Measuring Insight Into Multi-dimensional Data from a Combination of a Scatterplot Matrix and a HyperSlice Visualization

André Calero Valdez[1], Sascha Gebhardt[2], Torsten W. Kuhlen[2], and Martina Ziefle[1]

Human-Computer Interaction Center, Campus-Boulevard 57, RWTH Aachen University, Germany
{calero-valdez,ziefle}@comm.rwth-aachen.de

Virtual Reality Group, Kopernikusstr. 6, RWTH Aachen University, Germany
{gebhardt,kuhlen}@vr.rwth-aachen.de

Abstract. Understanding multi-dimensional data and in particular multi-dimensional dependencies is hard. Information visualization can help to understand this type of data. Still, the problem of how users gain insights from such visualizations is not well understood. Both the visualizations and the users play a role in understanding the data. In a case study, using both, a scatterplot matrix and a HyperSlice with six-dimensional data, we asked 16 participants to think aloud and measured insights during the process of analyzing the data. The amount of insights was strongly correlated with spatial abilities. Interestingly, all users were able to complete an optimization task independently of self-reported understanding of the data.

Keywords: Information Visualization; Insight; Multi-dimensional Visualization; Scatterplot; HyperSlice; Evaluation

1 Introduction

In a heavily technology-assisted work environment, many tasks will shift from manual labor to monitoring, control, and decision-making tasks in the future. To fulfill these tasks process data must be integrated. The underlying data that will be used for these purposes is often intricately interconnected and has multivariate dependencies (e.g., in control-parameters in laser-based welding). Therefore, many of the problems that workers will have to deal with in the future are highly complex. One approach to address this problem is accessing data by visualizations.

An intrinsic attribute of multi-dimensional data is that for visualizing data on a 2D-screen, dimensionality-reduction techniques (e.g., statistical reduction, projections, higher-order data) are necessary. Here lies a core challenge of multidimensional data visualizations. High-level dependencies are not easily visualized or detected, when the dependencies are embedded in the dimensions that are hidden by the dimension reduction approaches.
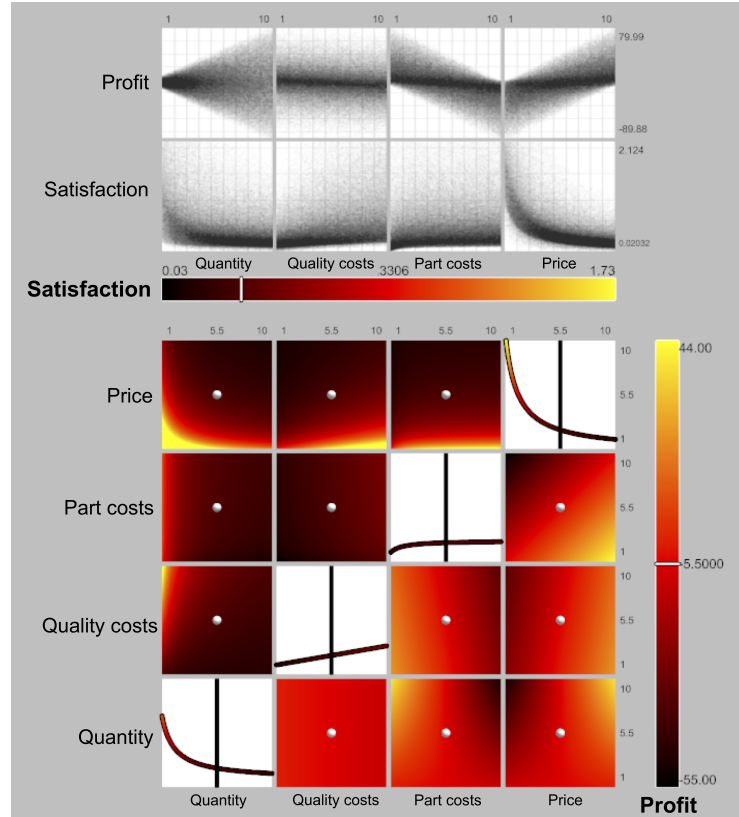
Fig. 1. Example of a scatterplot matrix in combination with a HyperSlice visualization.

