

Uncovering learning outcomes: explicating obscurity in learning of aesthetics in design and technology education

Grietjie Haupt and Seugnet Blignaut

Abstract

Education and training interventions can be evaluated through the success of learning outcomes. Kirkpatrick's four-level model is a widely accepted and highly popular evaluation tool. However, some criticise the model's shortcomings. This article will examine the extent to which the four-level model can evaluate design and technology students' learning about aesthetics after an intervention by reporting our use of an augmented version of the four-level model. We examine the results in terms of students' reaction to the intervention, their long-term learning and their behaviour changes by studying their visual analyses and drawings through segment codes. We found that, in order to uncover the obscurities imbedded in aesthetics and to explicate the complexities, we could not use the four-level-model on its own, but had to revert to a more augmented version.

Keywords

Kirkpatrick four-level model - Developmental research - Technology education - Aesthetics - Design principles - Design elements - Design techniques - Visual analysis - Pre-service teacher training - Instructional evaluation

Introduction

Researchers in the field of human resource development widely use Kirkpatrick's (1994) four-level model (hereafter called "the four-level model") to evaluate the effectiveness of education and training interventions by studying the success of learning outcomes (Arthur et al. 2003; Bates 2004; Winfrey 2002). The four-level model is widely accepted, highly popular, and regarded as a useful tool for instructors to benchmark the outcomes of learning interventions (Alliger et al. 1997). On the other hand, some strongly criticise the shortcomings of the model, e.g. its oversimplification of evaluation procedures, its incompleteness of structure, its assumption of causality, and its importance in terms of associated increasing of levels of learning outcomes (Bates 2004; Bates 2005; Russ-Eft and Preskill 2005). Although well documented, these advantages and limitations as an evaluation tool, offer no evidence for conclusions on its value as a theoretical framework for research in technology education (Winfrey 2002).

Therefore, this paper reports on the manner in which we implemented an augmented version the four-level model in order to contribute to the development of learning and teaching support in the field of aesthetics in technology. We will examine the extent to

which the four-level model allows for an alternative approach to conceptualise, operationalise and interpret qualitative data through analyses of aesthetic reasoning after an educational intervention.

Background of the study

The study involved a group of twenty-two first-year education students enrolled in Technology Education courses as part of pre-service training at the Faculty of Education, University of Pretoria (UP). For the purpose of this article, a case study was conducted on three diverse students from the student group as participants in this qualitative approach. The reason for selecting only three participants was to gain deep and rich information rather than obtaining wide and superficial information. The content of the module under discussion was design aspects of products, namely function, ergonomics, aesthetics and value taught in an integrated manner. The focus of this study is on three areas within aesthetics. Firstly on the constructs of aesthetics (design principles, elements and techniques), partly determining the quality of a design; secondly on the content knowledge required to critically analyse existing designs and thirdly on drawing skills to visualise the understanding of design in students' own designs (Anderson 1998; Tversky 1999). One of the interventions for learning and teaching constructs of aesthetics as well as the analysis of designs is an electronic tutorial consisting of fifty PowerPoint™ slides. Formal drawing tutorials and exercises support drawing and visualisation skills.

The design module under discussion, coded JOT120, is a semester contact course. It is presented during the second semester of the group's first year of study in design and technology. It follows on two modules: Conceptual Framework of Technology and The Design Process, including creative and critical thinking skills, perceptual development and visualisation techniques. When students start with JOT120, it is assumed that they have a conceptual understanding of the technological processes involved when seeking practical solutions to needs and problems experienced by humans and that they understand the importance of the aesthetic aspect of product design in the context of a market driven approach¹ (Parr 2004) towards product design.

We approached the implementation of the module from a constructivist perspective as a learner-centred and project-driven course. Two constructivist principles drove the implementation of the module in the constructivist paradigm: (1) Scaffolding through learning support material, structured lectures, class discussions, and class activities such as visual analysis and drawing skills. (2) Via problem solving through the implementation of sequential phases of the design process (identifying the problem, investigating the need, researching content knowledge available to support understanding of the problem and suggesting possible solutions, communicating possible solutions through exploratory and representational drawings).

In this study we aimed to understand what about aesthetic design theory has been learned/understood in each level of learning by novice designers with an electronic tutorial as intervention. The tutorial constitutes an introduction to universal visual

language in design theory as a cognitive tool for generation, visualisation, detailing and analysis of ideas (Anderson 1998).

Some of these concepts need explaining. *Design*, as prescribed in the Grade 9 National Curriculum Statement, relates to the steps of the design process (Department of Education 2003, p. 37). In design studies, the concept of “design aspects” refers to the function, aesthetics, ergonomics and value of man-made products, materials and manufacturing methods (Department of Education 2002; Garratt 1996; Press and Cooper 2002) considered in an integrated way during the design process. The word “aesthetic” is derived from a Greek word meaning perceptive, and it has to do with the appearance of things rather than with their reality (Faber 1975). The concept *aesthetics* in a technological context would then refer to appearance of strength rather than with actual strength. Dominating the landscape, a bridge, designed by a civil engineer, “may make or mar its surroundings for centuries to come. Consequently a striving after beauty of form and harmony with surroundings is a social obligation which structural engineers must recognise and educate themselves to perform” (Inglis 1975). Neglecting the aesthetic aspect of a structural design, leads to offending the eye of the viewer, which the designer has no right to inflict (Holden 1975). Although much has been written about aesthetics in architecture, little is found in the literature about reasoning in aesthetics in engineering/technology education (De Vries 2005).

