How Do Decision Time and Realism Affect Map-Based Decision Making?

Jan Wilkening and Sara Irina Fabrikant

Geographic Information Visualization & Analysis Group (GIVA),
Department of Geography,
University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
{jan.wilkening,sara.fabrikant}@geo.uzh.ch

Abstract. We commonly make decisions based on different kinds of maps, and under varying time constraints. The accuracy of these decisions often can decide even over life and death. In this study, we investigate how varying time constraints and different map types can influence people's visuo-spatial decision making, specifically for a complex slope detection task involving three spatial dimensions. We find that participants' response accuracy and response confidence do not decrease linearly, as hypothesized, when given less response time. Assessing collected responses within the signal detection theory framework, we find that different inference error types occur with different map types. Finally, we replicate previous findings suggesting that while people might prefer more realistic looking maps, they do not necessarily perform better with them.

Keywords: time pressure, slope maps, shaded relief maps, empirical map evaluation.

1 Introduction

Many people have used maps for spatio-temporal decision-making under varying time constraints. For instance, commuters must choose alternative driving routes with road maps, depending on rapidly changing traffic situations. Hikers might need to rapidly select a different trail using a topographic map, when the weather suddenly deteriorates in the mountains, or a sailor might have to quickly consult a nautical chart when navigating in an area with sudden wind and water level changes. Time available for such kinds of map-based decisions can vary enormously. Sometimes the decision time window might consist merely of a few seconds, and the decision might decide over life and death. With increasing human mobility, and respective increased availability of mobile map devices, it seems crucial to investigate how decision time constraints and display types might affect the quality of map-based decision making under rapidly changing conditions. We have been investigating this rather under researched issue with a series of prior controlled experiments, which we review with other relevant research in the related work section.

In this study, involving a map-based slope detection task under varying time pressure scenarios, we specifically investigate decision-making within a three-dimensional context

M. Egenhofer et al. (Eds.): COSIT 2011, LNCS 6899, pp. 1–19, 2011.

and different display types, and analyze collected responses (i.e., accuracy) within the signal detection theory framework.

2 Related Work

Our research program lies at the intersection of time pressure and decision-making research, mainly carried out in psychology and economics, including empirical map design research in cartography. In the following, we review related work from these cognate research fields.

2.1 Decision Making under Time Pressure

Many external factors and cognitive biases impair optimal human decision-making (Simon, 1959; Payne, 1982; Gigerenzer, 2002). The effect of time constraints on decision making has been evaluated systematically by cognitive, developmental, and personality psychologists, as well as by human resources researchers, or economists (see Förster et al., 2003 for an extensive review). It is widely accepted that decision time influences the quality of decisions (Svenson et al., 1990), and that the negative effect of time pressure on decision making is robust and consistent (Pew, 1969; Ahituv et al., 1998).

Two concepts from the psychological literature on time pressure and decision-making seem relevant for our study: Firstly, the s*peed-accuracy trade-off* concept (Wickelgren, 1977) suggests that time pressure can reduce the overall quality of a decision, and, secondly, the *speed-confidence trade-off* (Smith et al., 1982) suggests that the confidence with which people make decisions might decrease with increasing time pressure. The characteristics of the speed-accuracy trade-off depend on task complexity: The more complex a task, the more likely the occurrence of the speed-accuracy trade-off (Johnson et al., 1993).

However, there are also instances when time pressure has a beneficial effect on decision-making. For example, in a long-term time pressure study with NASA scientists and engineers Andrews and Farris (1972) found that decision performance actually increased with increased time pressure, but only to a certain tipping point. Beyond that point, decision performance decreased again. Peters et al. (1984) replicated these findings in a related study involving commercial bankers as decision makers. Hwang (1994) argues that perhaps the best way to describe the interaction between decision performance and time pressure is not a linear relationship, but an inverted U-shaped curve: *"Increasing time pressure leads to better performance up to a certain point, beyond that point more time pressure reduces, rather than increases, performance." (Hwang, 1994, p. 198).*

A still open question remains whether map-based decisions follow a linear speed-accuracy trade-off relation, or an inverted U-shaped curve as found in previous research outside of GIScience, which would imply that time pressure could also have a positive effect on map-based decisions. At this point, it is also unclear how map design and task complexity might interact with map-based decision making under time pressure.

In empirical cartographic research, response time is typically employed as a dependent variable (i.e., efficiency measure) to evaluate cartographic design principles (Lloyd and Bunch, 2003; Garlandini and Fabrikant, 2009; Dillemuth, 2009). However, little work has been done until now to study the effect of time pressure (i.e., as a controlled, independent variable or factor) on map based decision-making. Baus, Krüger and Wahlster (2002) suggest to consider time pressure when designing displays of mobile devices for pedestrian navigation. They argue that changing travelling speeds during navigation create varying time pressure situations, which in turn should lead to different user requirements for navigation displays. They contend that different content should be displayed on a map used in different time pressure conditions. In another study involving user motivation in navigation, Srinivas and Hirtle (2010) offered a reward to one participant group as an incentive for faster task completion, while the other "control" group was not given any incentive to reduce task completion time. Indeed, the "more motivated" participants completed the routes significantly faster than the participants in the "control" group.

2.2 Map Design Issues in Decision Making

Numerous prior empirical studies in cartography have investigated how map design might influence human visuo-spatial inference and decision making, typically depending on a specific map use task (Fabrikant and Lobben, 2009). For our study on slope detection, research comparing 2D and 3D-looking maps for a task involving three spatial dimensions seems particularly relevant. For example, studying aviator navigation performance, Smallman et al. (2001) have shown that users' search time for selecting aircraft which meet certain criteria was significantly faster with 2D maps than with 3D-looking map displays. In related work on the design of cockpit displays, Thomas and Wickens (2006) found no significant performance differences in participants' accuracy and response times between 2D co-planar or 3D perspective displays. Coors et al. (2005) evaluated small-screen 3D and 2D mobile navigation aids, and found that the majority of the participants had a positive attitude towards 3D. Participants found that 3D maps were generally a "good idea", but also that 2D was already "sufficient" for mobile navigation. However, participants' response times were significantly slower with 3D maps compared to 2D maps. This suggests that 3D displays in the context of navigation might be more suitable when having more time available for decision-making, but less useful under time pressure.

