

## 2 Theoretical Background

### 2.1 Statistical misconceptions

Statistical misconceptions are systematic patterns of errors that occur during interpreting, understanding or applying statistical concepts (Castro Sotos, et al., 2007; Liu, 2010). Sources can be ambiguous use of language (Sedlmeier, 1998), inherited intuitions (Sedlmeier, 1998; Watson & Moritz, 2000), and current structure in statistical teaching material (Gliner, Leech, & Morgan, 2002; Haller & Krauss, 2002). There are three characteristics of statistical misconceptions: first, they are difficult to observe (Garfield & Ben-Zvi, 2007), delicate to correct (Garfield & Ben-Zvi, 2007), and they complicate further learning processes (Castro Sotos et al., 2007). For instance, concepts such as the *p-value* can be seemingly plausible but are built on counterintuitive facts and are therefore misunderstood very often and deeply (Kirk, 2001). Bodemer et al. (2004) and Liu (2010) mention that misconceptions occur and hinder learning as fragmented mental representations make the understanding of abstract concepts difficult. This section introduces misconceptions that occurred among students in empirical studies. Statistical misconceptions were observed among understanding of *data distribution and variation* (Batanero, Tauber, & Sánchez, 2005; Chance et al. 2004; Finch, 1998; Jazayeri, Fidler, & Cumming, 2010; Saldanha & Thompson, 2007; Sedlmeier, 1998; Well, Pollatsek, & Boyce, 1990), *statistical significance* in testing (Haller & Krauss, 2002; Kalinowski, Fidler, & Cumming, 2008), and *confidence interval* (Belia, Fidler, Williams, & Cumming, 2005; Cumming, Williams, & Fidler, 2004).

**Data distribution and variance.** Understanding data and statistics starts with thinking about the distribution of sample data. As there are already several types of distribution to understand data, statistical measures and different characteristics related to these types, students confuse these types, especially if the learning topic is about sampling distribution. There are several interesting studies that have identified and looked more closely into the issue where students misunderstood or mixed up types of distributions (Batanero, Tauber, & Sánchez, 2005; Chance, delMas & Garfield, 2004; Finch, 1998; Jazayeri, Fidler & Cumming, 2010; Saldanha & Thompson, 2007; Sedlmeier, 1998; Well, Pollatsek, & Boyce, 1990). The most interesting findings and suggestions to improve con-

cepts around distribution and variation were reported in delMas et al.'s (1999) study. They state that understanding representations of frequency and its connection to probability (e.g. in a density distribution) is not intuitive at all because we do not regularly think in terms of huge amounts of data but rather in small samples. They also identified that there are difficulties in several different dimensions of thinking about a problem: application of rules, terminology, confidence (i.e. 'degree of certainty in choices or statements' (Chance et al., 2004, p. 309), and connecting ideas (i.e. integration of concepts). delMas et al. (1999) applied a model called *predict-test-evaluate* to confront students with their misconceptions in a simulation task. They compared this simulation activity to a similar simulation activity without the predict-test-evaluate structure. In the predict-test-evaluate condition, students with misconceptions tested their own hypotheses and created a sampling distribution in a computerized simulation task to test their hypotheses and to confront their understanding. As they received the correct solution, they were asked to reflect on the outcome of their experiment by comparing it to the correct solution. This learning activity of comparison resulted in a large improvement in reasoning and giving correct answers for students in the predict-test-evaluate condition: from 16% of correct reasoning in the pretest up to 72% in the post-test and from 16% in the non predict-test-evaluate structured task to 36% in the predict-test-evaluated structured task.