Learning the theory and application of *aesthetics* in product design in a logic way (De Vries 2005) is a domain-specific construct divided in two: cognitive operations as visual vocabulary for the designer (Tversky 1999) and a universal visual design language (Anderson 1998). We embrace the definition of Lauer (1985) and Wong (1993) of a universal visual language comprising of design principles, elements, and techniques constructs of aesthetics. The design principles under discussion are unity, balance, emphasis, scale, proportion, illusion of space and rhythm; the design elements are line, shape or volume, colour and texture (Lauer 1985). Design techniques involve arrangements in space, as well as order and distance achieved through repetition, addition, omission and distortion (Tversky 1999). Aesthetics refers to the formal design theory—the principles that rule and govern the arrangement of elements (Lauer 1985; Wong 1993).

Aesthetic design theory enables designers to reveal ideas through sketching and visual analyses using a universal design language (Anderson 1998), based on gestalt psychology (Arnheim 1986). In addition, sketching and visual analysis of products form an important part of the initial stages of the design process. Scholars in the field of visual literacy regard drawings as cognitive tools to disclose not only thought, but also to reveal a conceptual demonstration of analytical skills (Anderson 1998; Garner 1993; Tversky 1999). In educational contexts sketches provide clues to conceptualisations of specific domains as revelations of participants’ conceptions (understanding) rather than their perceptions (sensing) of things, representations of reality rather than the presentations of reality and naturally segmentation of reality into elements that can be schematised and spatially arranged (Tversky 1999). It makes sense to study technology education students’ sketches and their visual analyses in order to learn what they know and

understand about technology and design generally (Goel 1995) and aesthetics specifically (Anderson 1998).

The four-level model for evaluating learning (Kirkpatrick 1994) has a long history in the evaluation of instructional interventions. Since its first creation in 1959, the four-level model has become the most widely used approach to training evaluation in the corporate, government, and academic worlds. Evaluation researchers generally agree that the four-level model contributes to researchers' conceptual thinking on evaluation, but disagree on the extent of the empirical testing of the model (Alliger et al. 1997; Bates 2004; Holton-III 1996; Newstrom 1995). While many acclaim its usefulness in its original as well as its augmented variations (Alliger et al. 1997), others criticise it for its limitations (Bates 2004; Bates 2005; Russ-Eft and Preskill 2005). However, practitioners and researchers alike acknowledge its simplicity, practicality, and focus on learning outcomes in professional environments (Holton-III 1996). Other researchers recognise the contributions of the model to intervention-evaluation-thinking and practice. The model

- contributes to focused evaluation practice on learning outcomes
- recognises that single-outcome measures do not adequately reflect the complexity of instructional programmes
- underscores the importance of examining multiple measures of instructional effectiveness
- differentiates between learning and behaviour
- emphasises the importance of learning-transfer processes in making learning interventions effective.

The four levels of learning in the model represent a sequence of ways to evaluate instruction and learning support material. The evaluation process on each consecutive level becomes increasingly more difficult and time-consuming, but it provides increasingly valuable information on the process of learning (Kirkpatrick 1994).

Level one evaluates participants' reaction, much like measuring customer satisfaction. This level examines how well learners like the instruction and instructional interventions. It measures general satisfaction with the learning material by establishing the support material's ease of use, comprehensibility and usefulness. Previously, cognitive researchers explored mental processes from the perspective of cognition rather than

affect. However, recent research indicates that every sensation gives rise to an affect or emotion (De Villiers 2002), and learners' initial reaction to instruction influences the quality and the quantity of learning. Although a positive reaction does not guarantee learning, a negative reaction certainly reduces the possibility (Kirkpatrick 1994).

Level two, learning, measures change on an intellectual level: increased knowledge, improved skills and changed attitudes. No behaviour change occurs without learning. Increased knowledge is measured by content learned, the principles absorbed and mastered, the improved performance techniques and a positive attitude towards the learning (Clementz 2002; Kirkpatrick 1994). Alliger et al. (1997) refined the Kirkpatrick model to both immediate knowledge and knowledge retention. They define immediate retention as the knowledge acquired at the conclusion of an intervention, while knowledge retention is knowledge retained after the intervention. This study used Alliger et al (1997) refined version of the four-level model, and will be referred to as the "augmented model".

Level three, changed behaviour by transferring knowledge gained from level two, includes the application of trained strategies within a different context from the result of learning (De Villiers 2002; Hannafin and Peck 1988; Osman and Hannafin 1992). While commenting on this level of evaluation, Clark (2004) maintains that "behaviour is the action that is performed, while the final results of the behaviour is the performance." Learners should be allowed the opportunity to demonstrate changed behaviour. Unless such an opportunity is offered, it is impossible to predict when a behaviour change will occur (Kirkpatrick 1994).