In this context, the potential discrepancy between user preferences and actual task performance is also relevant. For instance, Canham, Smallman and Hegarty (2007) and Hegarty et al. (2009) have shown that users tend to prefer more realistic, 3D-looking weather maps that on the surface seem to contain more information for the decision-making task at hand than more abstract 2D maps. However, while users prefer 3D, these displays do not necessarily seem to positively influence users' task performance. In fact, Hegarty and colleagues (2009) found that performance was generally better with the less realistic-looking maps, while users' preference ratings indicated just the opposite. They interpret these results as "another good, empirically validated illustration of the common-sense notion that what people think they want is not always what is best for them" (Fabrikant and Lobben, 2009). According to Hegarty and colleagues, "naïve cartographers" seem to prefer 3D displays to 2D displays, and also seem to prefer more realistic depictions to simpler, more abstract ones. Cartographic design theories and principles, however, aim for reducing graphic complexity (Bertin, 1967). Similarly, the claims by designers for maximizing the data-ink ratio and for minimizing chart junk (Tufte, 1983), or the empirically validated clutter principle by Rosenholtz and colleagues (2007) also call for more abstraction, and less gratuitous realism to facilitate visuo-spatial decision making. From these related studies, we can derive an initial research hypothesis that users prefer more realistic-looking maps (e.g., satellite image maps) and 3D maps (e.g., shaded relief maps), but might perform better with traditional 2D cartographic maps (i.e., topographic maps).

While on the surface it might seem obvious that certain map types are suitable for certain kinds of tasks, it is less obvious how variations of map display designs might influence the quality of map-based decisions under varying temporal usage constraints.

3 Previous Own Work: Experiments and Expert Interviews

In order to fill the existing research gap between time pressure research and empirical map design and map use studies, we have been conducting a series of controlled experiments on map-based decision making under time pressure. We complemented these studies by expert interviews with professionals in the field of map-based decision-making under time pressure. In the following, we summarize the main findings of this work which set the context for the slope detection experiment reported in Section 4.

In a first experiment on map use preferences for a road selection task under various time pressure conditions, we found that participants preferred realistic-looking orthographic satellite image maps and perspective views with hill shaded relief when they were not under time pressure (Wilkening, 2009). However, these preferred image maps were rated significantly less useful when under time pressure. In contrast, preference ratings for the more abstract looking topographic or road maps (i.e., without hill shading) were not affected by time pressure.

In a second experiment, we assessed users' road selection task performance in flat urban terrain. The roads were depicted either on a satellite image map or on a standard road map, under varying time pressure scenarios (Wilkening, 2010). The map display type did not affect participants' accuracy scores. However, participants reported a significantly higher confidence in their performance with satellite images compared to the more abstract road maps. This over-confidence in realistic depictions has been discovered in prior work (Hegarty et al., 2009; Smallman and St. John, 2005; Fabrikant and Boughman, 2006).

In our road selection experiment, shorter decision time limits resulted in a significant decrease in participants' confidence, but not in accuracy. In other words, while we did find a speed-confidence trade-off effect, we did not find strong evidence for a speed-accuracy trade-off.

After having obtained some first insights on map type preferences and task performance under time pressure by non expert map users, we were interested in interviewing professionals who perform map-based decisions under time pressure on a daily basis, specifically within a more complex three-dimensional context. For this reason we interviewed, amongst others, search and rescue helicopter pilots and professionally trained mountain guides. Both professional groups mentioned that they were generally satisfied with using the “classic” 2D topographic map for their routine work. In the age of 3D interactive globe viewers, and location-aware mobile displays, we found that the static, two-dimensional topographic map on paper is still the state-of-the-art for professionals dealing with real world emergency situations under time pressure. One reason could be that the majority of search and rescue personnel have been specifically trained with these maps, can read them well, and thus are generally comfortable with using them. These interviews confirm findings from our first experiment that familiarity (and training) with a display can positively influence usage preference, especially when under time pressure (Wilkening, 2009).

For both helicopter flying and mountaineering activities, accurate slope identification is very important. For example, a helicopter must assess the steepness of the terrain for landing (Bloom, 2007), and for a mountain guide the steepness of a slope needs to be regularly assessed for determining the avalanche potential when on a ski tour during the snow season (Suter, 2007). As the depiction style of the thematically relevant third dimension might be important for these kinds of tasks, we specifically chose a slope detection task for our next experiment, which is described in detail the next sections.

4 Experiment

As mentioned earlier, in own prior work we discovered a significant effect of time pressure on user preferences and response confidence for realistic 3D-looking maps in a 2D task context, while actual performance did not seem to be affected by the verisimilitude of the display. In this study, we are interested how 3D realism might affect participants’ response accuracy and confidence for a task under time pressure that specifically involves decision-making within a 3D context.

We asked task domain novices, that is, people who might be familiar with maps, but have never used maps for landing a helicopter, to identify locations on various map stimuli where a helicopter could land. The previously interviewed professional helicopter pilots had mentioned inclines of less than 14% (or 8 degrees) for safe helicopter landing. This threshold seems to be a standard in the literature (e.g., Bloom, 2007).

We again selected three time pressure scenarios with time limits what were identified through pilot testing. The experiment follows a within-subject design, where each participant was exposed to all time constraints and display types.

4.1 Participants

Fifty-five (32 male and 23 female) participants took part in this study. Participants were either students or staff at the Department of Geography at the University of Zurich and the Institute of Cartography at the Swiss Federal Institute of Technology in Zurich. The majority of participants stated to be “rather familiar” with topographic maps (58.2%) and 3D displays (61.8%), while 32.7% reported to be “very familiar” with topographic maps, and 14.5% to be very familiar with 3D depictions. While our sample represents the more experienced map designer and user, the participants are not experts in the slope detection task domain, and do not represent experts in map-based decision making under time pressure.

4.2 Materials

We created twelve map displays in total, depicting mountainous areas in Switzerland. All maps were of identical size (389x355 pixels), and included a scale bar on the upper right of the display (see Figure 1). The elevation data for the stimuli were derived from the SRTM3 Digital Elevation Model (Jarvis et al., 2008). Slope information could be identified with two pieces of information depicted in the stimulus: the scale bar next to the map and the contour lines in the map. The map scale was held constant at 1:20,000 (run), and the contour line interval was held constant at 100m (rise). The twelve maps represent the elevation data in four different ways:

1. Contour lines only (map a)
2. Contour lines plus light hill shading (map b)
3. Contour lines plus dark hill shading (map c)
4. Contour lines plus colored slope classes (map d)

While all map types are suitable for identifying slope, the depiction methods systematically vary in the degree of the depicted realism (i.e., shaded relief vs. contour line maps), in the apparent visual clutter (i.e., contour line vs. shaded relief maps), and in the information content for detecting slope suitability (slope vs. contour line maps). In other words, the maps are neither computationally nor informationally equivalent (Simon and Larkin, 1987; Fabrikant et al., 2008).

