Bridging the Gap between Expert-Novice Differences: The Model-Based Feedback Approach

Dirk Ifenthaler
University of Mannheim

Abstract

The study adds to the body of knowledge about different types of feedback. Feedback is considered a fundamental component for supporting and regulating learning processes. Especially in computer-based and self-regulated learning environments, the nature of feedback is of critical importance. Hence, this study investigates different types of automatically generated feedback. Seventy-four students participated in the experimental study, in which they had to write texts and create concept maps reflecting their understanding of climate change while receiving three different forms of model-based feedback. Results indicate that concept maps were structurally and semantically more similar to an expert solution than written texts. However, the form of the automatically created model-based feedback did not influence the type of representation. (Keywords: Feedback, expertise, mental model, concept map, Highly Integrated Model Assessment Technology and Tools, HIMATT)

Feedback is any type of information provided to learners (Wagner & Wagner, 1985). Accordingly, feedback can take many forms, depending on theoretical perspective, learning and instructional goals, objectives, research purposes, and methodological approaches. Moreover, feedback is considered a fundamental component for supporting and regulating learning processes. Especially in computer-based and self-regulated learning environments, the nature of feedback plays a critical role in learning and instruction, especially in computer-based and self-regulated learning environments (Simons & de Jong, 1992).

Unlike this initial general understanding of feedback, the term informative feedback refers to all kinds of external postresponse information used to inform the learner of his or her current state of learning or performance (Narciss, 2006, 2008). Furthermore, from an instructional point of view, feedback can be provided by internal (individual cognitive monitoring processes) or external (various types of correction variables) sources of information. Internal feedback may validate the externally provided feedback, or it may lead to resistance against it (Narciss, 2008). However, the empirical evidence regarding the effects of different types of feedback is rather inconsistent and somewhat contradictory (e.g. Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Clariana, 1993; Kluger & DeNisi, 1996; Kulhavy, 1977; Mory, 2004).

Feedback on mental model construction, such as the use of conceptual models to help build mental models of the system being studied, has also been investigated and discussed (see, for example, Mayer, 1989). Conceptual models highlight the most important objects and associated causal relations of the phenomenon in question. However, not only do new developments in computer technology enable us to dynamically generate simple conceptual models and expert representations, but they may also be used to generate direct responses to the learner's interaction with the learning environment. We define this form of feedback as model-based feedback (Ifenthaler, 2009).

The current research aims to add to the body of knowledge about different types of feedback and perhaps resolve some of the inconsistencies. This paper reports the investigation of different types of automated model-based feedback. The first section highlights the underlying theoretical framework of model-based feedback and the research questions. The next section presents the research design used to investigate the effects of model-based feedback using different methods (concept mapping and written text) for presenting the solution of a task to be solved.

Theoretical Framework

Numerous studies in the field of educational research have provided evidence that "mental models guide and regulate all human perceptions of the physical and social world" (Seel & Dinter, 1995, p. 5). Mental models are dynamic ad hoc constructions that provide subjectively plausible explanations on the basis of restricted domain-specific information (Johnson-Laird, 1989; Seel, 1991). Research studies have shown that it is very difficult but possible to influence such subjectively plausible mental models by providing specific information (see Anzai & Yokoyama, 1984; Ifenthaler, Masduki, & Seel, 2009; Mayer, 1989; Seel, 1995; Seel & Dinter, 1995). Ifenthaler and Seel (2005) argue that it is important to consider how such information is provided to the learner at specific times during the learning process and how it is structured. In accordance with the general definition of feedback introduced above (Wagner & Wagner, 1985), an important aspect of model-based feedback is providing dynamic feedback generated purposively and individually to student-constructed models (Ifenthaler, 2009).

Model-Based Feedback

The importance of feedback for improving knowledge and skill acquisition has been controversial in educational research (e.g., Azevedo & Bernard, 1995; Bangert-Drowns et al., 1991; Narciss, 2008; Narciss & Huth, 2004; Shute, 2008). Widely accepted forms of feedback include (a) knowledge of result, (b) knowledge of correct result, (c) knowledge of performance, (d) answer until correct, (e) knowledge of task constraints, (f) knowledge