One approach to overcome this barrier is HyperSlice [1] (see Fig. 1). This visualization technique displays all pairwise orthogonal, axis-aligned slices through one common point, the focal point, of a multi-dimensional data space. By swiping through the multi-dimensional hyperspace, users can try to reconstruct this space in their mind. While the approach works well in the reconstruction of 3D models from 2D slices, higher dimensions are harder [2] to reconstruct in a mind shaped by evolution in a 3D world. Research even indicates that 5-way interactions can not be processed by the human mind [3].

However, users should be able to analytically think about multi-dimensional dependencies and derive them from visualizations such as HyperSlice. But, all users? Spatial cognitive abilities of users are different. Intelligence differs, experience differs, self-efficacy differs. The question we ask ourselves is: how important are these user differences when using multi-dimensional data visualizations?

In order to get a first understanding of how multi-dimensional data is understood, we conduct a qualitative user study and measure user diversity criteria to understand their influence on visualization insight.

2   Related work

The question of how to assess the quality of a visualization is hard, because all aspects relevant for a visualization come in a plethora of options [4,5]. The purpose of a visualization can be very different (e.g., reveal new facts of old data, monitor real-time sensor data, visual proofs, etc.), the data can be very different (e.g., static, dynamic, high-dimensional, structured, etc.) and the visual representation can be very different (e.g., HyperSlice, star-coordinates, Cherrnoff-faces, etc.). In this article we address visualizations that are made to be insightful and address multi-dimensional data. We focus on a combination of two visualizations: a scatterplot matrix and a HyperSlice visualization.

2.1   Multi-dimensional Data Visualization

A large body of research exists that is relevant to multi-dimensional data visualization. Each have different benefits and drawbacks [6] depending on their usage. A scatterplot matrix can be used to visualize the probability distributions of multiple variables. Columns and rows indicate two variables, and the cell plot is a scatter plot of these two variables (see Fig. 2 [7]).
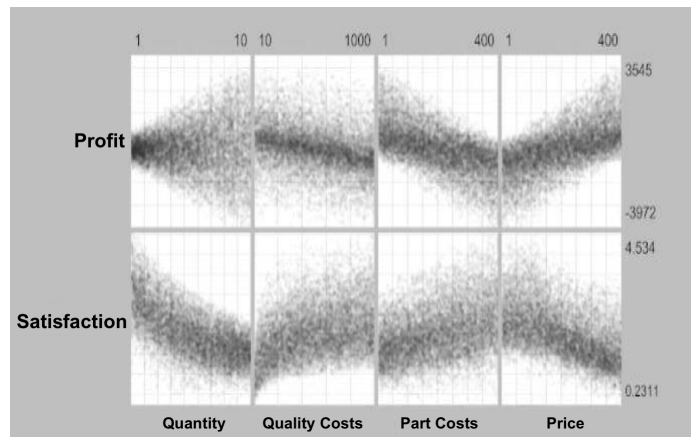

Fig. 2. memoSlice scatterplot matrix. The first row shows the scatterplots for profit and four input variables. The second row shows the scatterplots for customer satisfaction and the same input variables.

A multi-dimensional data visualization that is regularly used is HyperSlice [8]. It represents multiple variables in several plots. In its matrix layout columns and rows represent input variables. Each tile of this layout depicts a slice that maps these inputs to an output variable via color coding.

When the task in such a visualization is to find optimal parameters it can be helpful to add gradient trajectories [9,10]. These indicate the steepest ascent and descent in the mutli-dimensional space. Thus, they ultimately lead to the next local extrema and are a useful tool for optimization tasks. By combining both, the scatterplot matrix and the HyperSlice it is possible to re-use the spatial encoding of the columns [7,11] by aligning the plots in the same column. This combination allows users to gain insights on the data on two different levels: first, an overview is granted via projections of randomly sampled points within the data domain through scatterplots. Second, local detail are presented via slices through the focal point in the HyperSlice. Additionally, an overlay of projections of the multi-dimensional gradient trajectories through the focal point enables them to easily identify improved input combinations.

2.2   Visualization Insight

One approach to assess the suitability of a visualization for a given purpose and a given set of data is to measure how many correct facts are derived from data. North [12] proposed measuring insight from a visualization as a key goal for visualization evaluation. Insight in his definition is not just recognition of data but *"complex*, involving all or large amounts of the given data in a synergistic way, not simply individual data values."[12] Furthermore, insight should be *deep* and built up over time raising new questions in the process. Insight is also *qualitative* in nature – not exact, but uncertain and "can have multiple levels of resolution". A key aspect of insight is that it is unpredictable and *unexpected* for the user. The user should not simply validate their prior expectations. The insights should thus be *relevant*, as in meaningful and more than mere data analytic findings. They should connect the underlying theory of the model with new relevant findings that have domain impact.