Map (a) in Figure 1, containing only contour lines, represents the most abstract of the tested map types, and with the least amount of information (i.e., implicit slope information). Maps (b) and (c) additionally contain a shaded relief (i.e., explicit relative slope information), thus more information than map (a). Users can obtain slope information not only (implicitly) from the distance between the contour lines (a), but also from the relative darkness of the pixels (maps b+c). The steeper the slope, the darker is the appearance of the relief. To investigate the potential effect of the graphic quality of the hill shading, we created a lighter version (b) and a darker version (c) of the hill shaded relief maps. We employed the hill shade function available with the 3D Analyst Toolbox in ESRI's ArcGIS. The light source for the hill shading was set to a 45° angle for the lighter relief (map b) and to a 22° angle for the dark relief (map c), respectively.

Based on Simon and Larkin (1987), we hypothesize that the more implicit the depiction of the task relevant information (i.e., map a), and the higher the amount of task irrelevant information (i.e., maps b+c) the more reasoning effort is needed when making decisions with these maps.

While our interviewed map-based decision-making professionals did not use slope maps (map d) for their daily work, they considered them as "nice-to-have", so we included them in our study. The slope maps contain most task-relevant information in our tested maps. They show slope information explicitly in the map, and the respective information is explained in the accompanying legend, thus, should be easiest to use for task domain novices. Slope was calculated in ESRI's ArcGIS and depicted in a diverging color scheme, employing the traffic light metaphor (green = go, red = stop). Slopes that are flat enough for a helicopter to land (i.e., below 14% steepness) are depicted with green shades, while slopes that are too steep for landing (i.e., above 14%) are shown in magenta shades (Figure 1d). We define amount of realism as a degree of verisimilitude with the real world (Zanola et al., 2009). We thus contend that shaded relief maps (b+c in Figure 1) look more realistic than a contour map (Figure 1a), because contours cannot be seen in the real world.

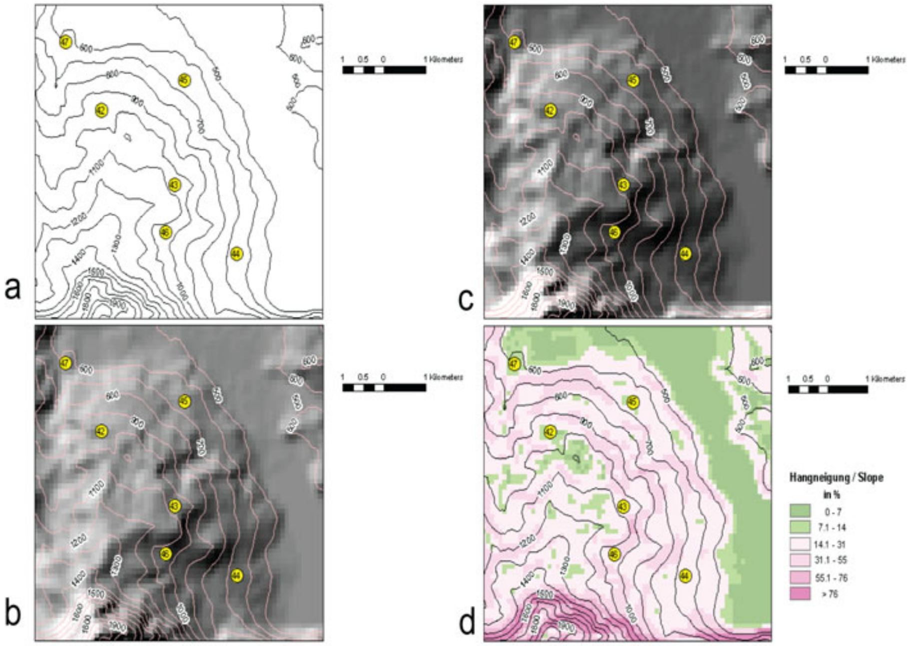

Fig. 1. Reduced examples of employed map stimuli: only contour line map (a), light hill shaded relief map (b), dark hill shaded relief map (c), and slope map (d)

Finally, the four tested map types also vary in graphic quality, or negatively put, in their degree of visual clutter. In Rosenholtz et al. (2007)'s terminology, clutter relates to the degree of perceptual organization of information in a display. The more organized a display, the less visual clutter (detracting information) it contains for a given task. We quantitatively assessed this purely bottom-up vision concept in our test displays by means of the Subband Entropy clutter measure, proposed by Rosenholtz and colleagues. This measure, also empirically validated with map displays, is "based on the notion of clutter as related to the efficiency with which the image can be encoded and inversely related to the amount of redundancy and grouping in the image" (Rosenholtz et al., 2007, p. 18). Subband entropy seems to be a good predictor for human map-reading performance under time pressure. The higher the subband entropy measure for a display (i.e., the more clutter), the less computationally efficient the extraction of information encoded in the image (Simon and Larkin 1987). To exemplify this measure, we computed subband entropy for the four stimuli shown in Figure 1, and find most clutter in the slope map (3.75), followed by the dark hill shaded relief (3.31), the light hill shaded relief (3.27), and lastly, the contour map (3.25). While the slope map is graphically more cluttered than the others (e.g., it includes an additional visual variable color), it shows the task relevant information explicitly (i.e., in the legend), thus despite perceptual clutter one would expect this map to need less reasoning effort to extract the task relevant information. The investigated factors are summarized in Table 1 below.

Table 1. Comparison of the map types used in the slope detection experiment. The amount of information is indicated with + (low) to +++ (very high).

	contour map	shaded relief maps	slope map
degree of realism	(+)	(++)	(+)
depiction type (elevation information)	lines of equal elevation (absolute)	lines of equal elevation (absolute) & shaded relief (relative)	lines of equal elevation (absolute) & slope classes (absolute)
slope information type (amount)	implicit (+)	implicit (++)	explicit (+++)
visual clutter (Subband Entropy)	(+)	(++)	(+++)
reasoning effort	(+++)	(++)	(+)

The locations participants had to assess for potential helicopter landing were represented with black labels (numbers) on a yellow background to maximize saliency. Each stimulus contained six such locations for assessment. No other pieces of information (such as labels of place names) were contained in the map. We ran a saliency model (Itti and Koch, 2001) on our test stimuli, to make sure that the saliency of the decision points was not significantly influenced by the tested map types.