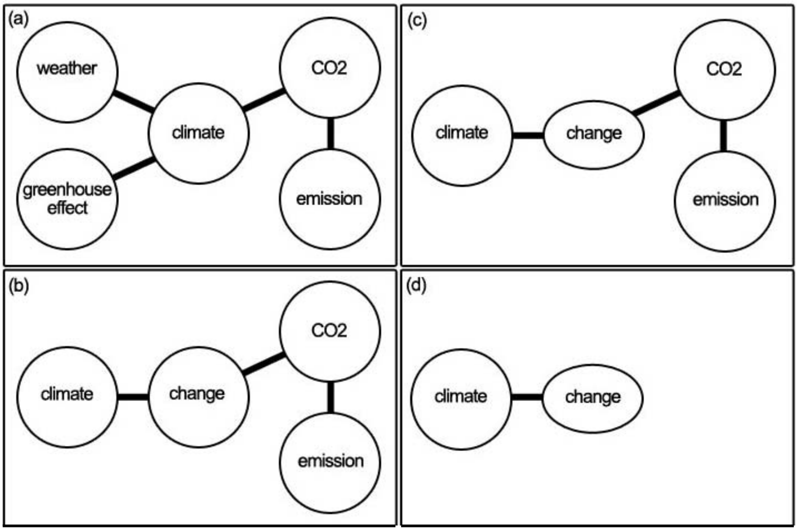

Figure 1: (a) Reference, (b) subject, (c) cutaway, and (d) discrepancy re-representations.

about concepts, (g) knowledge about mistakes, (h) knowledge about how to proceed, and (i) knowledge about metacognition (see Jacobs, 1998; Narciss, 2008). Additionally, Schimmel (1983) found that feedback is most effective under conditions that encourage the learner's conscious reception and engage the learner in reflecting on the response.

In accordance with empirical findings on feedback (Schimmel, 1983) and mental model theory (Ifenthaler, Pirnay-Dummer, & Spector, 2008), we argue that effective model-based feedback is composed of externalized representations (re-representations) of mental models. An externalization of a mental model of a learner or expert may be a causal model or a concept map, and it might involve written or spoken text as well as drawings and annotated diagrams (Ifenthaler, 2008). Such externalized representations induce positive effects on internal information processing (Galbraith, 1999). Additionally, model-based feedback aims to develop mental models for the improvement of expertise and expert performance (Johnson-Laird, 1989). Accordingly, model-based feedback is associated with the development of expertise and expert performance in a specific subject or task domain.

Past research studies have shown how providing conceptual models (i.e., explicit and consistent causal explanations of a given phenomenon) can improve one's understanding of a specific problem in a given context (Mayer, 1989; Norman, 1983; Seel & Dinter, 1995). However, we argue that model-based feedback should include more than an expert's solution to a given phenomenon. To be more effective, the feedback should also take into account the person's prior understanding (initial mental model, preconception), because such preconceptions are in many cases resistant to change, as they have a high subjective plausibility (Ifenthaler & Seel, 2005; Seel & Dinter, 1995).

To fulfill this requirement, this article introduces two new forms of model-based feedback: (a) cutaway model-based feedback and (b) discrepancy model-based feedback. These two forms of model-based feedback are graphical re-representations constructed from a set of vertices whose relationships are represented by edges (Ifenthaler, 2010b; Ifenthaler et al., 2009).

The cutaway model-based feedback is based on the individual's preconception or on a more elaborated mental model constructed during the learning process. Additionally, an expert's understanding of the phenomenon in question is taken into account. Combining both the individual's re-representation (preconception) and the expert's re-representation creates the cutaway model-based feedback re-representation. This re-representation includes all propositions (vertex-edge-vertex) of the individual's re-representation and highlights semantically correct vertices based on comparison with an expert's re-representation (see Figure 1c).

The discrepancy model-based feedback is also based on the individual's preconception or on a more elaborate mental model constructed during the learning process. However, it includes only the propositions (vertex-edge-vertex) that have no semantic similarity to the expert's re-representation. Additionally, semantically correct vertices (based on a comparison with an expert's re-representation) are highlighted (see Figure 1d).

Hence, model-based feedback aims to restructure the underlying representations and a reconceptualization of the related concepts (vertices and edges). This is in keeping with Piaget's epistemology (1950, 1976). New information provided through model-based feedback can be assimilated through the activation of an existing schema, adjustment by accretion, or tuning of existing schema. Otherwise, it is accommodated by means of a reorganization process that involves building new mental models (Ifenthaler et al., 2009; Seel, Ifenthaler, & Pirnay-Dummer, 2009).

To fulfill the requirement of providing model-based feedback to the learner on the fly, it is necessary to implement the cutaway and discrepancy model-based feedback in a computer-based environment. Accordingly, the following section describes the automated model-based feedback generation.