2.3   User Diversity

Whether a user is able to gain insight from any visualization is a question that not only depends on the quality of the visualization. Users themselves are different and many effects of user diversity must be considered when evaluating insights from a visualization. Three aspects of user diversity are considered in this paper, with no claim of being complete.

As one aspect of user diversity that could influence multi-dimensional understanding *fluid intelligence* comes to mind. A higher fluid intelligence is associated with being able to hold more information in working memory and change information more quickly [13].

The second aspect obviously relevant to this topic, is *spatial visualization ability*. Not all humans are able to manipulate spatial figures in their mind equally
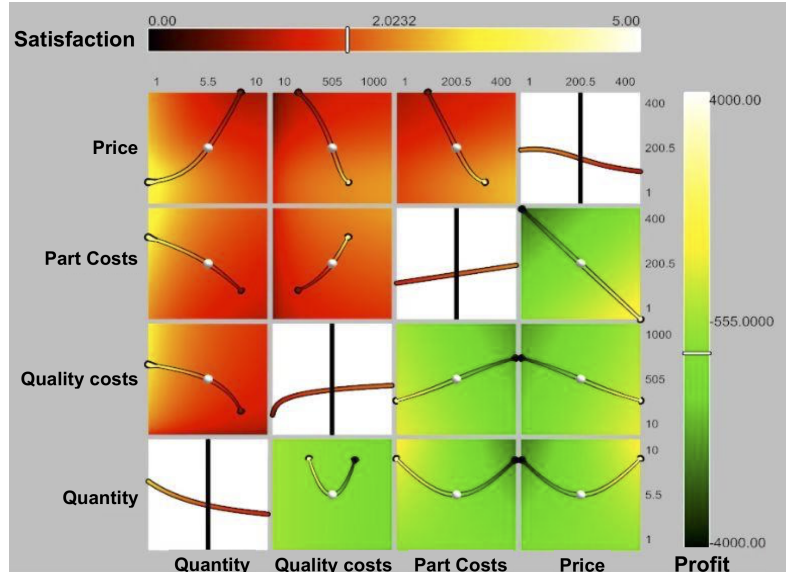
Fig. 3. HyperSlice of a data set with four input variables (columns and rows) and two output variables (upper and lower part). The graphs on the diagonal map one input variable to the same output variable as the upper part. By dragging the white focal point in any tile, all other tiles' slices are shifted in the 4-D hyperspace along the 2 of from the selected cell.

well. Older people typically perform worse than younger ones. Furthermore, men tend to outperform women [14], but the underlying hormone testosterone seems to have non-linear effects on spatial cognition. Low testosterone males, and high testosterone females perform best [14]. Beyond the general effect of spatial abilities on visualization evaluation, domain expertise may also play a role [15].

The last aspect of user diversity that might play a role in using a multi-dimensional visualization is *self-efficacy*. When using a computer system, the users' believe in being effective in doing so differs and influences how well they perform. One scale to measure this influence is the computer self-efficacy scale by Beier *KUT* [16].

3   Method

The purpose of this study was to investigate how users gain insight from multi-dimensional visualizations and the effect of user diversity. For this purpose, we used the visualization application memoSlice [7,11] in a user study with sixteen (n=16) participants.

Our approach is based on a no-benchmark task methodology by North [12]. However, letting participants explore the tool freely is not a fruitful approach, as

memoSlice provides two related views (scatterplot matrix and hyperslice matrix) that each consist of several plots. The amount of visual information is very high, as is the complexity of the underlying data-model (i.e., multi-dimensional dependencies, multiple outcome variables).

In order to get to interesting questions in a limited time-frame a tutorial is necessary. Therefore, we first asked participants to view an introductory video of memoSlice that explained the necessary features to complete six tasks afterwards. Completion of these tasks was not the actual target of the study, but the tasks were a mere guidance to structure the exploration of the tool. Participants were asked to think aloud and explain their insights during their exploration. After the exploration, we conducted a short interview, asking participants about what they liked in this tool and asked them to rate usability of the tool and how well their understanding of the visualizations was on a scale of 1–5 (bad–good).