Level four, results, refers to achieved goals of training in terms of reduced costs, increased quality, improved production, and a decreased rate of employee turnover and absenteeism (Kirkpatrick 1994). This level measures the success of a programme in terminology used by managers and executives; therefore, level four results are not considered relevant to educational contexts (Winfrey 2002) and are not included in this study.

The augmented model (Alliger et al. 1997) provides a tool for the developers of course information. It feeds the cyclic learning process during the subsequent loops of design and development of learning and teaching support (Van den Akker 1999). Evaluation of learning outcomes remains a key element of instructional design to provide decision making direction, while planning aims at improving new interventions: "Benefits that can be derived from evaluation are directly related to the capacity of the evaluation to develop information that increases the clarity of judgement and reduces the uncertainty of action for the stakeholders" (Bates 2004, p. 343). For learning and instruction, the goal includes developmental work necessary to design learning environments, to formulate curricula, and to assess cognitive and learning achievements. Development research addresses a fundamental understanding of learning in order to develop better learning and teaching support strategies. The information revealed in the evaluation can inform such development and design (Van den Akker 1999).

Research methodology

We examined the procedural guidelines for using the four-level model (Kirkpatrick 1994) to plan our approach of operationalising and conceptualising our investigation and evaluation of the learning outcomes. We then based our operationalisation on the number, format, and chronological order of the assessment strategies devised for the course. We designed these opportunities to correspond with the evaluation of three of the four levels of the four-level model for the evaluation of learning.

The simplicity of the four-level model's structure offered several advantages for our study. First, it allowed for gathering, ordering, analysing and interpreting the data qualitatively. Second, it let us deal with the data interpretively. We selected the qualitative approach because it offered the most appropriate way to analyse the aesthetic qualities in the participants' drawings. In addition, it allowed for the interpretation of their analyses in an interpretive way (Patton 2002).

The four-level-model, however, also had serious limitations for our research, especially its lack of provision for validation, triangulation and trustworthiness (Holton-III 1996). In addition, we encountered challenges in the model's restrictive data interpretation. We questioned the trustworthiness and the validity of the data and how it interfered with the interpretation of data to produce grounded theory. We therefore decided to use the basic structure of the four-level model (Kirkpatrick 1994) to systemise and order the data gathering process and to implement the augmented model by Alliger et al (1997) for validation purposes.

Operationalisation

Operationalisation, or the after-the-fact descriptions refers to our observations, including a description of how we collected data, what we thought about the data and our explanation of the data. Collecting data on the three levels in our study entailed many data gathering techniques (Cohen et al. 2000), such as self-reporting measures, performance tests, and on-the-job performance (Kirkpatrick 1994). As we aimed to study the learning outcomes, and not to evaluate directly the effectiveness of the intervention, we used the existing data collecting strategies to achieve value neutrality, validity and reliability through crystallisation (Cohen et al. 2000; McMillan and Schumacher 2001). We soon realised that the simplistic structure of the model did not cover the nuances and sub-levels of learning and subsequently implemented the augmented the model by Alliger et al (1997) model to strengthen the operationalisation of our investigation. Table 1 provides a summary of the criteria and definitions used in our research.

We designed a combination of formative and summative assessments to assess students' progress, the data generating instruments summarised in Table 1. Summative assessments comprised a follow-up test as well as a formal examination at the end of the course. The combination of assessment strategies formed an integral part of our development research methodology (Van den Akker 1999), which will be described in the following section.

Evaluating aesthetics in technology education

Table 1. Alliger's augmented version of Kirkpatrick's four level model*

Level	Criterion *	Definition	Data generating instruments
Level 1 reaction:	Affective reactions	Participants' perceptions of the intervention-generally their satisfaction with the training	Questionnaires
	Utility judgements	Participants' belief about the value and usefulness of the training; the extent to which they believe they will use the knowledge and skills in other relevant situations	Interviews
Level 2 learning:	Immediate retention	The assessment of knowledge acquisition at the conclusion of the intervention	Exercise
	Learning after a period of time	The retention of knowledge at some point after the immediate conclusion of the intervention	Sequential test
Level 3 behaviour:	Transfer of knowledge and skills	Demonstrated performance after a period of time	Project
	Sustained performance	Demonstrated sustained performance after a period of time	End-of-year examination

* Kirkpatrick's Four Level Model does not include a list of criterion, but only a list of suggestion of how to proceed from one level to another (Kirkpatrick and Kirkpatrick 2006)

On level one (reaction) of the Kirkpatrick model, we designed two data collection strategies, a paper-and-pencil questionnaire and semi-structured interviews to validate the participants' reaction to the intervention as well as to capture their affective and attitudinal responses to the intervention. We collected learners' beliefs and perceptions of the intervention—their satisfaction with the training. We also noted their beliefs about the usefulness of the intervention and the potential value of the instruction to practically apply their knowledge and skill in related tasks (Alliger et al. 1997).