**Statistical significance.** Another concept that is difficult to understand and causes a lot of misconceptions is the interpretation and understanding of *statistical significance* (Castro Sotos et al., 2007; Haller & Krauss, 2002; Kalinowski, Fidler, & Cumming, 2008). These studies and the review of Castro Sotos et al. (2007) state that understanding statistical significance is difficult because it is an abstract concept similar to the sampling distribution. In addition, there is another obstacle according to Gigerenzer, Krauss, and Vitouch (2004): the p-value seems to provide the information that a researcher or student really *wants to know*, which is that the p-value indicates the probability for the null hypothesis. They outline that the correct way of thinking of this concept is a counterintuitive way of thinking that has to be remembered every time. Furthermore, Fisher's approach and Neyman and Pearson's approach to the logic of statistical significance are often communicated as one common theory and not clearly separated in educational material (Gigerenzer et al., 2004; Haller and Krauss, 2002). To the authors' knowledge, there are only two attempts to correct this misconception in an empirical study. Kalinowski, Fidler, and Cumming (2008) tried to resolve the p-value misconceptions by using distinctly false applied logical sentences in discussions with students. By doing this, they confronted students with these wrong conceptions and corrected their misconceptions significantly. Another attempt that managed to resolve p-value misconceptions of students by applying

a similar strategy to Kalinowski et al. (2008) was the dissertation of Baglin (2013). In both studies, the improvements were on the linguistic level.

**Confidence interval.** Finally, a concept easily misunderstood is the *confidence interval* (Belia, Fidler, Williams, & Cumming, 2005; Cumming, Williams, & Fidler, 2004; Fidler, 2006). The confidence interval is also related to the understanding of probability and significance. According to the review written by Castro Sotos and colleagues (2007), there are more studies reporting misconceptions among researchers than among students. Fidler (2006) found that psychology and ecology students with prior knowledge in statistics misunderstood the confidence interval as a range of individual scores or that this interval increases with sample size. These misconceptions were mentioned among others such as the confidence interval contains ‘plausible values for sample mean’ (Fidler, 2006, p. 4) or the ‘90% CI [is] wider than the 95% CI (for same data)’ (Fidler, 2006, p. 5). However, Fidler (2006) also mentioned that by understanding how to interpret the confidence interval, the interpretation of the statistical significance improved. Similarly, Lipson (2002) discovered that the more students embedded and linked the sampling distribution in their statistical concept map, the better was their understanding of statistical inference including *p*-value and confidence interval. Fidler’s (2006) and Lipson’s (2002) studies therefore indicate that understanding of some concepts depends on the understanding of other concepts. Thus, connecting concepts could help with remembering and understanding concepts.

Approaches to simulate statistical processes have helped students in applying rules and relating concepts. This resulted in the discovery and correction of some statistical misconceptions (Chance et al., 2004; Jazayeri, Fidler, & Cumming, 2010). Both studies focused on the explanation of the sampling distribution on a visual level, but neither on the *p*-value, nor on the *confidence interval*. Therefore, it would be of interest to simulate processes related to these two statistical concepts. To create an effective statistical learning program with graphical simulations, empirically tested cognitive principles have to be applied.

## 2.2 Interactive visualized statistical concepts

This section focuses on cognitive principles that have been applied to create online learning tools. According to Rey’s (2009) review of theories in e-learning and Moreno and Mayer’s (2005) study, there are several learning principles that have to be considered when creating an interactive learning tool such as simulations.

**Structure and guidance.** In Mayer's cognitive theory of multimedia learning (CTML) (first overview: Mayer & Moreno, 1998), several cognitive processes are described (Mayer, 2005; Moreno & Mayer, 2005). This model includes theoretical attempts such as the cognitive load (Sweller, Van Merriënboer, & Paas, 1998) and Baddeley's working memory model (Baddeley, 1992). According to Mayer (2005), there are three important assumptions that can be derived from the CTML for the creation of learning material: First, representation of information should be on a verbal as well as on a pictorial level so that information can be processed more deeply in the working memory. Second, a learning person can only process a limited amount of information (Baddeley, 1992; Sweller et al., 1998), hence presented information should be short and clear in a learning environment (Rey, 2009; Sweller, 1994). Third, a learning person has to process information actively in order to acquire a concept in a coherent and meaningful way. Because of his third assumption, Mayer (2005), proposes that information is structured; for instance, in a hierarchical manner where concepts are represented in categories or subcategories. De Jong and van Joolingen (1998) emphasize in their review that *structured simulations* as learning environments were especially effective in the sense that students learned concepts long-term.