3.1   Measuring User Diversity

Since multidimensional visualizations might be heavily influenced by spatial thinking capability we asked participants to complete three standardized tests. The *KUT* [16] measures the computer self-efficacy (CSE), a variable that explains how able users perceive themselves in using a computer. The *KAI-N* [17] measures the fluid intelligence in "bit". Fluid intelligence measures the capacity of working memory, by measuring memory span (how long can you retain random items in working memory) and processing speed (how fast can you take in new information). The Paper-Folding test (VZ-2) is used to assess an individuals spatio-cognitive abilities. Users are asked to predict the location of holes after puncturing a folded sheet of paper. This test is one of three tests to measure mental visualization skills [18,19] and derived from L.L. Thurstone's punched hole test (see Fig. 4). This test requires mental folding and thus also mental rotation and visualization of objects.
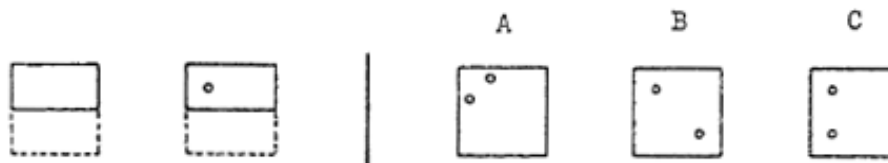

Fig. 4. Example of a paper-folding task. After folding a paper in half and punching a hole into the top left corner, two holes are in the unfolded paper. Participants now have to rate which option (A, B, C) is the layout of the holes in the unfolded paper. C would be correct in this case.

3.2 Measuring Insight

The model task in our study was an optimization task in a production setting. To have full control over the data and the insights that could be gained, we used a forged multi-dimensional function as underlying data model. As the two output variables we chose *profit* and *customer satisfaction.* The four determining input variables were *produced parts*, *quality costs*, *part costs*, and *price.* The probabilistic associations of the variables can be seen in the scatterplot matrix of memoSlice (see Fig. 2).

In order to analyze what actual insights users had, we analyzed all think-aloud user studies and recorded insights about both the tool and the underlying data model. We then measured the amount of correctly derived insights about the software and the data.
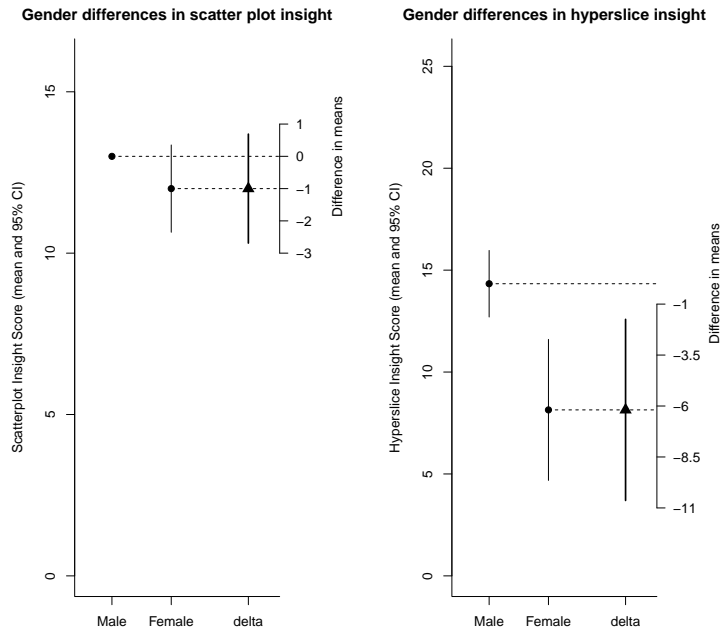

Fig. 5. No Gender differences in the insights on the scatter plot matrix (left plot). All males had all 13 insights. Comparison of means of insights found in the hyperslice visualization show gender differences (right plot). Error bars denote CIs.

3.3 Procedure

The sampling method was convenience sampling. We aimed to test a 50% gender ratio and asked only students from an engineering subject, approached directly

by us. The tests took between 30 to 90 minutes. Participants started with the tutorial video, before completing the task-driven think aloud experiment. Next, they completed the questionnaire survey and were then asked to comment on the experiment.

