Enhancing understandability of process models through cultural-dependent color adjustments

Author 1: **Tyge-F. Kummer**

Griffith University

170 Kessels Road, Nathan QLD 4111, Australia

Author 2: **Jan Recker**

Queensland University of Technology

2 George Street, P Block Level 8, Brisbane QLD 4000, Australia

Author 3: **Jan Mendling**

Wirtschaftsuniversität Wien

Welthandelsplatz 1, 1020 Vienna, Austria

Author biographies

Tyge-F. Kummer is a Lecturer at the Griffith Business School, Griffith University. His research interests include information systems, accounting, and business process management. He has published in journals including I&M, CAIS and BISE, as well as leading international conference proceedings (e.g., ICIS, ECIS and AMCIS).

Jan Recker is Full Professor of Information Systems and Retail Innovation at Queensland University of Technology. He has published in journals such as MISQ, JAIS, JIT, EJIS and others. He is Editor-in-Chief for CAIS and an Associate Editor for MISQ.

Jan Mendling is a Full Professor with the Institute for Information Business at Wirtschaftsuniversität Wien, Austria. His research interests include business process management and information systems. He has published more than 250 papers and is co-author of the textbooks „Fundamentals of BPM“ and „Wirtschaftsinformatik“.

Enhancing understandability of process models through cultural-dependent color adjustments

Abstract

National culture influences natural language communication. Yet, semi-formal or documented communication media such as process models have largely ignored these influences. In process models, secondary notation elements such as colors, however, provides designers with visual cues to potentially increase the efficiency and effectiveness without changing the semantics of the model itself. We propose that colors are a promising mechanism in tailoring process models to meet cultural preferences in order to enhance understandability. We test this assumption through an experiment with postgraduate students from a Confucian culture (China) and a Germanic Culture (Germany and Austria). Past research has shown that people understand such models better if important elements are highlighted through colors. We hypothesize that this general design principle only works if the applied color schemes match cultural preferences while mismatches can even diminish the level of understanding. Our results show that colors that are preferred in Asian cultures aid process model understandability of Confucian participants. In contrast, diverse effects occur if models with these colors are provided to members of a Germanic culture. Based on our findings, we derive implications for the culturally appropriate presentation of conceptual process models and we emphasize a need to construct modeling studies with cultural values and norms in mind.

Keywords: Process modeling, cross-cultural, quasi-experiment, design, secondary notation

1. Introduction

The documentation of business processes is of central importance for the analysis and design decisions in the contexts of information systems implementation, process redesign or organizational innovation initiatives [1-3]. Semi-formal graphical notations such as BPMN are used to facilitate communication among analysis and design team members by establishing a shared understanding of organizational business processes [4]. If effective, process models can be regarded as communication aids for analysts that support problem solving, and if understandable, can improve decision making [5].

It is known from marketing, accounting and decision support systems literature [6-8] that the mode representation influences problem solving and decision making abilities. Likewise, various factors are known to influence the accuracy and ease of understanding of process models by analysts [9, 10] when they try to make decisions about processes on the basis of such models. It is important to distinguish between those factors that are pre-defined by a modeling notation such as the symbols, their meaning and shape (*primary notation*) and other visual cues such as coloring, annotation, or positioning that can be used without impeding the meaning of the symbols [3, 11]. Such visual cues are referred to as *secondary notation* [12]. The usage of secondary notation is relevant as it provides the designer with a mechanism to potentially increase the efficiency and effectiveness of a model if it is tailored to meet characteristics of human cognition without changing the semantics of the model itself [13].

Experiments with process models have been largely conducted with samples from western cultures such as the United States [14], Australia [15] or Europe [10]. In fact, we are not aware of a single study that examined process model understanding with subjects from Asia or Africa. This sampling bias is a limitation of our knowledge to date. Research into technical visualization, marketing and user interface design indicates that customizations for different national cultures improve communication efficiency [16-18]. Yet, even though the importance of culture for process modeling has been highlighted [19], it has so far not been examined in how far tasks based on process models might be influenced.

In this paper, we focus on the effect of the secondary notation variation on process model understanding in two different cultures. Our central thesis is that we believe that coloring will have adverse effects for users from different cultural backgrounds. We study coloring for three reasons.

First, coloring is the most widely used secondary notation mechanism for process models [3, 20]. Second, prior studies have shown that both usage and perception of color varies across cultures [21]. Third, research findings from anthropology and intercultural communication suggest that cultural differences in color perception can be exploited to increase performance [e.g., 22, 23].