4.3 Procedure

The experiment took place in a lab equipped with standard personal computers connected to the Internet. The experiment was carried out digitally in a web browser displayed on a 17-inch computer screen set to 1280x768 pixel screen resolution. After filling out a background questionnaire, participants were then asked to safely land a helicopter on slopes not steeper than 14%. To assure that participants all had the necessary background to complete the task, they were first introduced to the slope concept and how slope can be calculated. They were shown how slope can be identified in a contour line map using the elevation information displayed with labels on the contour lines, and the ground distance information contained in the map scale bar. No other task relevant information was given to the participants. Participants were then asked to solve two warm-up tasks, which were identical to the actual experiment, which is described below.

Then, they were shown the sequence of twelve maps described in the previous section. The order of the stimuli was systematically rotated to prevent learning biases due to potential ordering effects. For each map, participants had to select one or more locations that were flat enough for a helicopter to land, by clicking the respective checkbox below the map. For each map, six locations had to be assessed. The number of correct locations varied randomly from 1 to 5 per map. Overall, 50% of the labeled slopes were too steep to land a helicopter.

Subjects had to solve the slope detection task under all three time constraint conditions, including 20s (most severe), 40s (moderate), and 60s (least severe) time limits and for all map display types described earlier. After completing each task, participants were asked to rate their confidence of response on a scale from “1 – not confident at all” to “4 – very confident”. Participants were not under time pressure when asked to rate their response confidence. Responses were collected digitally and included participants’ accuracy (percentage of correct answers) as well as (self-reported) confidence as success measures. After completing the digital portion of the experiment, participants were debriefed, and given a meal voucher for the university cafeteria in return for their participation. The experiment took approximately 15 minutes to complete.

4.4 Signal Detection Analysis

The conceptual framework of signal detection theory (SDT), which was originally developed for research on visual perception (Tanner and Swets, 1954), can generally be employed for decision-making under uncertainty, and especially when decisions have to be made based on two or more alternatives. The benefits of using this framework for our research context is that response accuracy can be assessed with more analytical depth than just comparing correct and false answers. In SDT correct answers are coined “hits” or “correct rejects”, and errors are called “misses” or “false alarms”, respectively. This analysis framework can especially help us to identify which kinds of errors participants might make, due to varying time constraints and map display types, and thus if errors might follow a particular pattern.

Applying this concept to our slope detection experiment, correctly selected locations per question are classified as “hits” (<14% steepness), and thus those (correctly) not selected locations are classified as “correct rejects” (>14%, see Table 2 below). Participant answers that are incorrect are classified as either “misses” or “false alarms”, respectively. A miss indicates a location that was not selected, even though it is correct (<14% steepness), and a false alarm occurs when participants incorrectly selected a location with a slope that is too steep (>14%). In other words, a miss is an overestimation of slope, while a false alarm represents the underestimation of a slope. Table 2 illustrates how we classify four possible types of responses within the data analysis framework of Signal Detection Theory.

Table 2. Classification of correct and false answers according to Signal Detection Theory

	Reality	
Participants’ decisions	Slope too steep (> 14%)	Slope flat enough (<14%)
Slope too steep (> 14%)	correct reject (true)	miss (false)
Slope flat enough (<14%)	false alarm (false)	hit (true)

5 Results

We first present the results regarding time pressure, followed by the results for the different map types, and finally, we report on the interaction of time pressure with map types on participants’ accuracy and confidence ratings.

5.1 Time Pressure Effect

Overall participants' average response accuracy and confidence ratings shown in Figure 2 reveal a surprising and counterintuitive pattern. Participants are most accurate with the moderate time limit of 40s (M=82.8%, SD=11.4%), followed by the most generous time limit with 60s (M=72.2%, SD=15.2%), and lastly, as expected, the most severe time limit of 20s (M=66.7%, SD=20.4%). Similarly, participants' self-reported confidence is also highest for the moderate 40s time limit (M=2.86, SD=0.44), followed by the least severe 60s limit (M=2.73, SD=0.15) and, lowest, again as expected, for the most severe 20s limit (M=2.67, SD=0.20). The performance increase from the 20s time limit to the 40s limit, and the performance decrease from the 40s to the 60s time limit are all significant for both accuracy and confidence (p< .001).

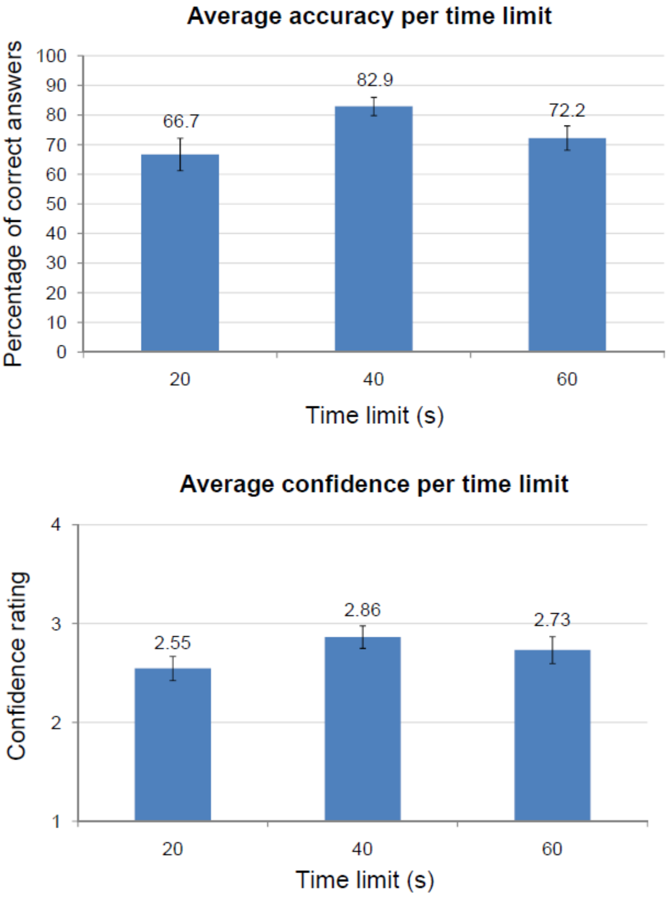

Fig. 2. Average accuracy and confidence per time pressure limit. Error bars: ±2 Standard Error (SE).