Model-Based Feedback Generation

The model-based feedback generation is realized within the Highly Integrated Model Assessment Technology and Tools (HIMATT) environment. HIMATT is a combined toolset that conveys the benefits of various methodological approaches in a single Web-based environment. It is implemented and runs on a Web server using Apache, MySQL, PERL, and additional packages (Pirnay-Dummer, Ifenthaler, & Spector, 2010). The HIMATT architecture consists of two major platforms: the HIMATT Research Engine (functions for conducting and analyzing experiments) and the HIMATT Subject Environment (functions for dynamically providing assigned experiments to individual subjects). Text and conceptual graphs (e.g., concept maps) can be analyzed quantitatively with the automated comparison functions. Additionally, Ifenthaler (2010a) introduced an automated feature to generate standardized graphical re-representations of subjects' data with the help of the open source graph visualization software GraphViz (Ellson, Gansner, Koutsofios, North, & Woodhull, 2003). This algorithm enables the generation of domain-specific automated model-based feedback.

The model-based feedback function automatically generates standardized reference (e.g., expert), participant (e.g., learner), cutaway, and discrepancy re-representations. A cutaway re-representation includes all propositions (vertex-edge-vertex) of the individual's re-representation. Additionally, the semantically correct vertices (compared to a reference re-representation such as an expert solution) are graphically highlighted as circles (ellipses for dissimilar vertices). The discrepancy re-representation of an individual only includes propositions (vertex-edge-vertex) that have no semantic similarity to a reference re-representation. Additionally, the semantically correct vertices (compared to a reference re-representation) are graphically highlighted as circles (ellipses for dissimilar vertices). Figure 1 provides examples of (a) simplified reference, (b) participant, (c) cutaway, and (d) discrepancy re-representations. The model-based feedback function generates these automated and standardized re-representations on the fly while participants work within the HIMATT environment. They are then used for individual model-based feedback during work on a learning task.

The reference model (see Figure 1a) represents an expert's best-practice solution to complete the task. The participant's model (b) is a solution found after a specified time working on the task. With the reference (a) and participant (b) models at hand, HIMATT automatically generates the cutaway (c) and discrepancy (d) feedback models. The cutaway model allows the learner to see how many vertices are semantically correct (graphically highlighted circles compared to the expert solution). Additionally, the cutaway model provides information on the semantically incorrect vertices (ellipses). The discrepancy model provides information only on the semantically incorrect propositions as compared to the expert solution (vertex-edge-vertex). Additionally, semantically correct vertices are highlighted. We argue that either feedback model (c or d) will have different effects when presented during the learning process. As the cutaway feedback model (c) helps to confirm the correct understanding of the phenomenon in question (compared with an expert), the discrepancy feedback model (d) causes a cognitive conflict, because correct propositions (vertex-edge-vertex) of the person's understanding are deleted from the re-representation.

Purpose of the Study

Each of the above described feedback models could help to improve expertise and expert performance in various subject domains. In line with the experimental study mentioned above, this paper investigates the effects of model-based feedback using different methods (concept mapping and written text) for presenting the solution of a task to be solved.

Usually, learners have different options for externalization at their disposal using language (spoken and written) or graphical notes (abbreviations) of various commonly known (mind mapping, concept mapping) or idiosyncratic formats (Ifenthaler, 2008; Seel, 1999). Accordingly, this study examines two specific sources of externalized knowledge, written text and concept maps, and examine how model-based feedback influences these forms of externalization. Accordingly, the first research question was: Are there differences between learners' written texts and concept maps before a learning intervention compared to an expert's representation?

Furthermore, feedback plays a particularly important role in highly self-regulated model-centered learning environments because it facilitates the development of mental models, thus improving expertise and expert performance (Ifenthaler, 2009; Ifenthaler & Seel, 2005). Past research studies have demonstrated how conceptual models can be provided to improve a person's understanding of a specific problem in a given context (Mayer, 1989; Norman, 1983; Seel & Dinter, 1995). Conversely, model-based feedback includes not only a conceptual or expert solution to the given phenomenon, but also the person's prior understanding (i.e., initial mental model, preconception). Therefore, Ifenthaler (2009) introduced three forms of model-based feedback: (a) cutaway model-based feedback, (b) discrepancy model-based feedback, and (c) expert feedback (as control group). Thus, the second research question was: Do different forms of model-based feedback (cutaway, discrepancy, and expert) influence the forms of externalized understanding (written text and concept map) of a specific phenomenon in different ways?

Because this study assesses both written text and concept maps on the same subject domain and at the same point in time, we expect that these different forms of externalization will represent the same structural and semantic content. More specifically, due to the close assessment and the short time between writing and concept mapping, we expected a close match between the structural and semantic HIMATT measures (see Pirnay-Dummer et al., 2010; a description of all of the applied measures will be provided in the following section) for both externalizations. Accordingly, the third research question investigated in this article was: Do written text and concept maps represent the same structural and semantic content?