We designed and conducted an experiment in which we used two samples, postgraduate students from a Chinese background (referred to as Confucian) and from German-speaking countries (referred to as Germanic in accordance with House et al.'s [24] culture clusters), to examine our thesis. Our results show that color has *different effects* on understanding efficiency and perceived difficulty across these two cultural groups. These results suggest that process models should be customized according to cultural preferences in order to achieve the best level of understanding efficiency.

We proceed as follows. First, we discuss prior research on process modeling and relevant cultural research. Based on this discussion, we develop our thesis into a set of hypotheses. Next, we describe design of the experiment and the results. We discuss our findings and emphasize implications for research and practice. We conclude with a reflection of limitations and a summary of contributions.

2. Background

2.1 Color Usage in Process Models

Process models are a type of conceptual model, that is, they provide a typically graphical representation of some features of a real-world domain. They typically include graphical depictions of process steps, agents, actors, roles and artifacts that constitute a business process to analyze current or future operations [14].

Much research has been conducted to examine the key modeling design choices that determine how well individuals develop an understanding of the process and thus develop an ability to communicate and use process information for organizational and systems analysis tasks. This research focuses less on the question *what* is represented, but rather *how* a process is represented. One research stream has focused on *primary notation* elements [12] such as symbol sets and shapes [2, 15]. A second stream has been studying *secondary notation* elements such as layout, colors, and annotations [e.g., 3, 11, 25]. Our interest is in secondary notation elements as it contains *additional* visual cues that can support the understandability of the process model, but which are not part of the formal notation [12], and do not change the semantics of the grammar constructs in the model. Thus, they describe *free design choices* for the model creator.

The theory of effective visual notations provides a theoretical framework to analyze symbols and symbol sets according to desirable properties [13] such as semantic transparency and perceptual discriminability. These properties describe how clear (i.e., transparent) the meaning of a model element may be to a reader, and how easy it is to differentiate (i.e., discriminate) elements with different meanings from another. Coloring has been suggested as a mechanism to improve the perceptual discriminability of modeling elements. Because color highlighting can reduce visual search in a diagram, it can lead to a reduction in cognitive load, which translates into better understanding accuracy [3].

In summing up, much of the research on design choices in process modeling is explicitly or implicitly based on the assumptions that (a) model users are homogenous in terms of their national or cultural background or (b) cultural values and norms of users do not significantly influence their use of process models. We will present arguments from research on cultural influences on communication that suggest that these assumptions about effective process modeling guidelines may not necessarily hold.

2.2 Intercultural factors in technical visualization

Technical visualizations depict complex relationships between various objects in a simplified manner by using shapes, lines, arrows or other visual links in a structured manner [26, pp. 9]. Process models can be regarded as a specific instance of technical visualizations [27, p. 71]. The processing of technical visualizations is based on innate perceptual abilities that enable us to decide if we should pay attention to specific visual cues and how we should classify them in order to create a mental model [28]. Yet, knowledge about arrangements, relations and hierarchy are also partially learned and trained in recognition and dependence of the culture of the individual [29], suggesting that “culture matters” also to how users interpret process models.

Culture generally describes a set of shared norms, values and orientation patterns that influence behavior of individuals within a specific group [30]. Such differences can be observed between individuals from different countries [24, 31]. Cultural differences have been shown to be a key challenge in global organizations and projects [32, 33], technology use [34, 35], or the perception of values relevant to systems design practices [36]. In turn, it appears likely that cultural differences will also manifest in projects involving process models.

There is some evidence from research on forms of technical visualization that supports this assertion. If a domain expert creates a technical visualization to be read by experts from the same cultural background, it is likely that they have a similar mindset and can understand the message easily [37]. Many technical visualizations, however, rely on conventions developed in Western cultures such as tree charts or organizational diagrams [27, p. 87]. This can be problematic when they are applied in a different cultural setting. For example, studies found vast differences between the choice of diagrammatic illustrations between China and the United States [17] or in the use of illustrations [38].

In sum, the literature indicates that communications with technical visualizations such as process models may be perceived and interpreted differently by individuals from different cultures.

2.3 Color-dependent differences across cultures

Color differences between cultures have been studied in cultural anthropology, linguistics, psychology, user interface or website design, e-commerce and marketing. To make sense of this literature and to guide our research design, we conducted a literature review. Appendix A summarizes this review and highlights key findings and the implications on our research project.