5.2 Map Type Effect

As expected, participants' accuracy was significantly better with the slope map (M=83.6%, SD=14.8%) compared to all other maps, as shown in Figure 3. However, accuracy was not better as predicted, but even worse with the shaded relief maps compared to all other maps. Participants' mean accuracy for the light hill shaded relief map is 73.1% (SD=16.8%), and with 65.4% (SD=18.5%) it is lowest overall for the dark hill shaded relief map. Surprisingly, participants perform even worse with the hill shaded relief maps that look more realistic, and contain more information than the most abstract contour map (M=73.5%, SD=13%). The difference between the contour map and the light hill shaded relief map is significant (p< .01), as well as the difference between the slope map and all other maps (p< .001).

As can be seen in Figure 3, in congruence with the accuracy response pattern, participants' confidence ratings are also highest for the slope map (M=3.13,

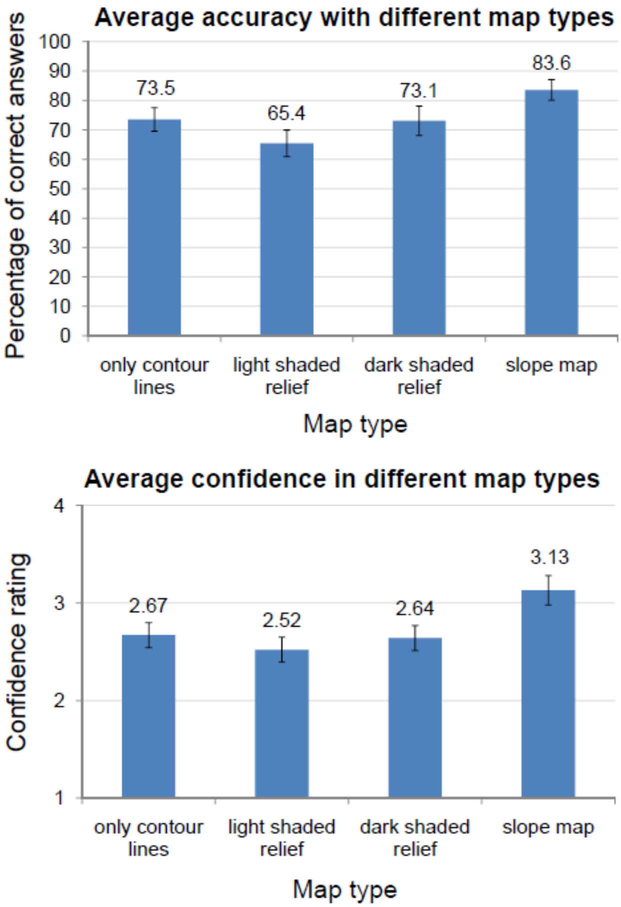

Fig. 3. Accuracy and confidence ratings for tested map display types (± 2 SE)

SD=0.55), and higher for the contour maps (M=2.67, SD=0.47), compared to the lowest scoring hill shaded relief maps (dark: M=2.64, SD=0.47 and light: M=2.52, SD=0.47). Surprisingly again, participants have significantly higher confidence in their performance with the contour map, compared to the light hill shaded relief map ($p < .05$) that contains more information.

The power of signal detection theory lets us analyze response accuracy in more detail. Overall, regardless of map type, misses (i.e., slope overestimation) occur more frequently than false alarms. Misses occur also more frequently than false alarms, independent of the tested time limits. The “correct rejection” is overall the more frequent correct answer than the “hit”, also for all map types and all temporal conditions. As expected, the number of false alarms (i.e., slope underestimation) shown in Figure 4 is significantly higher for the light hill shaded relief map (M=2.00, SD=1.75) than for the dark hill shaded relief map (M=1.18, SD=1.23). In contrast, the number of misses (i.e., slope overestimation) is, again as expected, higher with the darker hill shaded relief map (M=2.67, SD=1.77) compared to the light shaded relief (M=2.05, SD=1.69).

As shown in Figure 4, SDT provides additional insights on what kinds of decision errors might have specifically contributed to the unexpectedly low accuracy for the shaded relief maps. Similar to the other map types, participants seem to overestimate the steepness of the slopes more frequently with the dark hill shaded relief maps (i.e., higher number of misses) compared to the light shaded relief maps. Hence, a map with a lighter shaded relief might help reduce this potential source of error. However, one can also see in Figure 4 that one drawback of light hill shaded relief maps might be their relatively high rate of false alarms.

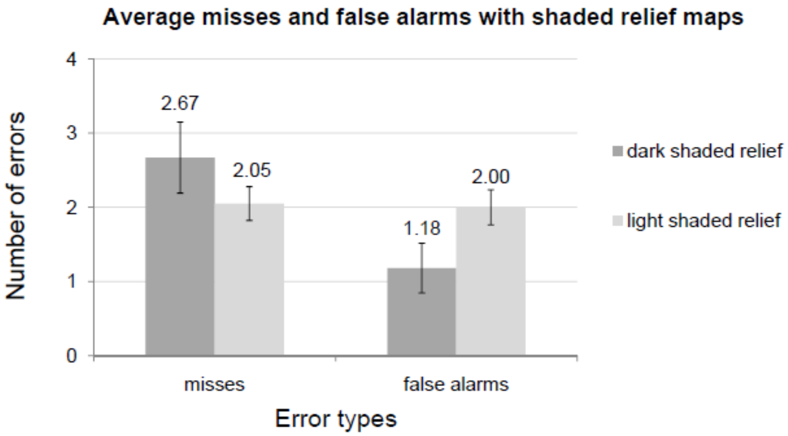

Fig. 4. Misses and false alarms with shaded relief maps (± 2 SE). The maximum number of possible errors is 9 per map type.

5.3 Interaction of Map Type and Time Pressure

We now turn to the research question how map types might support participants in their decision-making under varying time pressure scenarios. Participants gave most accurate answers with the (explicit) slope map under all time constraint conditions (see Figure 5). In the most severe time limit condition (20s), participants scored better with the most abstract contour map (M=71.2%, SD=30.5%), containing least amount of information, compared to the more realistic looking shaded relief maps (dark: M=63.9%, SD=27.9% and light: M=53.3%, SD=40.1%). For this shortest time limit, the overall differences between tested map types are significant (p< .01 for both shaded relief maps). Accuracy scores generally increase from the most severe (20s) to the moderate (40s) time limit condition. The accuracy differences between maps are not significant in the moderate condition. Overall, accuracy scores drop again for the highest scoring slope and contour maps under the least severe time constraint condition (60s), while accuracy scores for the hill shaded relief maps do not change much for the 40s and 60s limit conditions. In other words, participants' accuracy with hill shaded relief maps only reaches the higher level of the other more abstract map types when participants are not under severe time pressure.