On the counted insights and user diversity criteria we used Pearson correlations (r) and comparison of means. We further report the test statistics with the level of significance (p) or the 95%-confidence intervals for point estimates. Confidence intervals (CIs) were calculated with the assumption of underlying normally distributed data, which can be assumed for the given standardized tests and are reported in square brackets.

4 Results

Our sample consisted of 7 female and 9 male students. Males scored higher in all diversity criteria (see Table 1). Although most differences can not be treated as such for the given sample size. Males showed a higher score in computer self efficacy ($M_{\male} = 5.32$, 95% CI $[5.06, 5.57]$, $M_{\female} = 4.59$, 95% CI $[3.60, 5.58]$). Since the CIs overlap to a large degree no difference in means must be assumed. Similarly males scored slightly higher in the KAI N scale ($M_{\male} = 154.2$, 95% CI $[126.5, 182.0]$, $M_{\female} = 120$, 95% CI $[95.8, 144.2]$), but again because of overlapping CIs no difference in means may be assumed. The difference in the paper folding test ($M_{\male} = 15.67$, 95% CI $[14.51, 16.82]$, $M_{\female} = 12.57$, 95% CI $[9.60, 15.54]$), where males also score higher, is also not statistically relevant.

The best possible score for insights two users received was 30 correct insights. The following insights were derived from the two individual views. The scatterplot matrix had 13 associated insights. The numbers in parentheses refer to the amount of different insights that were counted. For example, four fundamentally different high-dimensional associations were found in the scatterplot matrix (4).

- Rows are output variables (1)
- Columns are input variables (1)
- Identify association of two variables and their direction (positive, negative, or non-linear) (2)
- Interpretation of this association (2)
- Understand higher dimensional associations (e.g., price × profit × satisfaction) (4)
- Identify correct ranges of variables (2) and how to explore them (1).

HyperSlice yielded a maximum of 17 insights:

- Understand meaning of color in both parts of the matrix (2)
- Meaning of focal points and their manipulation (1)
- Meaning of gradient trajectories (3)
- Meaning of diagonal (maximal trajectory) (1)
- Meaning of upper half and lower half (2)
- Understand multi-dimensional associations (e.g., price × profit × part costs) (4)

- Identify correct ranges of variables (2) and how to explore them (1).
- Finding optimal production point (1).

	Scale Mean ♂	Scale Mean ♀
CSE	5.32	4.59
KAI N	154.20	120.00
Paper-Folding (VZ2)	15.67	12.57

Table 1. Means of computer self-efficacy (CSE), fluid intelligence (KAI N) and spatial capabilities (VZ2) by gender.

When looking at the amount of insights from both genders, we see no differences in the amount of insights derived from the scatterplot matrix (see Fig. 5). Yet it is interesting to note, that all male participants reported all thirteen insights. Differences between genders occurred only when looking at how many insights were derived from the HyperSlice visualization (see Fig. 5). Men on average had 14.3 insights (95% CI $[12.46, 16.25]$), while women only had 8.14 insights on average (95% CI $[3.82, 12.41]$).

Men reported a better subjective usability of the software ($M = 4.1$, 95% CI $[3.5, 4.7]$) than women ($M = 3.2$, 95% CI $[2.4, 4.2]$). They also report a higher subjective understanding of the underlying data ($M = 4.1$, 95% CI $[3.4, 4.8]$) than women ($M = 2.8$, 95% CI $[1.9, 3.8]$). But these differences are statistically not meaningful.

We also looked into how the different variables influence each other when used in correlation analysis (see Fig. 6). We found that both, computer self-efficacy and spacio-cognitive abilities influence scatterplot matrix insights, HyperSlice insights, and subjective understanding of the data model, respectively. HyperSlice insights also correlated strongly with subjective understanding. It is interesting to note, that only hyperslice insights correlate with subjective understanding and therefore no other associations of objective and subjective outcome measures exist in our study. However, this correlation is relatively strong ($r = .58$).