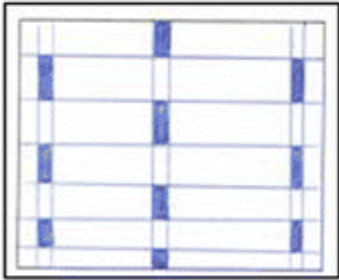
On level two (learning) we used an exercise and a sequential test as evaluation strategies. We focused on what the participants learned about design principles, elements and techniques during the intervention. We required instruments that could simultaneously measure visual analysis of drawings and design knowledge that would also validate the

data. We selected two types of data analysis: written visual analyses of existing designs and analyses of drawings. First we designed a pencil-and-paper visual-analysis exercise and presented it to the participants directly after the introduction of the intervention to measure immediate learning, in which they were required to explain the operational relations between the different constructs, i.e. design principles, design elements and design techniques. Second we designed a sequential paper-and-pencil test to measure retention of knowledge and skills and to validate the data. Participants were required to analyse existing designs and then to demonstrate their understanding of some constructs through free hand drawings.

We provided open-ended and non-directed test items. Participants saw three images of designs to analyse and discuss without any directions to specific design constructs. This complicated our investigation greatly as we needed to consider combinations of elements and techniques and their functional relationships. We could not foresee all combinations, but we uncovered them during reading and re-reading data. For example, when a participant identified a design principle without recognising the applied design elements and techniques, we considered this a poor demonstration of analytical skills. The opposite was also true. We established triangulation when participants' responses contained the appropriate terminology, accurate and detailed specific descriptions, and the participant explained the visual effect of the constructs in a design.

We designed a separate coding system to analyse and interpret participants' drawings as a standard of measuring the quality of participants' understanding of the constructs in the analysis and interpretation of participants' drawings.

The cognitive indicators for measuring participants' knowledge in their drawings were the clear evidence of design principles, the number of the design techniques applied, and the number of design elements used. Figures 1–3 (participants' drawings from the test) revealed the learning that had taken place. Participants applied design elements by employing techniques resembling the designs in the tutorial.



Participant 1's schema of the design principle "unity"

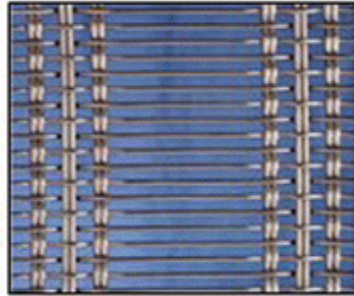
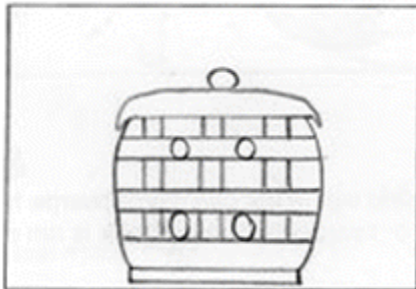


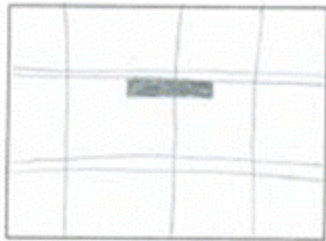
Image in the intervention depicting the design principle "unity" (slide 12)



Participant 2's schema of the design principle "unity"



Image in the intervention depicting the design principle "unity" (slide 13)



Participant 3's schema of the design principle "unity with variation"

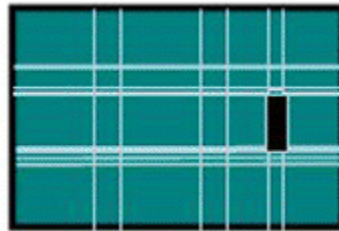


Image in intervention illustrating the design principle "unity with variation" (slide 14)

Fig. 1. Participant 14's exploratory drawing of a lamp design in the second assessment, indicating her prevailing lack of drawing skill