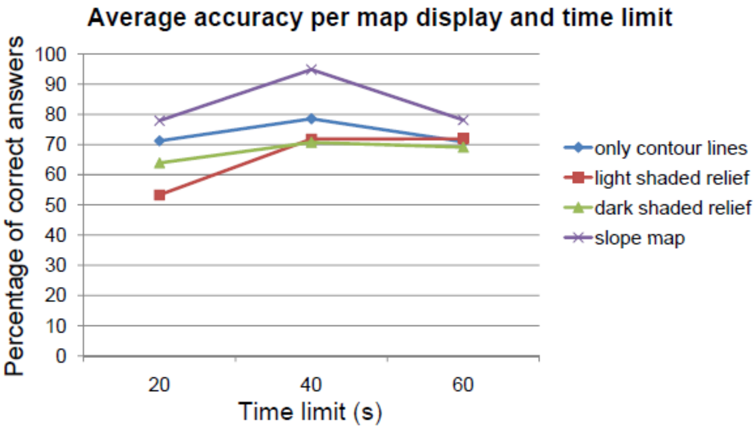

Fig. 5. Participant average accuracy per map display type and time limit

A very similar response pattern can be observed in Figure 6, when looking at participants' confidence ratings. Again, mirroring accuracy scores, participants are most confident in their responses with the slope map, regardless of the given time limit. Participants' confidence is also consistently high with the contour line map.

The difference between the average confidence ratings for the slope map and the shaded relief maps is only significant at the 20s time limit. For this shortest time limit, the average confidence rating with the contour map is 2.58 (SD=0.08) and 2.30 (SD=0.07) with the shaded relief map. The rating difference between the contour map and both hill shaded relief maps is significant (p< .001).

Only in the moderate time limit condition (40s), confidence ratings for the hill shaded relief maps are higher than for the contour map. This is in contrast to participants' accuracy scores shown in Figure 5 earlier, where participant performance is better with the contour map that the shaded relief maps.

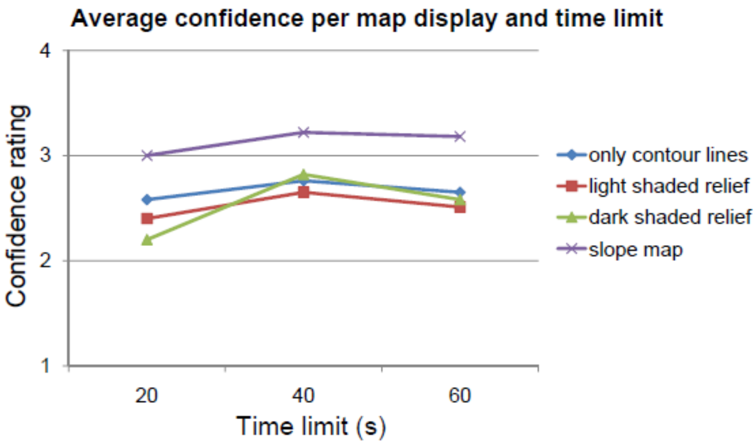

Fig. 6. Participant average confidence per map display and time limit

6 Discussion

Summarizing our results, we find that indeed, response accuracy and confidence ratings are worst when under highest time pressure, but best when participants are under a moderate response time limit. Both scores decrease significantly when participants have more decision time available. These results, which seem somewhat counterintuitive, do resemble the previously discovered inverted U-shaped response curve found by Hwang (1994), but not related to map-based decision making.

Based on Johnson et al.'s (1993) and Hwang's (1994) research results reviewed earlier, changes in speed-accuracy and speed-confidence trade-offs might be a consequence of task difficulty. As both response accuracy and confidence decreased with a time limit more severe than 40s, this slope detection task might have become significantly more difficult when participants had less than 40s to respond. As a result, we do find a clear speed-accuracy and speed-confidence trade-off effect. Participant performance did not further increase from the moderate to the least severe time limit, thus the slope detection task is not getting easier with more available decision time beyond the 40s time limit. In this study, the tipping point to which time pressure actually increases performance seems to be in the vicinity of 40s decision time. This pattern is in contrast to our previous road selection task study, where we did not find a time pressure effect on response accuracy, even with overall shorter decision time limits, down to even 10s decision time. One could argue that the road selection task on flat terrain is significantly less complex than a 3D slope detection task, and thus speed-accuracy and speed-confidence trade-offs are generally harder to find. This difference in task complexity might be one of the main reasons for the non-repeatability of the results of Experiment I.

Regarding the decision performance differences due to different the map types, surprisingly, participants' accuracy and response confidence was unexpectedly low with the shaded relief maps. This result supports prior work by Hegarty and colleagues (Hegarty et al., 2008; Hegarty et al., 2009), who have shown that more realistic, 3D-looking displays while often preferred by “naïve users”, do not necessarily increase performance. While three-dimensional shaded reliefs provide more task-relevant (but implicit) information, compared to the more abstract contour map, this more on information does not lead to more effective (accurate) or efficient (faster) decision-making. One reason for this is might, arguably, be that the implicit thematically relevant information is not presented in a cognitively and perceptually adequate and inspired way (Swienty et al., 2008; Fabrikant et al., 2010). While the hill shaded relief maps might contain more explicit task-related information than the contour maps, they are also more cluttered (Rosenholtz et al., 2007), and thus might require more time for participants to visually parse. As Tufte (1983) would put it, the task-relevant data to graphic ink ratio in the visuo-spatial display is not optimized for the task at hand. On the other hand, while the slope maps exhibit the highest clutter values of the tested displays (see Table 1), the task-relevant data to graphic ink-ratio is indeed optimized for the task at hand. In fact, running a saliency model (Itti and Koch, 2001) on the stimuli, we find only one significant difference between the slope map and the other three tested map types (see Figure 1). The area along the bottom edge of the maps, where the density of the elevation contours is highest (i.e., the steepest area in the map), the slope map also shows darkest magenta shades between the contour lines. Moreover, the visual variable color hue seems not to have much influence on this saliency map pattern, as running the saliency model on a gray scale version of the slope maps (i.e., removing color hue) yields an identical saliency map pattern. Another possibility why the more abstract maps could have performed better under time pressure is that our 3D maps with high graphic density might have a general relative disadvantage when shown at smaller screen sizes with lower spatial display resolution than the 2D maps.