Method

Participants

To test the effects of three types of model-based feedback (cutaway, discrepancy, expert), we conducted an experimental study. Students from a German university (N = 74) participated in this experiment (average age = 21.9, SD = 2.3; 66 female, 8 male). The participants were randomly assigned to one of the three experimental groups: (a) cutaway feedback ($n = 26$), (b) discrepancy feedback ($n = 24$), and (c) expert feedback ($n = 24$).

Instruments and Materials

The learning content was a German-language article on climate change (Schönwiese, 2005) with 1,417 words. The article included tables and figures that were closely related to the subject domain.

To assess the participants' understanding of the subject domain, climate change, we used the HIMATT concept map and text input tools. All participants' concept maps and written texts were automatically stored in the HIMATT database for further analysis. Additionally, the domain specific knowledge test included 27 multiple-choice questions on climate change. In a pilot study with five female und five male participants (average age = 26.3, SD = 3.49), we tested the average difficulty level to account for ceiling effects. The participants scored 10.5 out of 27 possible points on average (SD = 3.54, Min = 5, Max = 17). In the experiment, we administered two versions (in which the 27 multiple-choice questions appeared in a different order) of the domain-specific knowledge test (pre- and posttest). It took about 10 minutes to complete the test.

Additional instruments included two subsets of the I-S-T 2000 R (Amthauer, Brocke, Liepmann, & Beauducel, 2001) to test the participants' verbal and spatial abilities. This test is a widely used intelligence test in Germany with high reliability ($r = .88$ and $r = .96$; split-half reliability). We tested the participants' experience with concept mapping with a questionnaire comprised of eight items (Cronbach's alpha = .87). The participants answered the questions on a 5-point Likert scale (1 = totally disagree, 2 = disagree, 3 = partially agree, 4 = agree, 5 = totally agree). The feedback model quality test consisted of nine items meant to ascertain whether the provided feedback model helped the participant understand the text better (Cronbach's alpha = .66). The participants answered the questions on a 5-point Likert scale (1 = totally disagree, 2 = disagree, 3 = partially agree, 4 = agree, 5 = totally agree).

Procedure

First, the participants completed a demographic data questionnaire. Second, they completed the concept map and causal diagram experience questionnaire. Next, the participants completed the test on verbal (six minutes) and spatial abilities (nine minutes). Then they answered the 27 multiple choice questions

Table 1: Detailed Information about the Seven HIMATT Measures

Measure and Type	Description
Surface Matching *Structural Indicator*	The surface matching (Ifenthaler, 2010a) compares the number of vertices within two graphs. It is a simple and easy way to calculate values for surface complexity.
Graphical Matching *Structural Indicator*	The graphical matching (Ifenthaler, 2010a) compares the diameters of the spanning trees of the graphs, which is an indicator for the range of conceptual knowledge. It corresponds to structural matching, as it is also a measure for structural complexity only.
Structural Matching *Structural Indicator*	The structural matching (Pirnay-Dummer & Ifenthaler, 2010) compares the complete structures of two graphs without regard to their content. This measure is necessary for all hypotheses that make assumptions about general features of structure (e.g. assumptions that state that expert knowledge is structured differently from novice knowledge).
Gamma Matching *Structural Indicator*	The gamma or density of vertices (Pirnay-Dummer & Ifenthaler, 2010) describes the quotient of terms per vertex within a graph. Because both graphs connect every term with each other term (everything with everything) and graphs that connect only pairs of terms can be considered weak models, a medium density is expected for most good working models.
Concept Matching *Semantic Indicator*	Concept matching (Pirnay-Dummer & Ifenthaler, 2010) compares the sets of concepts (vertices) within a graph to determine the use of terms. This measure is especially important for different groups that operate in the same domain (e.g. use the same textbook). It determines differences in language use between the models.
Propositional Matching *Semantic Indicator*	The propositional matching (Ifenthaler, 2010a) value compares only fully identical propositions between two graphs. It is a good measure for quantifying semantic similarity between two graphs.
Balanced Propositional Matching *Semantic Indicator*	The balanced propositional matching (Pirnay-Dummer & Ifenthaler, 2010) is the quotient of propositional matching and concept matching. Especially when both indices are being interpreted, balanced propositional matching should be preferred over propositional matching.

