development. Accordingly, we agree that attention is needed to explore a wider variety of effective teaching practices with representations (Tytler et al., 2007; Waldrip et al., 2010) and join Eilam (2012) in arguing that more attention should be paid to representations within professional development programmes to adequately prepare teachers to foster students' understanding of this topic. Representational competence plays a fundamental role in fostering scientific literacy and enabling students to participate in discourse about scientific topics (Lemke, 2004; Yore & Hand, 2010; Kozma et al., 2000), consequently research studies such as the one described in this paper help us further understand the crucial role that teachers and teaching practices play in supporting student learning in science.

References

Ainsworth, S. E. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*(3), 183-198.

Ainsworth, S. E., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science, 333*(6046), 1096-1097.

Baumert, J., Blum, W., Brunner, M., Dubberke, T., Jordan, A., Klusmann, U. … Tsai, Y.-M. (2009). *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV): Dokumentation der Erhebungsinstrumente (Materialien aus der Bildungsforschung Nr. 83).*) [Professional competence of teachers, cognitively activating instruction, and development of students' mathematical literacy (COACTIV): Documentation of instruments (materials for empirical educational research No. 83)].Berlin, Germany: Max-Planck-Institut für Bildungsforschung.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., … Tsai, Y.-M (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133–180. doi:10.3102/0002831209345157

Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction, 20*, 155-165. doi: 10.1016/j.learninstruc.2009.02.014

Clausen, M. (2002). *Unterrichtsqualität - Eine Frage der Perspektive?* [Instructional quality – a question of perspectives?] Münster, Germany: Waxmann.

De Jong, R., & Westerhof, K. J. (2001). The quality of student ratings of teacher behavior. *Learning Environments Research, 4*, 51-85. doi: 10.1023/A:1011402608575

Eilam, B.(2012). *Teaching, learning, and visual literacy: The dual role of visual representation in the teaching profession*. New York, NY: Cambridge University Press.

Fauth, B., Decristan, J., Rieser, S. Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction, 29*, 1-9. doi: 10.1016/j.learninstruc.2013.07.001

Fraser, B. J. (2003). Science learning environments: Assessment, effects and determinants. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education* (pp. 527-564). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Gilbert, J. K. (2005). Visualization: A metacognitive skill in science and science education. In J. K. Gilbert (Ed.), *Visualization in science education* (pp. 9-27). Dordrecht, the Netherlands: Springer.

Greeno, J. G., & Hall, R. (1997). Practicing representations: Learning with and about representational forms. *Phi Delta Kappan, 78*, 361-367. doi: 10.2307/20405797

Hand, B., Prain, V., Lawrence, C., & Yore, L. D. (1999). A writing in science framework designed to enhance science literacy. *International Journal of Science Education, 21*, 1021-1035. doi: 10.1080/095006999290165

Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: A multidisciplinary journal*, *6*(1), 1-55. doi: 10.1080/10705519909540118

Hubber, P., Tytler, R., & Hastam, F. (2010). Teaching and learning about force with a representational focus: Pedagogy and teacher change. *Research in Science Education, 40*, 5-28. doi: 10.1007/s11165-009-9154-9

Kattmann, U. (2008). Texte [Texts]. In H. Gropengiesser & U. Kattmann (Eds.), *Fachdidaktik Biologie* (8th ed., pp. 357–366). Köln, Germany: Aulis-Verlag Deubner.

Kohl, P. B., & Finkelstein, N. D. (2006). Effect of instructional environment on physics students' representational skills. *Physical Review Special Topics - Physics Education Research, 2*, 1-19. doi: 10.1103/PhysRevSTPER.2.010102

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction, 13*, 205-226. doi: 10.1016/S0959-4752(02)00021-X

Kozma, R., Chin, E., Russell, J., & Marx, N. (2000). The roles of representations and tools in the chemistry laboratory and their implications for chemistry learning. *Journal of the Learning Sciences, 9*, 105-143. doi: 10.1207/s15327809jls0902_1

Kozma, R., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching, 34*, 949-968. doi: 10.1002/(SICI)1098-2736(199711)34:9<949::AID-TEA7>3.0.CO;2-U

Kozma, R., & Russell, J. (2005). Students becoming chemists: Developing representational competence. In J. K. Gilbert (Ed.), *Visualizations in science education* (pp. 121-146). Dordrecht, The Netherlands: Springer.

Krajcik, J. S., & Sutherland, L. M. (2010). Supporting students in developing literacy in science. *Science, 328*, 456-459. doi: 10.1126/science.1182593

Kress, G., Jewitt, C., Ogborn, J., & Tsatsarelis, C. (2001). *Multimodal teaching and learning: Rhetorics of the science classroom*. London, England: Continuum.

Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research, 9*, 231-251. doi: 10.1007/s10984-006-9015-7

Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction, 17*, 494-509. doi: 10.1016/j.learninstruc.2007.09.002

Lemke, J. L. (1990). *Talking science: Language, learning, and value*. Norwood, NJ: Ablex Publishing Corporation.

Lemke, J. L. (2004). The literacies of science. In E. W. Saul (Ed.), *Crossing borders in literacy and science instruction: Perspectives on theory and practice* (pp. 33-47). Arlington, VA: International Reading Association.