However, participants do perform better with the shaded relief maps compared to the more abstract contour line map when they have more decision time available, and also seem to be more confident in their responses when under less time pressure. In this case, participant performance and confidence seem to reflect participant preferences, when we compare results from this study with the results from a prior map use preference experiment (see previous work Section 3), in which more realistic 3D looking satellite image maps obtained higher preference ratings when participants have more decision time available. In other words, we did not find strong evidence for a “naïve realism” effect (Smallman and St. John, 2005), or over-confidence in realistic-looking maps in this experiment, as low accuracy scores co-occurred with equally low confidence ratings for the tested shaded relief maps. This could be due to the fact that our participant sample consisted mainly of cartographic (design) professionals, and thus not “naïve” cartographers.

Not surprisingly, the 2D slope maps, containing most of the thematically relevant information, outperformed all other map types with respect to effectiveness (i.e., accuracy) and efficiency (i.e., under all time limits), including participant confidence. In this case, in contrast to the shaded relief maps, the information increase had a positive effect on response accuracy and confidence, even though perceptually these maps appeared to be most cluttered (see Table 1). Reasons for this could be that the slope map already explicitly contains an intrinsic reasoning step (i.e., slope computation). This more on thematically relevant information is communicated in a cognitively adequate (explicit), and perceptually salient way, using empirically validated cartographic design principles (Fabrikant, Rebich-Hespanha and Hegarty 2010). In other words, participants can perform well and be confident in their decisions even with an abstract (but computationally efficient) depiction method, but only when thematically relevant information is communicated explicitly and rendered in a perceptually salient manner. It would be thus interesting to further investigate how different ways of representing slope information might affect the outcomes of map-based decision making tasks under time pressure. Although slope maps are not commonly known or used by map-based decision-making experts under time pressure, or the general public, our expert interviewees did find them useful, and had no problem in detecting the relevant information without any training.

7 Summary and Outlook

In this study, we investigated how display types might affect people's decision making when solving a complex slope detection task under varying time pressure conditions. Replicating previous work (Andrews and Farris, 1972; Hwang, 1994) we discovered an inverted U-shaped accuracy response curve which implies that moderate time pressure can have a positive effect on map-based decision-making, but only up to a certain tipping point, which seems to be around 40s in our study. Moreover, confirming long-standing (but rarely empirically validated) cartographic design theory (Bertin, 1967), we found that more abstract, but well designed contour and slope maps outperform more preferred realistic, 3D-looking hill shaded relief maps for the 3D slope detection task in our study. This might suggest that the benefit of explicitly communicating thematic relevant information, even in a graphically abstract way (i.e., higher cognitive cost), is greater for efficient and effective map-based decision-making, than adding preferred and attractive, but visually more cluttered realism (i.e., higher perceptual cost). Low participant performance with shaded relief maps—even lower than the more abstract contour maps, containing even less information—suggest that visual realism might negatively influence decision-making, especially when under time pressure.

Future experiments in varying map-based decision making contexts with different task complexity levels should be conducted to further investigate the generalizability of these somewhat counterintuitive findings, involving 1) performance decreases with more available decision time, and 2) surprisingly poor performance with shaded relief maps. For example, one could vary display sizes and the ways of representing slope information, in order to investigate the robustness of our findings.

It is unclear at this point how performance is affected by user background and training. In future related studies, participants with less cartography training could be tested in similar time pressure contexts, in order to compare previous results by Hegarty and colleagues (2009), who have found higher preferences for 3D maps among "naïve cartographers".

Finally, we also encourage like minded researchers in GIScience and cartography to more often try to analyze response accuracy with the signal detection approach and to explore in which context misses or false alarms are the dominant types of errors, and how individual and group differences might influence hit and false alarm rates. For example, we found a higher number of misses compared to the number of false alarms in our experiment. One explanation for this play-safe strategy in this safety-critical task context could be that our participants, not trained in helicopter landing, might have rather preferred to miss a suitable landing spot, than landing on unsuitable terrain, with potentially life-threatening consequences (e.g., see the work of Hofer and Schwaninger (2005) relating to baggage screening tasks).

Future empirical map design and map use studies could thus focus more on the question of what kinds of errors might result in low accuracy rates, and this might in turn lead to more focused map design guidelines.

Acknowledgements. We would especially like to thank all our participants who were willing to participate in this study. We are also indebted to Christian Häberling from the Institute of Cartography at the Swiss Federal Institute of Technology Zurich, who gave us the opportunity to conduct our experiments in his cartography classes. Finally, we are grateful to our interviewees at Protection & Rescue (Schutz & Rettung) and Swiss Air Rescue (Rega) in Zurich, and at the Institute for Snow and Avalanche Research (SLF) in Davos, Switzerland, for sharing their expert insights on map-based decision making under time pressure.

References

Ahituv, N., Igbaria, M., Sella, A.: The Effects of Time Pressure and Completeness of Information on Decision Making. Journal of Management Information Systems 15(2), 153–172 (1998)

Andrews, F.M., Farris, G.F.: Time Pressure and Performance of Scientists and Engineers: A Five-Year Panel Study. Organizational Behavior and Human Performance 8(2), 185–200 (1972)

Baus, J., Krüger, A., Wahlster, W.: A resource-adaptive mobile navigation system. In: Proceedings of the 7th International Conference on Intelligent User Interfaces, pp. 15–22. ACM, New York (2002)

Bertin, J.: Semiologie graphique, Mouton, Paris (1967)

Bloom, G.S.: Helicopters: How They Work (2007), `http://www.helicopterpage.com/html/forces.html` (last accessed June 16, 2011)

Canham, M., Smallman, H., Hegarty, M.: Using Complex Visual Displays: When Users Want More than is Good for Them. In: Mosier, K., Fischer, U. (eds.) Proceedings of the Eighth International NDM Conference, Pacific Grove, CA (2007)

Coors, V., Elting, C., Kray, C., Laakso, K.: Presenting Route Instructions on Mobile Devices: From Textual Directions to 3D Visualization. In: Dykes, J., Kraak, M.J., MacEachren, A.M. (eds.) Exploring Visualization, pp. 529–550. Elsevier, Oxford (2005)

Dillemuth, J.: Navigation Tasks with Small-Display Maps: The Sum of the Parts Does Not Equal the Whole. Cartographica 44(3), 187–200 (2009)

Fabrikant, S.I., Boughman, T.: Communicating Data Quality through Realism. In: Proceedings of GIScience 2006, pp. 59–60. Münster, Germany (2006)

Fabrikant, S.I., Lobben, A.: Introduction: Cognitive Issues in Geographic Information Visualization. Cartographica 44(3), 139–143 (2009)

Fabrikant, S.I., Rebich-Hespanha, H., Hegarty, M.: Cognitively Inspired and Perceptually Salient Graphic Displays for Efficient Spatial Inference Making. Annals of the Association of American Geographers 100(1), 13–29 (2010)