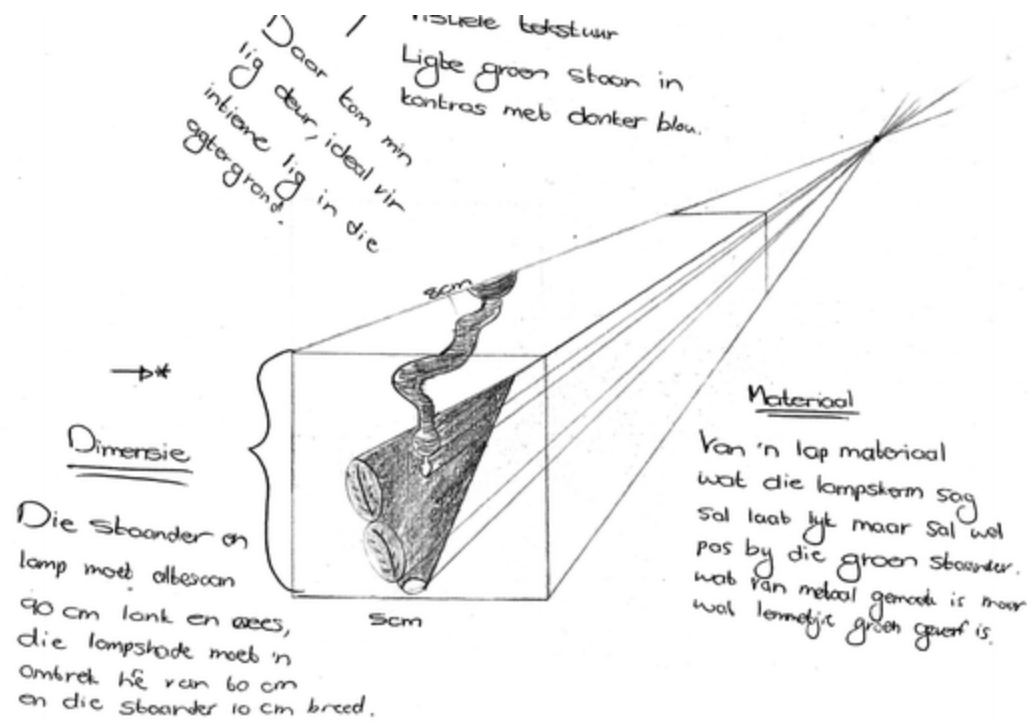


Fig. 2. Participant 9's exploratory drawing of a lamp design in the second assessment indicating her improved drawing skill as well as an increase in visual interest

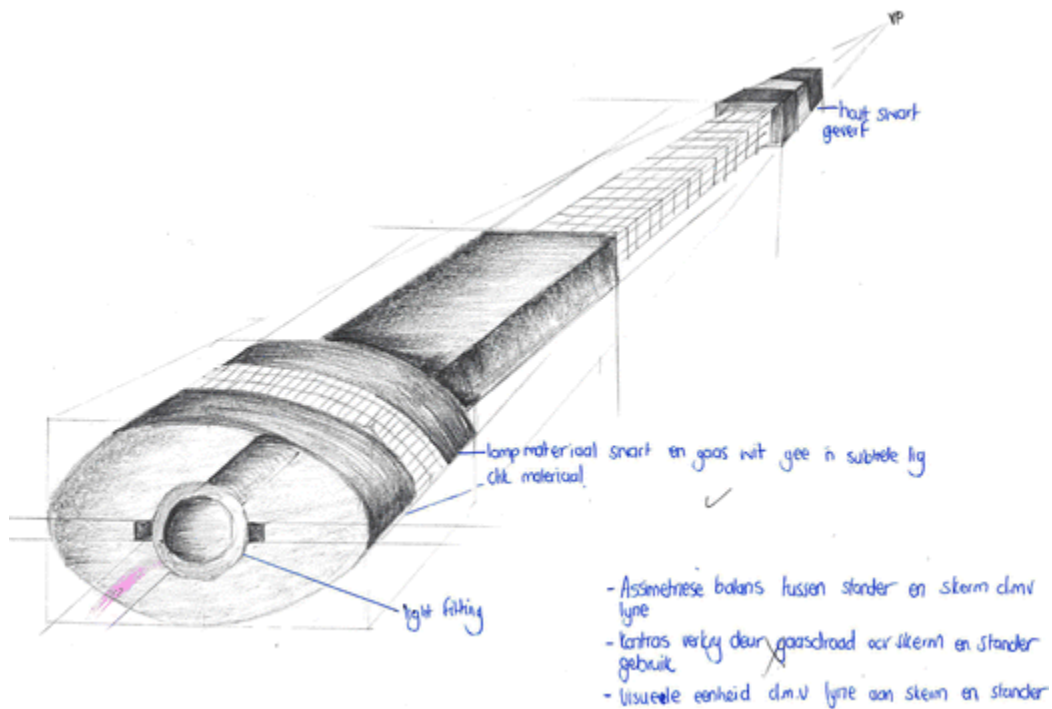


Fig. 3. Participant 21's exploratory drawing of a lamp design in the second assessment indicating her sustained ability to draw, make her intentions clear and apply the design elements and principles successfully

On level three (behaviour) we selected the participants' individual projects as primary data sources for data analysis. We evaluated participants' learning after a period of time by studying presentational drawings required for the project. In order to evaluate their sustained performance and as validation strategy we examined their development drawings from their end-of-year examination. We used the same coding system as indicators of behaviour for level three, as for the drawings made by the participants during level two. Therefore we experienced the same complications as we had under level two. Consequently, we accumulated rich detail that guided us towards creating patterns of knowledge, understanding aesthetics, and visual literacy from the wealth of the qualitative data.

Conceptualisation

Conceptualisation refers to our process of making sense of data, organising data, formulating preliminary ideas, developing concepts, formulating definitions, considering relations among concepts, and linking concepts and constructs. As previously mentioned, one of us was the only the subject matter expert also involved in the course delivery of the intervention, contributing inter-validity. We achieved crystallisation through applying multiple research instruments and constantly cross-checking findings (Cohen et al. 2000).

We achieved internal and external validation of the data interpretation through a process of crystallisation by developing clear and explicit definitions, linking them to relating ideas, and tying them closely to specific data. We scrutinised the data to understand the responses and achieved construct validity by categorising data according to classical domain-specifics of aesthetic design theory. To achieve criterion-related validity, we used the augmented model (Alliger et al. 1997) to categorise the levels of learning.