**Conflict.** Limón (2001) postulated another theory that is important in relation to learning environment: the *cognitive conflict theory*. According to Limón (2001), a conceptual change can take place if learners are confronted with correcting information that helps them to reduce confusion. His assumption is that learners are conscious of their understanding of a concept or a relation between two concepts. Next, some new information is presented to them, for example some data that disprove the previous understanding. This causes an uncomfortable feeling and learners will try to reduce this feeling either by adapting the prior understanding of the concept to the encountered information or by stopping the learning process. Therefore, a cognitive conflict can be produced if learners are confronted with their wrong answer or misconception and the correct solution (Jazayeri et al., 2010). To the authors' knowledge, the cognitive conflict theory was applied and could successfully improve students' statistical knowledge in four studies (Jazayeri et al., 2010; Kalinowski et al., 2008; Liu et al., 2010; Baglin, 2013).

**Explanatory feedback.** The feedback principle in the CTML postulates that the learner should receive not only a correct answer but also an explanation in order to benefit from the learning environment (Mayer, 2005). Similar to general information in the learning material, good explanatory feedback – such as an example solution – should be phrased as clearly and briefly as possible and should be well structured (Mayer, 2005; Rey, 2009; Sweller, 1994). However, as clearly explained and well structured a sample solution is, the task can still be

too demanding. Renkl (2005) reported in his study that students could be overwhelmed if a task demands means-end analysis. He explains the process as follows: Means-end analysis is when a learner has to process several steps to reach a goal. Subgoals have to be created and writing the answer increases cognitive load and can reduce cognitive capacity, which results in decreased understanding of the learning material.

**Reflection.** In a study of Moreno and Mayer (2005), participants – undergraduate students in psychology – selected appropriate characteristics of plants to adapt them to different environments. Moreno and Mayer’s (2005) results of their third experiment indicate that an interactive learning task is in general as good as a non-interactive task in improving knowledge of college students, as long as the task was guided and students could reflect correct system-generated answers in comparison to their answers. However, students were worse in answering knowledge questions when the task was interactive compared to when it was not interactive if they could *not* reflect on correct answers provided by the system. According to several studies cited by Moreno and Mayer (2005) (Chi, de Leeuw, Chiu, & La Vancher, 1994; Martin & Pressley, 1991; Willoughby, Waller, Wood, & MacKinnon, 1993; Woloshyn, Paivio, & Pressley, 1994; Wood, Pressley, & Winne, 1990), asking students to reflect about *correct learning content* in texts helped students to understand the content better. The argument goes that reflection initiates deeper cognitive processes such as inference (Seifert, 1993). Therefore, Moreno and Mayer (2005) assume that students integrate and organize old and new information if they are able to reflect about learning content that is correct.

**Visualization.** Mayer (2011) postulated in his *multimedia instruction hypothesis* that concepts are learned better when using both sensory channels: verbal and pictorial channels instead of just one channel. Corter and Zahner (2007) discovered in their structured interview study that students spontaneously created visual representations in order to understand probability problems. Moreover, to improve statistical misconceptions, attempts with simulation-based tasks worked when students saw how a sampling distribution is built (delMas et al., 1999; Lipson, 2002).

**Interactivity.** The term interactivity takes a central role in this area of research. Visualizations are interactive when a computer-generated ‘series of drawings [...] can be paced by the learner or [...] animation[s] [...] can be stopped and started by the learner’ (Mayer, 2011, p. 428), whereas visualizations are non-interactive when the learner only observes them. A study of Schwartz and Martin (2004) found that students could improve understanding of statistical concepts by learning interactively with graphical tools. That was especially the case when their learning context was framed by an experiment where they had to predict

outcomes and received more learning resources (a follow-up lecture). In Moreno and Mayer's (2005) study, where students learned about the growth factors of plants in a tutorial, students in the interactive condition selected answers, hence they decided on their own what the best answer might be. In studies about intervention approaches to resolve statistical misconceptions with simulation tasks (delMas et al., 1999), students had to observe simulation processes while changing parameters such as the sample size  $N$  were selected for them. In both studies, students had to select answers on their own. Another study that directly compared interactivity to non-interactivity in multimedia learning environment was conducted by Evans and Gibbons (2007) who found that students learning with interactive images and texts outperformed students learning with a non-interactive images and texts. Therefore, interactivity with the learning content seems to have an effect on learning outcome. However, we were specifically interested in what would happen if students selected the parameters on their own. Therefore, in this study 'interactivity' is defined as the process where students take an active part in learning and decide on their own how to interact with a graphical simulation. The cognitive process behind the interactivity is comparable to reflection (Moreno & Mayer, 2005). According to their explanation, in both cognitive processes students have to integrate and organize old and new information in order to make sense of a concept that they want to acquire.