Fabrikant, S.I., Rebich-Hespanha, H., Andrienko, N., Andrienko, G., Montello, D.R.: Novel Method to Measure Inference Affordance in Static Small-Multiple Map Displays Representing Dynamic Processes. The Cartographic Journal 45(3), 201–215 (2008)

Förster, J., Higgins, E.T., Taylor Bianco, A.: Speed/accuracy decisions in task performance: Built-in trade-off or separate strategic concerns? Organizational Behaviour and Human Decision Processes 90, 148–164 (2003)

Garlandini, S., Fabrikant, S.I.: Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization. In: Hornsby, K.S., Claramunt, C., Denis, M., Ligozat, G. (eds.) COSIT 2009. LNCS, vol. 5756, pp. 195–211. Springer, Heidelberg (2009)

Gigerenzer, G.: The Adaptive Toolbox. In: Gigerenzer, G., Selten, R. (eds.) Bounded Rationality: The Adaptive Toolbox, Dahlem Workshop Reports, pp. 37–50 (2002)

Hegarty, M., Smallman, H.S., Stull, A.T.: Decoupling of Intuitions and Performance in the Use of Complex Visual Displays. In: Love, B., McRae, K., Sloutsky, V. (eds.) Proceedings of the 30th Annual Conference of the Cognitive Science Society, Washington, DC, pp. 881–886 (2008)

Hegarty, M., Smallman, H.S., Stull, A.T., Canham, M.: Naive Geography: How Intuitions about Display Configuration Can Hurt Performance. Cartographica 44(3), 171–186 (2009)

Hofer, F., Schwaninger, A.: Using threat image projection data for assessing individual screener performance. In: Brebbia, C.A., Bucciarelli, T., Garzia, F., Guarascio, M. (eds.) WIT Transactions on the Built Environment (82), Safety and Security Engineering, vol. 82, pp. 417–426. WIT Press, Wessex (2005)

Hwang, M.I.: Decision making under time pressure: a model for information systems research. Information and Management 27, 197–203 (1994)

Itti, L., Koch, C.: Computational Modelling of Visual Attention. Nature Reviews Neuroscience 2(3), 194–203 (2001)

Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E: Hole-filled seamless SRMT data V4: International Centre for Tropical Agriculture, CIAT (2008)

Johnson, E.J., Payne, J.W., Bettman, J.R.: Adapting to Time Constraints. In: Svenson, O., Maule, J. (eds.) Time Pressure and Stress in Human Judgment and Devision Making, pp. 103–1116. Plenum Press, New York (1993)

Lloyd, R.E., Bunch, R.L.: Technology and Map-Learning: Users, Methods and Symbols. Annals of the Association of American Geographers 93(4), 828–850 (2003)

Payne, J.: Contingent decision behavior. Psychological Bulletin 93, 382–402 (1982)

Peters, L.H., O'Connor, E.J., Pooyon, A., Quick, J.C.: The Relationship between Time Pressure and Performance: A Field Test of Parkinson's Law. Journal of Occupational Behaviour 5, 293–999 (1984)

Pew, R.W.: The speed-accuracy operating characteristics. In: Koster, W.G. (ed.) Acta Psychologica 30: Attention and Performance, vol. II, pp. 16–26. North-Holland Publishing Company, Amsterdam (1969)

Rosenholtz, R., Li, Y., Nakano, L.: Measuring Visual Clutter. Journal of Vision 7(2), 1–22 (2007)

Simon, H.: Theories of decision making in economics and behavioural science. American Economic Review 49(3), 253–283 (1959)

Simon, H.A., Larkin, J.H.: Why a diagram is (sometimes) worth then thousand words. Cognitive Science 11, 65–100 (1987)

Smallman, H.S., St. John, M.: Naive Realism: Misplaced Faith in Realistic Displays. Ergonomics in Design 13(3), 6–13 (2005)

Smallman, H.S., St. John, M., Oonk, H.M., Cowen, M.B.: Information Availabity in 2D and 3D Displays. Applied Perception 21(5), 51–57 (2001)

Smith, J.F., Mitchell, T.R., Beach, L.R.: A cost-benefit mechanism for selecting problem-solving strategies: Some extensions and empirical tests. Organizational Behaviour and Human Performance 29(3), 370–396 (1982)

Srinivas, S., Hirtle, S.C.: The Role of Motivation and Complexity on Wayfinding Performance. In: Purves, R., Weibel, R. (eds.) GIScience 2010 Extended Abstracts Volume, Zurich, Switzerland (2010)

Suter, C.: Lawinenausbildung mit mobilen Systemen. In: Unpublished Diploma Thesis at the Department of Geography, University of Zurich, Switzerland (2007)

Svenson, O., Edland, A., Slovic, P.: Choices and judgments of incompletely described decision alternatives under time pressure. Acta Psychologica 75, 153–169 (1990)

Swienty, O., Reichenbacher, T., Reppermund, S., Zihl, J.: The Role Of Relevance and Cognition in Attention-guiding Geovisualization. The Cartographic Journal 45(3), 227–238 (2008)

Tanner, W.P., Swets, J.A.: A Decision-Making Theory of Visual Detections. Psychological Review 61, 401–409 (1954)

Thomas, L.C., Wickens, C.D.: Display Dimensionality, Conflict Geometry, and Time Pressure Effects on Conflict Detection and Resolution Performance Using Cockpit Displays of Traffic Information. The International Journal of Aviation Psychology 16(3), 321–342 (2006)

Tufte, E.R.: The Visual Display of Quantitative Information. Graphics Press, Cheshire (1983)

Wickelgren, W.A.: Speed-accuracy tradeoff and information processing dynamics. Acta Psychologica 41(1), 67–85 (1977)

Wilkening, J.: User Preferences for Map-Based Decision Making Under Time Pressure. In: COSIT 2009 Doctoral Colloquium Proceedings, Aber Wrac'h, France, pp. 91–98 (2009)

Wilkening, J.: Map Users' Preferences and Performance under Time Pressure. In: Purves, R., Weibel, R. (eds.) GIScience 2010 Extended Abstracts Volume, Zurich, Switzerland (2010)

Zanola, S., Fabrikant, S.I., Çöltekin, A.: The Effect of Realism on the Confidence in Spatial Data Quality in Stereoscopic 3D Displays. In: Proceedings of the 24th International Cartography Conference (ICC 2009), Santiago, Chile, November 15-21 (2009) (on CD-ROM)