On level one, we read and re-read the data gathered on level one for the domain-specific constructs. Subsequently we analysed the content of the responses to the questionnaire and semi-structured interviews. We used a basic content analysis approach to analyse interviews and questionnaire. We searched for words, phrases, and statements in the three participants' reactions as they verbalised their experiences, feelings, knowledge, and formation of the constructs. In order to establish reliability and validity in the analysis, we used a criterion-related coding system to standardise the analysis of the responses (Cohen et al. 2000). We created codes for analysing and categorising the comments into two types of reactions— affective reactions and utility judgements, (Table 1) and then categorised the affective reactions into two sub-groups, based on the coding system, of general satisfaction and dissatisfaction. Similarly, we also categorised the utility judgements into three sub-categories: the intervention's perceived relevance, the perceived practicality and worth-the-while interaction (Alliger et al. 1997).

We developed definitions of affective reactions and utility judgements. These definitions guided us to organise the data into categories according to themes, concepts and commonality from the definitions. We searched for phrases such as "I like ...", "I did not

like ...” “I would like ... to be different” and then scrutinised the participants’ responses. Subsequently, we classified the responses into three groups: responses indicating general satisfaction or dissatisfaction, and their perceived usefulness.

We created two categories of comments for the participants’ utility judgements captured from the questionnaire and interviews. We divided comments on the relevance of the intervention and its potential for learning from comments relating to participants’ perceptions of their learning with the intervention. This analysis provided rich descriptions, and allowed us to describe how the participants reacted to the intervention. We gained insight into participants’ satisfaction, dissatisfaction, and their perceived usefulness of the intervention. Although measuring reaction is not an indication of learning (Kirkpatrick 1994), a positive reaction to the intervention indicated the participants were motivated to interact with it. Through their utility judgements the participants indicated perceived usefulness of the intervention.

The affective responses indicated the participants’ general satisfaction with the intervention—how much they liked or enjoyed it; how easy the content was, and how comfortable they were using the intervention. The affective responses also indicated the extent to which participants understood how to interact with the intervention, considered the intervention relevant and practical, and thought it a worthwhile learning tool. On level one, our analysis revealed that the participants were motivated to interact with the intervention because they enjoyed it, perceived it as practical and useful learning tool - despite some negative reactions.

On level two, domain-specific constructs formed the basis of the coding system. We captured data from the exercise and the sequential test (Table 1) and analysed and categorised them. Evaluating the research data for level two involved analysing two types of data: visual analyses of existing designs, and drawings representing participants’ understanding of formal design theory. The exercise generated data revealing participants’ immediate retention by visual analyses. The sequential test provided visual analyses as well as drawings.

In order to ensure validity and reliability in analysing and interpreting data generated by participants’ visual analyses, we created a coding system. We used the same coding system for the exercise as well as for the sequential test. Words, phrases and definitions indicating a conceptual understanding of the design principles and design elements, including the use of the correct terminology used in the tutorial, were considered a demonstration of knowledge retention. Students were given images of designs to be analysed and discussed without directing them towards specific constructs. This presented a difficulty in the research as numerous possible combinations and relationships (Tversky 1999) could not be foreseen and were established in the reading and re-reading process. If a student could identify only the principles without recognising the elements and techniques applied, it revealed poor understanding of the operational constructs and poor application of analytic skills.

We used Tversky's (1999) theory to assess how the participants' drawings revealed their cognition. According to Tversky, drawings are *inter alia*, clues to mental conceptualisations of the domain—participants' conception of things, not their perception of things. Drawings represent reality; they do not present reality. They segmented into elements that can be schematised and spatially arranged endlessly. Some segments or elements can be used in combinations to produce a potentially infinite set of drawings. The analysis of the designs revealed that the participants did not differentiate well between design principles and elements. The findings indicate significant links between learning on the immediate retention level and after a period of time about the participants' understanding of the domain constructs. They

- understood the design principles fairly well, except for the concept balance
- understood the design elements fairly well, except for texture
- did not understand design techniques to manipulate elements
- only after a period of time started to understand arrangements in designs. This indicated an improvement of learning from one level to another. It also confirmed the gestalt theory that perceptual patterns cannot only be recognised by tracing the relations between elements, but that an approach “from above” (describing the overall structure of the pattern imbedded in the design principles achieved) is needed to understand the place and function of each element and technique to one another (Arnheim 1986).

On level two, the findings of the investigation provided evidence of similarities between the explanatory text in the tutorial, the responses of participants in the exercise, in the test relating to visual analysis. This indicated a direct link between the content of the tutorial and what the participants learned. Schemas in the test indicated similarities between existing designs in the tutorial and in the participants' schemas.

Measuring learning outcomes and validating the knowledge and skills became complicated on level two. The complication resulted from the complex criteria for visual analysis, the intricate analysis of drawings, and the complexity of domain constructs (Tversky 1999). To ensure validity, we designed a coding system for the analysis and interpretation of the participants' responses to the visual analysis exercise. We searched for cognitive indicators such as words, phrases, and definitions that indicated conceptual understanding of the design principles and elements.