of the domain-specific knowledge test on climate change (pretest). After a short relaxation phase, the participants received an introduction to concept maps and causal diagrams and learned how to use the HIMATT software. Then, the participants used the username and password they had been assigned to log in to the HIMATT system, where they constructed a concept map on their understanding of climate change (10 minutes). Immediately afterward, they wrote a text about their understanding of climate change (10 minutes). A short relaxation phase followed, during which we automatically generated the individual feedback models for each participant. After that, the participants received the text on climate change and the automatically generated feedback model (cutaway, discrepancy, or expert model, depending on which experimental group they had been assigned to). All three types of feedback models were automatically generated with HIMATT. The cutaway feedback model included all propositions (vertex-edge-vertex) of the participant's pretest causal diagram. Additionally, the semantically correct vertices (compared to the expert re-representation) were graphically highlighted (circles are semantically correct, ellipses semantically incorrect as compared to the expert re-representation). The discrepancy feedback model included only propositions (vertex-edge-vertex) of the participant's pretest causal diagram that had no semantic similarity to the expert re-representation. The expert feedback model consisted of an expert's standardized re-representation on climate change.

The participants had 15 minutes to read the text and examine their feedback model. Immediately after working on the text, the participants completed the model feedback quality test. Then they answered the 27 multiple choice questions of the posttest on declarative knowledge. After another short relaxation phase, the participants used their username and password to log in to the HIMATT system for the second time. In the HIMATT posttest, they constructed a second concept map on their understanding of climate change (10 minutes) and wrote a second text regarding their understanding of climate change (10 minutes). Finally, the participants had to complete a short usability test on HIMATT features (Pirnay-Dummer et al., 2010).

Analysis

To analyze the concept maps and written text the participants created in the HIMATT environment, we used the seven core measures implemented in HIMATT (Pirnay-Dummer et al., 2010). Table 1 describes the seven measures of HIMATT, which include four structural indicators and three semantic indicators (Ifenthaler, 2010a; Pirnay-Dummer et al., 2010).

Reliability measures exist for the single indicators integrated into HIMATT. They range from $r = .79$ to $r = .94$ and are tested for the semantic and structural indicators separately and across different knowledge domains (Pirnay-Dummer et al., 2010). Validity measures are also reported separately for the structural and semantic indicators. Convergent validity lies between $r = .71$ and $r = .91$ for semantic comparison indices and between $r = .48$ and $r = .79$ for structural comparison indicators (Pirnay-Dummer et al., 2010).

Results

More than two thirds of the participants (68%) did not use concept maps or causal diagrams to structure their own learning materials before our experiment. Only 12% of the participants used concept mapping software to create their own concept maps beforehand. On the other hand, more than 40% of the participants answered that they did not find it difficult to create a concept map or causal diagram. Consequently, there was no significant difference in learning outcomes as measured by the domain-specific knowledge posttest between participants who used concept mapping software before the experiment and those who did not use concept mapping software at all, $t(72) = .508$, *ns*.

Domain-Specific Knowledge

On the domain-specific knowledge test (pre- and posttest), participants could score a maximum of 27 correct answers. On the pretest they scored an average of M = 7.78 correct answers (SD = 2.10), and in the posttest they scored an average of M = 18.16 correct answers (SD = 3.80). The increase in correct answers was significant, $t(73) = 28.32$, $p < .001$, $d = 3.096$ (strong

Table 2: Average Similarity Scores for Written Text and Concept Maps of Participants Compared to the Expert's Representation (N = 74)

	Written Text	SD	Concept Map	SD
Surface Matching	.494	.283	.338	.114
Graphical Matching	.592	.296	.607	.162
Structural Matching	.674	.208	.738	.125
Gamma Matching	.517	.293	.769	.193
Concept Matching	.084	.057	.124	.059
Propositional Matching	.006	.017	.007	.018
Balanced Propositional Matching	.040	.108	.044	.111

Note: 0 = no similarity, 1 = total similarity

effect). The cutaway feedback group (M = 10.88, SD = 3.32) scored higher than the discrepancy (M = 10.42, SD = 2.92) and expert group (M = 9.79, SD = 3.23) in knowledge gain. However, these differences were not significant.

Expert-Novice Differences between Written Text and Concept Maps

To answer the first research question, HIMATT analysis feature (see above) automatically compared the written text and concept maps the participants constructed during the pretest to an expert representation. Hence, for both written text and concept maps, seven similarity scores (0 = no similarity; 1 = total similarity; for measures surface, graphical, structural, gamma, concept, propositional, and balanced propositional matching) are available for further statistical analysis. To identify possible expert-novice differences between written text and concept maps, we computed paired-sample t-tests for the seven HIMATT similarity scores (see Table 2).

The seven paired-sample t-tests revealed significant differences between written text and concept maps for the HIMATT measures surface matching, $t(73) = 4.05$, $p < .001$, $d = .666$ (medium effect), structural matching, $t(73) = -2.12$, $p = .038$, $d = .349$ (small effect), gamma matching, $t(73) = -5.77$, $p < .001$, $d = .949$ (strong effect), and concept matching, $t(73) = -4.54$, $p < .001$, $d = .746$ (medium effect).