Several key findings emerge. First, the perception of colors between Asian and western countries varies. Purple, for instance, is a color that is perceived as positive in Asian countries such as China, Japan and South Korea where it represents love and expensiveness. In western countries, the color purple is often associated with inexpensiveness [39, 40]. While some colors such as white seem to be neutral between cultures, in particular loud and bright colors are preferred in Asian countries [18, 40]. Findings in website and user interface design research also suggest that color differences in website design exist between cultures [41] and websites that apply cultural adjusted color schemes are preferred in specific cultures. The findings specifically indicate that Asian cultures prefer many bright colors while western countries prefer few colors with minimal intensity [42].

Second, categorical color perception varies in accordance with language variations. Most existing color categories are based on western art and culture and ignore cultural diversity in cognition and language [22]. Yet, innate visual capabilities to distinguish color categories are shaped based on linguistic learning [23, 43]. For instance, comparative experiments investigating color categories of 2 to 4 years old children did not find cross-cultural differences [44], suggesting that physiological differences in the color perception can be ruled out. By contrast, Athanasopoulos et al. [45] find that bilingual Japanese that speak English frequently were not able to distinguish shades of blue as well as the monolingual norm. This suggests that culture-dependent learning strategies appear to at least influence color perception.

Third, research on differences between cultural groups [e.g., 24, 30] identifies three properties that relate to differences in color perception.

a) The cultural dimension *individualism* [30] measures the degree of personal freedom and individual rights in a society (e.g., freedom of opinion and press). This affects design choices. Individuals from individualistic cultures such as Germany tend to prefer monotonously colored interfaces while individuals from collectivistic cultures such as China prefer colorful designs [16, 46]

b) The cultural dimension *uncertainty avoidance* manifests in preferences for strict rules, a high degree of formality and punctuality, and a low tolerance for ambiguity [30]. Individuals from cultures with a high degree of uncertainty avoidance such as Germany tend to prefer redundant cues (color, typography, sound, etc.) to structure content [47]. In contrast, individuals from cultures with a low degree of uncertainty avoidance such as China share a higher tolerance for ambiguity that is expressed through less written or unwritten rules to maintain predictability. Such cultures prefer a flexible coding using a variety of colors and sound to maximize information [46].

c) *High and low context communication* may also induce differences in color perception [42]. In a high-context communication culture, messages tend to include only a small amount of words and many things are left unsaid [48]. As a result, in Scandinavian countries and Germany functionality and aesthetic experience is favored leading to consistent layout and color schemes [42]. In contrast, in a low-context culture communication tends to be more direct; and many words are used, leaving little information to be deduced from the context [48]. Therefore, it is argued that Asian cultures prefer diverse layout and color schemes using many bright colors, fonts, and shapes [42].

Importantly, such culturally appropriate design preferences are not only relevant for stimulating a particular sentiment, but also for performance. For example, users with a culturally adjusted interface performed tasks faster and more accurately in the study by Reinecke and Bernstein [46].

We believe that these findings suggest that culture can be a prominent factor in explaining differences in process model understanding in dependence on the usage of colors within the models and the cultural background of the model user.

3. Hypothesis development

Our central thesis in this paper is that differences in cultural values about the use of color will induce differences in the *perceived difficulty, understanding* and *efficiency* of process models as visual communication artifacts. This is because the effective usage of color in intercultural communication can work as a visual cue to assign significance to certain stimuli and guides the attention of the viewer [28]. Yet, inconsistent usage of colors can also undermine communication [49], which suggests that coloring in some contexts may also have adverse effects. The purpose of our hypothesis development is, therefore, to detail our expectations about the cultural contexts in which coloring induce positive – or negative – effects on process models understanding.

In developing our hypotheses, we draw on cognitive load theory to view process models as explanative multimedia messages from which individuals can develop understanding [2, 14, 50-52]. In this theory, process model viewers actively integrate information content in the process model that is presented in a particular presentation format with their own previous experience and existing mental models. The theory differentiates three forms of cognitive load: *Intrinsic* load refers to inherent level of difficulty, *extraneous* load to the presentation in which information is presented and *germane* load to the processing and construction of schemas [53]. The ability of viewers to integrate new knowledge with their existing knowledge is limited by their cognitive capacity and influenced by cognitive load [54], which in turn explains how much understanding of the process being modeled will be achieved. [55], for instance, discovered that the ability to learn a foreign language is reduced for English-speaking college students if they received additional visual annotations. This effect did not occur when the annotations were presented verbally.