On level three, the data revealed a sustained change in behaviour in the participants' on-the-job performance. We posed the following questions to explore the extent of the

participants' applied design knowledge and skills: (1) How well did the drawings reveal the design elements and techniques to realise design principles? (2) Was it sufficiently clear? (3) How well did the drawing represent the idea? (4) Was sufficient visual interest created through variety? First, all the participants uncovered the design principles and elements with sufficient clarity. This indicated sufficient transfer of knowledge of design principles and elements to change the participants' behaviour. Also, this change in behaviour was sustained after a time lapse and demonstrated again during the examination. Second, some of the participants improved in transferring their knowledge of design techniques from the first to the second assessment on level three, while other participants, who performed well in the first assessment, did not improve much. Third, some participants represented their ideas well but did not accomplish the required drawing skills. Fourth, most participants did not create good visual interest in the first assessment but improved in the drawings of the second assessment.

Figures 1–3 reveal the improvement in drawing skills and creation of visual interest in the second assessment.

On level three, similarities between participants' final representational drawings in their projects and their developmental drawings in the examination were less apparent. However, the evidence of the participants' acquired learning in the projects and examination seems to align well with Arnheim's (1954) longstanding theory that new images contact memory traces from past experiences; therefore, we infer that the intervention had an ongoing effect on participants' learning on level two and on level three.

Conclusions

From our interpretation and conceptualisation of the data, we developed the following theory on the design and development of electronic interventions for the domain of aesthetics:

- Visual and verbal reference to the constructs and segments of the design constructs should be explicit and repetitious
- Participants' verbal referencing should be reinforced with interactive activities
- Behaviour change (level two–learning) is not permanent unless accompanied by sustained change in level three is apparent.

Our investigation illustrated that employing the four-level model for the evaluation of educational interventions carries both advantages and limitations for research in the field of design and technology education. Although learning and behavioural criteria are conceptually related, we found little evidence of this link in the literature (Arthur et al.

2003). We also could not show conclusive empirical evidence of the relationship between level two and level three. These weaknesses inherent in the four-level model were substantiated by critics of the model (Bates 2004), reminding us that the substantive and procedural qualities of the model have significant limitations for accurate evaluation:

- The model does not endorse key contextual input variables in the evaluation of training, thereby weakening the evaluation process and masking training complexities.
- The model focuses only on outcomes. It generates little or no formative data on the efficacy of the training.
- The model assumes that all learning contexts have equal effects on the transfer of knowledge.

Although the model provided a pragmatic tool for the evaluation of learning in an especially systematic way, it also oversimplified the evaluation process to such an extent that we could not draw clear conclusions. Although the Kirkpatrick model (1994) assumes a link between the different levels of learning, we were not able to show evidence of such a causal chain. In addition, we saw limited evidence of substantial relationships between findings on the different levels.

The advantage of using the augmented model (Alliger et al. 1997) provided an approach for operationalising and conceptualising our investigation. The four-level model (Kirkpatrick 1994) guided us to generate and capture data that clearly illustrated the participants' satisfaction and dissatisfaction with the intervention. In addition, it provided a basis for understanding the evaluation of educational interventions based on the achieved learning outcomes in a systematic way. Furthermore, it simplified the complex process of the evaluation of an intervention by providing straightforward guidelines on the type of questions that should be asked as well as on appropriate criteria for the intervention. This approach reduced the number of measurements for the evaluation of the intervention, e.g. pre-tests could be eliminated. We could collect data after the intervention was implemented. The model's focus on outcomes also reduced the complex network of variables that evaluators of educational interventions must contend. We also learned that advancing our case study from a qualitative approach by inter alia using Alliger's (1997) augmented version, and reporting the findings in an interpretative way, we overcame many limitations inherent of the Kirkpatrick model. The augmented four-level model developed by Alliger enabled us to generate "thick and rich" findings leading to deep understanding and interpretation for our research perspectives (Cohen et al. 2000; McMillan and Schumacher 2001). Emergent and grounded theory developed from our research findings resulted from the flexible and open-ended approach allowed by the model (Cohen et al. 2000).

Recommendations

The four-level model should be implemented for qualitative research and evaluation investigations even if Kirkpatrick recommends quantitative measures. In addition, each level of the model should be elaborated and expanded in order to accommodate validating measuring strategies (Kirkpatrick 1994) according to the specific study.

To expand on our findings of this study, we recommend follow-up research on Kirkpatrick's level four—results in terms of a longitudinal investigation where professionally accepted behaviour will be measured when the participants (pre-service teachers) are employed in real-world teaching environments. Such an investigation would conclude the full cycle evaluation of the participants' learning, not only the partial implementing of the intervention (Holton-III 1996). Further investigations should also be initiated on the four-level model's failure to offer causal links between the interventions and the participants' learning (Bates 2004).

References

- Alliger, G. M., Tannenbaum, S. I., Bennett, W., Traver, H., & Shotland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology*, 50(1), 341–358.
- Anderson, E. (1998). *Enhancing Visual Literacy through Cognitive Activities*. (Paper presented at the 2002 ASEE/SEF/TUB Colloquium, Glasgow).
- Arnheim, R. (1954). *Art and visual perception: a psychology of the creative eye*. Berkeley, CA: University of California Press.
- Arnheim, R. (1986). *New essays on the psychology of art*. Berkeley: University of California Press.
- Arthur, W., Bennet, W., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology*, 88(2), 234–245.
- Bates, R. (2004). A critical analysis of evaluation practice: The Kirkpatrick Model and the principle of beneficence. *Evaluation and Program Planning*, 27, 341–347.
- Bates, R. (2005). Kirkpatrick Four-level Evaluation Model. In S. Mattheson (Ed.), *Encyclopedia of Evaluation* (pp. 221–222). London: SAGE.