Effects of Model-Based Feedback

Regarding the second research question, we investigated the influence of different types of feedback (cutaway, discrepancy, expert feedback) on the two forms of externalization (written text and concept map). Table 3 shows the average change (pre/posttest) of the HIMATT similarity scores (surface, graphical, structural, gamma, concept, propositional, and balanced propositional matching) for the three experimental groups (cutaway feedback, discrepancy feedback, and expert feedback), separated for written text and concept map (type of representation).

We conducted seven separate multivariate analyses of variance (MANOVA) with experimental group (cutaway feedback, discrepancy feedback,

Table 3: Average Change (Pre/Posttest) of Similarity Scores for Written Text and Concept Map of Participants Compared to the Expert's Representation (N = 74)

	Experimental Group	Written Text	SD	Concept Map	SD
	CF	-.102	.362	.028	.121
Surface Matching	DF	-.117	.366	.090	.100
	EF	-.004	.428	.084	.119
	CF	.023	.466	.035	.169
Graphical Matching	DF	-.124	.421	.068	.144
	EF	.074	.462	.104	.191
	CF	-.024	.273	.006	.152
Structural Matching	DF	-.081	.260	.004	.162
	EF	-.033	.335	.073	.152
	CF	.014	.478	.066	.144
Gamma Matching	DF	-.049	.332	.063	.216
	EF	-.005	.254	.022	.201
	CF	.066	.067	.044	.062
Concept Matching	DF	.027	.071	.069	.077
	EF	.057	.082	.051	.088
	CF	.001	.017	.007	.030
Propositional Matching	DF	-.001	.017	.003	.029
	EF	-.001	.018	.004	.015
	CF	-.003	.105	.012	.179
Balanced Propositional Matching	DF	-.005	.095	.017	.161
	EF	-.014	.134	.017	.089

Note: 0 = no similarity, 1 = total similarity

Note: EXPERIMENTAL GROUP: CF = cutaway feedback (n = 26), DF = discrepancy feedback (n = 24), EF = expert feedback (n = 24)

expert feedback) and type of representation (written text or concept map) as between-subject factors. The average change in the seven HIMATT similarity scores (surface, graphical, structural, gamma, concept, propositional, and balanced propositional matching) served as the dependent measure.

The seven MANOVAs showed no main effects of the experimental groups or type of externalization on the average change in the HIMATT similarity scores. However, we found a main effect for type of externalization on the average change of the surface matching similarity score, $F\ (1, 142) = 5.556$, $p = .020$, $\eta^2 = .038$. The concept maps that the participants constructed became more similar to the expert's representation than the written text. Accordingly, the form of model-based feedback did not influence the type of externalization.

Comparison between Written Text and Concept Maps

To address the third research question, we analyzed the written text and concept maps from the posttest to identify structural and semantic similarities or differences between these two forms of externalization.

Table 4: Average Similarity between Individuals' Written Text and Concept Map for HIMATT Measures (N = 74)

	M	SD	Min	Max	KS-Z	*p*
Surface Matching	.280	.155	.07	.80	1.136	.151
Graphical Matching	.601	.258	.11	1.00	.932	.350
Structural Matching	.608	.175	.22	1.00	1.173	.127
Gamma Matching	.525	.242	.00	1.00	.339	1.000
Concept Matching	.179	.104	.00	.45	.678	.748
Propositional Matching	.015	.031	.00	.14	3.662	.000**
Balanced Propositional Matching	.062	.123	.00	.48	3.690	.000**

Note: Min = 0, Max = 1

*Note: KS-Z = Kolmogorov-Smirnov one-sample test; * $p < .05$; ** $p < .01$*

Table 4 shows the descriptive measures and the results of the Kolmogorov-Smirnov one-sample tests. We found no interindividual differences between the subjects, except for in the measures propositional matching and balanced propositional matching. Interestingly, written text and concept maps seem to represent different structure and content, because we found only low similarities between the seven HIMATT measures. Exceptions are the structural measures graphical matching, structural matching, and gamma matching. Here we found more than 50% similarity, indicating that the structural complexity of the representations is closely connected. Additionally, to locate differences between the three experimental groups (cutaway feedback, discrepancy feedback, expert feedback), we computed conservative Kruskal-Wallis H-tests. However, we did not find significant differences for any of the HIMATT measures.