4.1 Qualitative Results

Since all think-aloud studies were audio-recorded and transcribed, more detailed results can be drawn from what participants actually said. For this purpose we categorized all mentions and counted the occurrence of various topics in those mentions. The most frequent topic was the positive mention of *associations*. Eleven participants mentioned that the tool helped them to identify how variables are associated (*"Oh, I can click the points here and they are highlighted in the other scatterplots, too. So I can see the relationship of two variables at the*
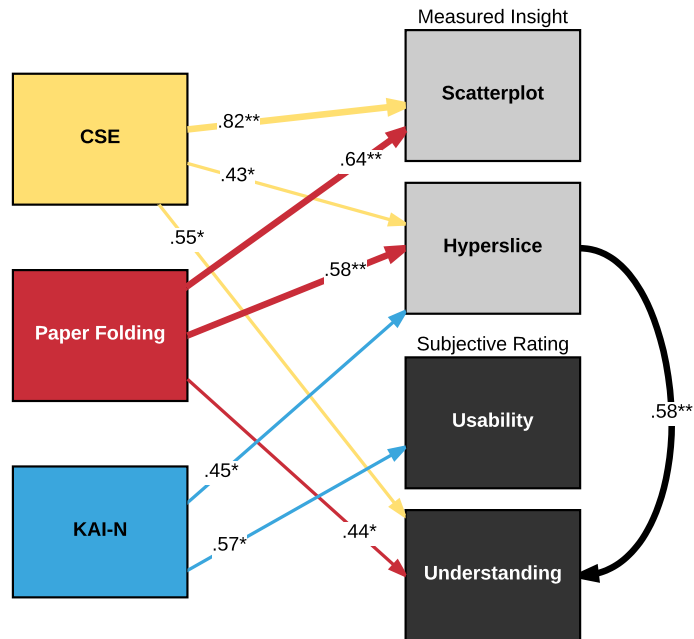
Fig. 6. Correlations of all variables. Numbers denote Pearson's r. Asterisks and line thickness indicate level of significance (thin* $p < .05$; thick** $p < .01$). Gender is dummy-coded, with male coded as 1. CSE is computer self-efficacy, and KAI-N is a standardized test for fluid intelligence.

top and three in the bottom."). The most often mentioned negative topic was the problem of *multi-dimensionality* (8 mentions, e.g.: *"The association of, for example, quality costs and costs per part and satisfaction. How do I set this? I can't. [frustrated] This is because of multi-dimensionality. This is hard to imagine. What does this point, that I see, mean in relation to the other two?"*). A close follow-up in negative mentions were the HyperSlice visualization itself (6 mentions) and the gradient trajectories (4 mentions). Only men commented on the latter and mostly negative (e.g.: *"What do these little worms mean? I have a suspicion [...], the tendency to neighboring points. But what neighbors? Hmm, no idea."*).

Most interestingly, all participants were able to complete the optimization task, independently of how well the users reported to understand the underlying data set. Thus, they were able to use the information visualization for one of its intended purposes.

5   Discussion

The data generated from this study focused on measuring insights, is only one part of the story of this case study. We found confirmatory evidence to expected outcomes (e.g. spatial ability predicts understanding of multi-dimensional dependencies). We also found the stereo-typical stronger results for male participants in the HyperSlice insights, based on the assumption of differing spatial abilities. However, one must note that there is no single spatial ability, but a multitude of different abilities. The mental folding of objects, which was used in this study, might be the most relevant for understanding hyperslices, as one has to "fold" dimensions onto each other. Mental rotation, which is important for mental folding, might also be highly relevant. Other Interestingly, almost all users were able to gain insights from the scatterplot matrix, and this also to a large extent. The scatterplot matrix can also be used to find three-dimensional associations by looking at two 2D-associations at the same time in the same row or column. Of course higher order conclusions are present in the data, yet they are not immediately visible to the naked eye.

Only the HyperSlice matrix naturally yields four-dimensional associations by combining two three-dimensional associations. This is were some of the participants struggled, and this is in line with the assumption of Halford et al. who claim no more than four variables can be assessed at the same time [3]. However, even higher order-associations are present in the data. No participant reported any of these. Yet, all were able to find satisfactory input settings in the optimization task.

In order to prevent interviewer effects, we asked a student to conduct the studies, without helping the participants. Interestingly, she reported to have increased insights herself after each trial. But, in the end the gradient trajectories remained a mystery to her.

Lastly, designing a study to quantify insights into complex data is also quite challenging. The sample we selected were a best case scenario of possible users, yet all were new to the visualization application. Domain knowledge effects (i.e. how to set-up quality costs for production) have been neglected in this study.