Clark, D. (2004). *Evaluation phase—instructional systems*. Retrieved 14 September, 2004, from www.nwlink.com/~donclark/hrd/sat6.html

Clementz, A. R. (2002). *Program level evaluation: Using kirkpatrick's four levels of evaluation to conduct systemic evaluation of undergraduate college programs*. 24 September 2003, from http://web.bryant.edu/~assess/Program_Level_Evaluation.doc

Cohen, L., Manion, K., & Morrison, K. (2000). *Research Methods in Education* (5 ed.). London: Routledge & Farmer.

De Villiers, M. R. (2002). *The dynamics of theory and practice in instructional systems design*. Thesis, University of Pretoria.

De Vries, M. J. (2005). *Teaching about technology. An introduction to the philosophy of technology for non-philosophers* (Vol. 27). Dordrecht: Springer.

Department of Education. (2002). *National curriculum statement for grades R–9 for technology*. Pretoria: Department of Education.

Department of Education. (2003). *White paper on e-Education*. Pretoria: Department of Education.

Faber O. (1975). Aesthetic aspect of civil engineering design. In Institution of Civil Engineers (Ed.), *Aesthetic Aspect of Civil Engineering Design* (pp. 1–15). London: Institution of Civil Engineers.

Garner, S. W. (1993). The importance of graphic modelling in design activity in teaching and learning technology. In R. McCormick, P. Murphy & M. Harrison (Eds.). Workingham: Addison-Wesley Publishing Company.

Garratt, J. (1996). *Design and technology* (2 ed.). Cambridge: Cambridge University Press.

Goel, V. (1995). *How sketches work. A bradford book*. Cambridge: MIT Press.

Hannafin, M. J., & Peck, K. L. (1988). *The design, development and evaluation of instructional software*. New York: Macmillan.

Holden C. (1975). Aesthetic aspect of civil engineering design. In Institution of Civil Engineers (Ed.), *Lecture at Seminar for Institution of Civil Engineers* (pp. 41–50). London: Institution of Civil Engineers.

- Holton-III, E. F. (1996). The flawed four-level evaluation model. *Human Resource Development Quarterly*, 7(1), 5–21.
- Inglis, R. (1975). *Aesthetic aspect of civil engineering design*. London: Institution of Civil Engineers.
- Kirkpatrick, D. L. (1994). *Evaluating training programs: The four levels*. San Francisco: Berrett Koehler.
- Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels* (3rd ed.). San Francisco: Berrett-Koehler Publishing.
- Lauer, D. A. (1985). *Design Basics* (2nd ed.). New York: Holt, Rinehart and Winston.
- McMillan, J. H., & Schumacher, S. (2001). *Research in education: A conceptual introduction* (5th ed.). London: Longman Inc.
- Newstrom, J. W. (1995). Review of evaluation training programs: The four levels by D.L. Kirkpatrick. *Human Resource Development Quarterly*, 6(1), 317–319.
- Osman, M. E., & Hannafin, M. J. (1992). Metacognition research and theory: Analysis and implications for instruction design. *Educational Technology Research & Development*, 40(2), 83–99.
- Parr, J. W. (2004). *Aesthetic intentions in product design market driven or alternative form*. Retrieved 4 September, 2004, from <http://folk.uio.no/parr/text/aesthetic.pdf>.
- Patton, M. (2002). Two decades of developments in qualitative inquiry. *Qualitative Social Work*, 1(3), 261–283.
- Press, M., & Cooper, R. (2002). *The design experience*. Burlington: Ashgate.
- Russ-Eft, D. F., & Preskill, H. S. (2005). Search of the Holy Grail: Return on investment evaluation in human resources development [Electronic Version]. *Advances in Developing Human Resources*, 7, 1:71–85. Retrieved 5 February 2007 from <http://adh.sagepub.com/cgi/content/abstract/7/1/71>.
- Tversky, B. (1999). *What does drawing reveal about thinking?* Retrieved 20 January, 2005, from <http://www.arch.usyd.edu.au/kcdc/books/VR99?Tversky.html>.
- Van den Akker, J. (1999). Principles and methods of development research. In J. Van den

Akker, N. Nieveen, R. M. Branch, K. L. Gustafson & T. Plomp (Eds.), *Design methodology and developmental research in education and training* (pp. 1–14). The Netherlands: Kluwer Academic Publishers.

Winfrey, E. C. (2002). *Kirkpatrick's four levels of evaluation*. Retrieved 24 September, 2003, from <http://coe.sdsu.edu/eet/Articles/k4levels/index.htm>.

Wong, W. (1993). *Principles of form and design*. New York: Van Nostrand Reinhold.

Footnotes

¹ Market driven approach: aesthetic choices are quite often based on and motivated by market surveys, studies of competing products, and studies of defined target user groups, in order to make the product fit a favourable niche among other products (Parr 2004).