Discussion

This study examined three forms of model-based feedback using different methods (concept mapping and written text) for presenting the solution of a task to be solved. We introduced new forms of model-based feedback, which we refer to as (a) cutaway model-based feedback and (b) discrepancy model-based feedback. As we were able to generate the model-based feedback automatically and on the fly, the participants received it just after finishing their pretests, which served to motivate them further. Additionally, our HIMATT analysis features enabled us to score the participants' solutions automatically within an instant. Not only do these automated process have very high objectivity, reliability, and validity (Pirnay-Dummer et al., 2010), they are also very economical, especially when large data sets need to be analyzed within a short period of time (Ifenthaler, 2010a; Johnson, Ifenthaler, Pirnay-Dummer, & Spector, 2009).

First, we looked at two specific sources of externalized knowledge, written text and concept maps, and examined how model-based feedback influences these forms of externalization. Accordingly, we looked for differences between learners' written texts and concept maps before a learning intervention compared to an expert's representation. Our results revealed significant differences between written text and concept maps for the HIMATT measures surface matching, structural matching, gamma matching, and concept matching.

Regarding the surface matching measure, it appears that the number of vertices in the written text is closer to the expert's representation than the number of vertices in the concept map; i.e., written text covers a broader area of the phenomenon in question. On the other hand, the concept map representations are significantly more closely related (structural matching and gamma matching) to the expert's representation than the written text representations. Additionally, more correct concepts were included in the concept maps (compared to the expert representation) than in the written text (concept matching). These findings suggest that concept maps and written texts represent different things, even when used in the same task. Further studies will have to be conducted to strengthen the theorem that concept maps and written texts have different dimensions. But the results clearly show that concept mapping techniques and systematic text analysis cannot be used as complements.

Second, we investigated whether the form of model-based feedback influences the type of representation differently. Our in-depth analysis showed only one significant effect for the surface matching similarity score. Here, the concept maps became more similar to the expert's representation than the written text in number of relations. As the semantic content showed no significant effects, we conclude that participants added more concepts and relations in the posttest, which also could contain several misconceptions.

Third, we looked at the structural and semantic similarities between the written text and concept maps of the posttest. Here, we found that the written text and concept maps represent different structure and content. Hence, the type of externalization strategy also influences the represented knowledge (structurally and semantically). These findings may have a large impact on future knowledge diagnosis and should be investigated in a future experimental study. Accordingly, this study will compare graphical representations and written texts from similar tasks but on completely different subject domains. If the structures turn out to be nonselective for any content or knowledge type, we assume that concept maps are better for rebuilding a concise structure. If they are selective, we assume that concept maps always have similar structures because their users apply only a certain range of structural patterns and therefore constrain themselves to the available patterns.

To conclude, in further studies we will focus on learning trajectories while providing forms of model-based feedback on different types of externalizations. This will provide more detailed insight into the effects of model-based feedback and the role of externalization and how it helps bridge the gap between expert–novice differences.

Author Note

Dirk Ifenthaler is temporary chair and professor of educational technology and learning sciences in the Department of Educational Science at the University of Mannheim, Germany. His research interests include technology and learning, mental models, problem solving, and longitudinal analysis methodologies. Correspondence regarding this article should be sent to Dirk Ifenthaler, University of Mannheim, A5, 6, 68159 Mannheim, Germany. E-mail: ifenthaler@uni-mannheim.de

References

Amthauer, R., Brocke, B., Liepmann, D., & Beauducel, A. (2001). *I-S-T 2000 R.* Göttingen: Hogrefe.

Anzai, Y., & Yokoyama, T. (1984). Internal models in physics problem solving. *Cognition and Instruction, 1*(4), 397–450.

Azevedo, R., & Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research, 13*(2), 111–127.

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213–238.

Clariana, R. B. (1993). A review of multiple-try feedback in traditional and computer-based instruction. *Journal of Computer-Based Instruction, 20*(3), 67–74.

Ellson, J., Gansner, E. R., Koutsofios, E., North, S. C., & Woodhull, G. (2003). GraphViz and Dynagraph. *Static and dynamic graph drawing tools.* Florham Park, NJ: AT&T Labs.

Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Knowing what to write. Conceptual processes in text production* (pp. 139–160). Amsterdam: University Press.

Ifenthaler, D. (2008). Practical solutions for the diagnosis of progressing mental models. In D. Ifenthaler, P. Pirnay-Dummer, & J. M. Spector (Eds.), *Understanding models for learning and instruction. Essays in honor of Norbert M. Seel* (pp. 43–61). New York: Springer.

Ifenthaler, D. (2009). Model-based feedback for improving expertise and expert performance. *Technology, Instruction, Cognition, and Learning, 7*(2), 83–101.