### 2.3 Improvement in knowledge

The aim of this study was to find out what role interactivity plays in learning with graphical simulations. To reach this goal, the described learning principles and empirical insights were combined to create an interactive e-learning environment in the form of a tutorial in which students could interactively change visualizations and answer questions.

Overall, Moreno and Mayer (2005) discovered that students performed well in knowledge transfer tests if they interacted with a structured program and could reflect on correct answers. Hence the question for this study is whether reflection is enough or whether *interactivity* is needed for deeper cognitive processing and improvement in knowledge and understanding statistics. Therefore we kept the e-learning principles structure and guidance, conflict, explanatory feedback and reflection the same for both test groups. Most importantly, we adapted the structure *predict-test-evaluate* (Chance et al., 2004; delMas et al., 1999) and created tasks where students had to hypothesize how statistical visualizations change when certain parameters are changed. Then the students conducted tests where they set parameters to change the visualization. By creating these kinds of tasks,

the principles structure and guidance, and conflict were applied. To investigate the influence of interactivity, we manipulated the way students interacted with visualizations – the statistical graphs. The group that could interact with the graphs was called the *dynamic test group* and the group that could not interact with the graphs was called the *static test group*. This manipulation might reveal whether interactivity is necessary or whether other e-learning principles are enough for a significant increase in statistical knowledge and understanding.

To detect changes in knowledge and understanding, we measured students' statistical knowledge and understanding three times, once before and twice after learning with the tutorial. The increase in knowledge and understanding was measured by questions that demanded knowing not only the definition of a concept but also its application to statistical graphs. Furthermore, the subjective perceived increase in knowledge and understanding was measured.

According to prior stated theoretical insights (delMas et al., 1999; Moreno & Mayer, 2005), the following outcomes in test performance for the experimental groups are expected. It is assumed that the dynamic test group will be supported in their learning process by the live interaction that they are allowed to perform: they can change diagrams by changing parameters related to the concept in order to understand how the concept works. As a result, they should have more cognitive capacity in order to understand how a statistical graph can change. Therefore, participants in the dynamic condition will be better in processing explained concepts in these tasks and will have a higher sum of test score. However, the knowledge and understanding performance should not differ in the pretest because measured prior knowledge is expected to be the same. Hence, the following interaction is expected:

$$\begin{aligned}
 (1) \quad & \mu_{\text{static\_pretest}} = \mu_{\text{dynamic\_pretest}} \\
 & \text{but } \mu_{\text{static\_post-test1}} < \mu_{\text{dynamic\_post-test1}} \\
 & \text{and } \mu_{\text{static\_post-test2}} < \mu_{\text{dynamic\_post-test2}}
 \end{aligned}$$

Second, in addition to an immediate post-test, a delayed post-test (post-test 2) was included in order to observe performance of retention after some weeks.

This outcome could be interesting because studies about statistical misconceptions observed how concepts were learned over time (Lipson, 2002). The e-learning principles – structure and guidance, conflict, explanatory feedback and reflection – are integrated in both interactive versions of the tutorial, static and dynamic. Therefore we expect an increase in knowledge and understanding in post-test 1 (immediate post-test) and post-test 2 (delayed post-test) compared to pretest:

$$(2) \quad \mu_{\text{pretest}} < \mu_{\text{post-test1}} \text{ and } \mu_{\text{pretest}} < \mu_{\text{post-test2}}.$$

The predicted outcomes (1) and (2) are expected for objective as well as for subjective measures.



<http://www.springer.com/978-3-658-08334-2>

Impact of Visual Simulations in Statistics  
The Role of Interactive Visualizations in Improving  
Statistical Knowledge

Iten, G.

2015, IX, 48 p. 6 illus., Softcover

ISBN: 978-3-658-08334-2