6   Conclusion

Overall, our case study yielded results that agreed with our hypotheses drawn from theoretical models. The relatively small sample size was caused by the large workload for each individual experiment. However, to our knowledge no research exists that analyzes how user-diversity factors influence insights in multi-dimensional visualizations. Therefore results even with limited statistical relevance are important. It seems to be necessary to develop methods that incorporate user diversity in the measurement process of insights, as without these aspects it will never be fully understood, who is really responsible for understanding multi-dimensional data: Human intelligence, good information visualization, or both? Probably both.

7   Acknowledgments

References

1. van Wijk, J.J., van Liere, R.: Hyperslice: Visualization of scalar functions of many variables. In: Proceedings of the 4th Conference on Visualization '93. VIS '93, Washington, DC, USA, IEEE Computer Society (1993) 119–125
2. Murata, M., Hashimoto, S.: Interactive environment for intuitive understanding of 4d object and space. In: Proceedings of the International Conference on Multimedia Modeling—MMM. (2000) 383–401
3. Halford, G.S., Baker, R., McCredden, J.E., Bain, J.D.: How many variables can humans process? Psychological science **16**(1) (2005) 70–76
4. Munzner, T.: Visualization Analysis and Design. CRC Press (2014)
5. Isenberg, T., Isenberg, P., Chen, J., Sedlmair, M., Möller, T.: A systematic review on the practice of evaluating visualization. IEEE Transactions on Visualization and Computer Graphics **19**(12) (2013) 2818–2827
6. Hoffman, P.E., Grinstein, G.G.: A survey of visualizations for high-dimensional data mining. Information visualization in data mining and knowledge discovery (2002) 47–82
7. Gebhardt, S., Pick, S., Voet, H., Utsch, J., al Khawli, T., Eppelt, U., Reinhard, R., Büscher, C., Hentschel, B., Kuhlen, T.W.: flapassist: How the integration of vr and visualization tools fosters the factory planning process. In: 2015 IEEE Virtual Reality (VR). (March 2015) 181–182
8. Hilda, J.J., Srimathi, C., Bonthu, B.: A review on the development of big data analytics and effective data visualization techniques in the context of massive and multidimensional data. Indian Journal of Science and Technology **9**(27) (2016)
9. Dos Santos, S., Brodlie, K.: Gaining understanding of multivariate and multidimensional data through visualization. Computers & Graphics **28**(3) (2004) 311–325
10. Al Khawli, T., Gebhardt, S., Eppelt, U., Hermanns, T., Kuhlen, T., Schulz, W.: An integrated approach for the knowledge discovery in computer simulation models with a multi-dimensional parameter space. In: International conference of numerical analysis and mathematics 2015 (ICNAAM 2015). Volume 1738., AIP Publishing (2016) 370003
11. Gebhardt, S., Al Khawli, T., Hentschel, B., Kuhlen, T., Schulz, W.: Hyperslice visualization of metamodels for manufacturing processes. In: IEEE Visualization Conference (VIS): Atlanta, GA, USA. Volume 13. (2013)
12. North, C.: Toward measuring visualization insight. Computer Graphics and Applications, IEEE **26**(3) (2006) 6–9
13. Lehrl, S., Fischer, B.: A basic information psychological parameter (bip) for the reconstruction of concepts of intelligence. European Journal of Personality **4**(4) (1990) 259–286
14. Gouchie, C., Kimura, D.: The relationship between testosterone levels and cognitive ability patterns. Psychoneuroendocrinology **16**(4) (1991) 323–334
15. Downing, R.E., Moore, J.L., Brown, S.W.: The effects and interaction of spatial visualization and domain expertise on information seeking. Computers in Human Behavior **21**(2) (2005) 195–209

16. Beier, G.: Kontrollüberzeugungen im Umgang mit Technik; self-efficacy in the use of technology (engl.). Report Psychologie **9** (1999) 684–693
17. Lehrl, S., Zipp, A., Schwarzfischer, C., Eissing, G.: Kurztest für allgemeine Intelligenz (KAI). (2016)
18. Carroll, J.B.: Human cognitive abilities: A survey of factor-analytic studies. Cambridge University Press (1993)
19. Eliot, J., Smith, I.: An International Directory of Spatial Tests. NFER-Nelson (1983)