Ifenthaler, D. (2010a). Relational, structural, and semantic analysis of graphical representations and concept maps. *Educational Technology Research and Development, 58*(1), 81–97. DOI: 10.1007/s11423-008-9087-4

Ifenthaler, D. (2010b). Scope of graphical indices in educational diagnostics. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge* (pp. 213–234). New York: Springer.

Ifenthaler, D., Masduki, I., & Seel, N. M. (2009). The mystery of cognitive structure and how we can detect it. Tracking the development of cognitive structures over time. *Instructional Science*. DOI: 10.1007/s11251-009-9097-6

Ifenthaler, D., Pirnay-Dummer, P., & Spector, J. M. (Eds.). (2008). *Understanding models for learning and instruction. Essays in honor of Norbert M. Seel.* New York: Springer.

Ifenthaler, D., & Seel, N. M. (2005). The measurement of change: Learning-dependent progression of mental models. T*echnology, Instruction, Cognition, and Learning, 2*(4), 317–336.

Jacobs, B. (1998). *Aufgaben stellen und Feedback geben.* Retrieved October 6, 2008, from http://www.phil.uni-sb.de/~jakobs/wwwartikel/feedback/index.htm

Johnson-Laird, P. N. (1989). Mental models. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 469–499). Cambridge, MA: MIT Press.

Johnson, T. E., Ifenthaler, D., Pirnay-Dummer, P., & Spector, J. M. (2009). Using concept maps to assess individuals and team in collaborative learning environments. In P. L. Torres & R. C. V. Marriott (Eds.), *Handbook of research on collaborative learning using concept mapping* (pp. 358–381). Hershey, PA: Information Science Publishing.

Kluger, A. N., & DeNisi, A. (1996). Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284.

Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research, 47*(2), 211–232.

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research, 59*(1), 43–64.

Mory, E. H. (2004). Feedback research revisited. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 745–783). Mahwah, NJ: Lawrence Erlbaum.

Narciss, S. (2006). *Informatives tutorielles Feedback. Entwicklung- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse.* Münster: Waxmann.

Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. van Merrienboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (pp. 125–143). New York: Taylor & Francis Group.

Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. In H. M. Niegemann, D. Leutner, & R. Brünken (Eds.), *Instructional design for multimedia learning* (pp. 181–195). Münster: Waxmann.

Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 7–14). Hilsdale, NJ: Lawrence Erlbaum Associates.

Piaget, J. (1950). *La construction du réel chez l'enfant.* Neuchatel: Delachaux et Niestlé S.A.

Piaget, J. (1976). *Die Äquilibration der kognitiven Strukturen.* Stuttgart: Klett.

Pirnay-Dummer, P., & Ifenthaler, D. (2010). Automated knowledge visualization and assessment. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge* (pp. 77–115). New York: Springer.

Pirnay-Dummer, P., Ifenthaler, D., & Spector, J. M. (2010). Highly integrated model assessment technology and tools. *Educational Technology Research and Development, 58*(1), 3–18. DOI: 10.1007/s11423-009-9119-8

Schimmel, B. J. (1983). *A meta-analysis of feedback to learners in computerized and programmed instruction.* Paper presented at the AREA 1983, Montreal.

Schönwiese, C.-D. (2005). Klimawandel—Tatsache oder Fiktion? *Energiewirt, 104*, 26–29.

Seel, N. M. (1991). *Weltwissen und mentale Modelle.* Göttingen: Hogrefe.

Seel, N. M. (1995). Mental models, knowledge transfer, and teaching strategies. *Journal of Structural Learning and Intelligent Systems, 12*(3), 197–213.

Seel, N. M. (1999). Educational semiotics: School learning reconsidered. *Journal of Structural Learning and Intelligent Systems, 14*(1), 11–28.

Seel, N. M., & Dinter, F. R. (1995). Instruction and mental model progression: Learner-dependent effects of teaching strategies on knowledge acquisition and analogical transfer. *Educational Research and Evaluation, 1*(1), 4–35.

Seel, N. M., Ifenthaler, D., & Pirnay-Dummer, P. (2009). Mental models and problem solving: Technological solutions for measurement and assessment of the development of expertise. In P. Blumschein, W. Hung, D. H. Jonassen, & J. Strobel (Eds.), *Model-based approaches to learning: Using systems models and simulations to improve understanding and problem solving in complex domains* (pp. 17–40). Rotterdam: Sense Publishers.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189.

Simons, P. R. J., & de Jong, F. P. C. M. (1992). Self-regulation and computer-aided instruction. *Applied Psychology: An International Review, 41*(4), 333–346.

Wagner, W., & Wagner, S. U. (1985). Presenting questions, processing responses, and providing feedback in CAI. *Journal of Instructional Development, 8*(4), 2–8.