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction, 19*, 527-537. doi: 10.1016/j.learninstruc.2008.11.001

Lowe, R., Schnotz, W., & Rasch, T. (2011). Aligning affordances of graphics with learning task requirements. *Applied Cognitive Psychology, 25*, 452-459. doi: 10.1002/acp.1712

Lüdtke, O., Robitzsch, A., Trautwein, U., & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology, 34*, 120-131. doi: 10.1016/j.cedpsych.2008.12.001

Lüdtke, O., Trautwein, U., Kunter, M., & Baumert, J. (2006). Reliability and agreement of student ratings of the classroom environment: A reanalysis of TIMSS data. *Learning Environments Research, 9*, 215-230. doi: 10.1007/s10984-006-9014-8

Merzyn, G. (1996). A comparison of some linguistic variables in fifteen science texts. In G. Welford, J. Osborne & P. Scott (Eds.), *Research in science education in europe: Current issues and themes* (pp. 361-369). London, England: Falmer Press.

Mortimer, E., & Scott, P. (2000). Analysing discourse in the science classroom. In R. Millar, J. Leach & J. Osborne (Eds.), *Improving science education* (pp. 127-143). Buckingham, PA: Open University Press.

Nitz, S., Nerdel, C., & Prechtl, H. (2012). Die Verwendung von Fachsprache im Biologieunterricht: Entwicklung eines Erhebungsinstruments. *Zeitschrift für Didaktik der Naturwissenschafte, 18,* 117-139.

Nitz, S., Prechtl, H., & Nerdel, C. (in press). Survey of classroom use of representations: Development, field test and multilevel analysis. *Learning Environments Research*.

Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education, 87*, 224-240. doi: 10.1002/sce.10066

Prain, V., & Waldrip, B. G. (2006). An exploratory study of teachers' and students' use of multi-modal representations of concepts in primary science. *International Journal of Science Education, 28*, 1843 -1866. doi: 10.1080/09500690600718294

Schnotz, W. (2001). Sign systems, technologies, and the acquisition of knowledge. In J.-F. Rouet, J. Levonen & A. Biardeau (Eds.), *Multimedia Learning: Cognitive and Instructional Issues* (pp. 9-29). Amsterdam, The Netherlands: Pergamon.

Schroeder, S., Richter, T., McElvany, N., Hachfeld, A., Baumert, J., Schnotz, W. … Ullrich, M. (2011). Teachers' beliefs, instructional behaviors, and students' engagement in learning from texts with instructional pictures. *Learning and Instruction, 21*, 403-415. doi: 10.1016/j.learninstruc.2010.06.001

Scott, P. (1998). Teacher talk and meaning making in science classrooms: a Vygotskian analysis and review. *Studies in Science Education, 32*, 45-80. doi: 10.1080/03057269808560127

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*(4), 454-499. doi: 10.3102/0034654307310317

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London, England: Sage.

Stieff, M. (2011). Improving representational competence using molecular simulations embedded in inquiry activities. *Journal of Research in Science Teaching, 48*, 1137–1158. doi: 10.1002/tea.20438

Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying representational competence with multi-representational displays. *Cognition and Instruction, 29*, 123-145. doi: 10.1080/07370008.2010.507318

Taskin, V., & Bernholt, S: (2012). Students' understanding of chemical formulae: A review of empirical research. *International Journal of Science Education*, 1-29. doi: 10.1080/09500693.2012.744492

Tytler, R., Prain, V., & Peterson, S. (2007). Representational issues in students learning about evaporation. *Research in Science Education, 37*, 313-331. doi: 10.1007/s11165-006-9028-3

Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology, 31*, 142-166. doi: 10.1016/j.cedpsych.2005.04.001

Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review, 17*, 285-325. doi: 10.1007/s10648-005-8136-3

Wagner, W., Göllner, R., Helmke, A., Trautwein, U., & Lüdtke, O. (2013). Construct validity of student perceptions of instructional quality is high, but not perfect: Dimensionality and generalizability of domain-independent assessments. *Learning and Instruction, 28*, 1-11. doi: 10.1016/j.learninstruc.2013.03.003

Waldrip, B. G., Prain, V., & Carolan, J. (2010). Using multi-modal representations to improve learning in junior secondary science. *Research in Science Education, 40*, 65-80. doi: 10.1007/s11165-009-9157-6

Wellington, J., & Osborne, J. (2001). *Language and literacy in science education*. Buckigham, PA: Open University Press.

Wu, H.-K., & Puntambekar, S. (2012). Pedagogical affordances of multiple external representations in scientific processes. *Journal of Science Education and Technology, 21*(6), 754-767. doi: 10.1007/s10956-011-9363-7

Yore, L. D., & Hand, B. (2010). Epilogue: Plotting a research agenda for multiple representations, multiple modality, and multimodal representational competency. *Research in Science Education, 40*, 93-101. doi: 10.1007/s11165-009-9160-y

Yore, L. D., Pimm, D., & Tuan, H.-L. (2007). The literacy component of mathematical and scientific literacy. *International Journal of Science and Mathematics Education, 5*, 559-589. doi: 10.1007/s10763-007-9089-4

Footnotes

[1] Realistic and logical pictures are visual-graphical representations. These are spatial configurations that represent a subject with structural similarities between the object and the representation. In realistic pictures, such as photographs and drawings, the similarity between the object and the representation is concrete. In logical pictures, such as diagrams and graphs, the similarity is abstract (Schnotz 2001).

[2]These are the most prevalent modes in teaching and learning photosynthesis in German biology class, the context of our study (Nitz et al., 2012)