In general, design choices such as the use of coloring in process models relate to the *presentation* *format*. Cultural values and norms are *user characteristics* that will influence how the information content in the chosen *presentation format* allows for *knowledge construction*, which is typically operationalized in three ways [9, 56]: the **difficulty of developing understanding**, the **accuracy of the understanding** that is being developed, and the resources invested in developing this understanding (i.e., **task efficiency**). We structure our hypotheses accordingly.

3.1 Influence on Perceived Difficulty

First, we expect that participants from the Confucian culture with a low degree of Individualism and Uncertainty Avoidance as well as a high Context Communication score will prefer colorful designs, and will therefore experience a lower cognitive load if a process model contains several different bright colors with high contrast. For Confucians, colors will support visual channel processing and thereby reduce the cognitive load for the Confucian Group. In turn, this group will perceive the knowledge construction to be easier and colorful models to be less difficult to understand. On the contrary, users from countries with a high degree of Individualism and Uncertainty Avoidance as well as a Low Context Communication score, like Germany and Austria, will likely prefer fewer and more monotonous colors. To that end, for Germanics, the use of colors in conceptual process models will block the visual channel of working memory processes and lead to additional cognitive load in the content presentation format. This will make the knowledge construction more challenging resulting in an increased perceived difficulty. We state:

H1a. Members of the Confucian culture cluster will perceive models that contain coloring as less difficult to understand than models that do not contain coloring.

H1b. Members of the Germanic culture cluster will perceive models that contain coloring as more difficult to understand than models that do not contain coloring.

H1c. Members of the Confucian culture cluster will perceive models that use coloring as less difficult than members of the Germanic culture cluster.

3.2 Influence on Understanding Accuracy

In countries such as China with a low degree of Individualism and Uncertainty Avoidance, users will prefer bright and colorful designs and it can be expected that these colors will reduce extraneous load. For Confucians, the lower cognitive load increases understanding accuracy of the learning material and thus improves model understanding for Confucian participants. In contrast, users from Germanic countries will be more likely to prefer monotonous colors. For Germanics, the colors will cause additional extraneous load leading to understanding difficulties. Consequently, participants will have a lower understanding score if a process model contains several different bright colors with high contrast. We state:

H2a. Members of the Confucian culture cluster will achieve higher levels of understanding accuracy if the model contains coloring than if the model does not contain coloring.

H2b. Members of the Germanic culture cluster will achieve higher levels of understanding accuracy if the model does not contain coloring than if the model contains coloring.

H2c. Members of the Confucian culture cluster will achieve higher levels of understanding accuracy for colored models than members of the Germanic culture cluster.

3.3 Influence on Understanding Task Efficiency

As proposed in hypothesis H2, we expect that the use of colors in process models will reduce cognitive load for the Confucian Group and add cognitive load in the Germanic Group. In consequence, we assume that participants from the Confucian Group will require a shorter amount of time to process the information in a process model while the Germanic Group will require a larger amount of time to complete the model comprehension tasks. In turn, the differences expected in understanding accuracy should also manifest in differences in task efficiency, that is the amount of understanding generated per unit of time [57, 58]. Formally, we state:

H3a. Members of the Confucian culture cluster will achieve higher levels of task efficiency in tasks with models that contain coloring than for tasks with models that do not contain coloring.

H3b. Members of the Germanic culture cluster will achieve higher levels of task efficiency in tasks with models that do not contain coloring than for tasks with models that contain coloring.

H3c. Members of the Confucian culture cluster will achieve higher task efficiency in tasks with models that contain coloring than members of the Germanic culture cluster.

4. Research Method

We used a laboratory quasi-experiment to test our hypotheses because (a) experiments are common in process modeling research, thus allowing comparison of results [59], and (b) there are few studies on cultural influences in model comprehension and thus our primary objective was to maximize internal validity.

4.1 Design and Measures

A (2x2)x2 mixed multi-trial quasi-experimental design was chosen with one between-groups factor (coloring) and one within-group factor (culture) in two trials. The defining characteristic of quasi-experiments (also called difference studies) is that the distinguishing factor – in our study the cultural background – is fixed and cannot be assigned randomly [60]. Quasi-experiments are a common and widely accepted research approach that has been applied successfully to explore for example differences in relation to gender, age, education, ethnicity, or socio economical background (e.g., see [61, 62]). A methodological weakness of this approach is that the internal validity suffers if confounding variables exist that could cause differences between the groups [63]. However, the bias can be controlled by determining the confounding variables and including them in the model as covariates [60, 64]. Quasi-experiments are suitable in our study as the fixed factor does not allow a random group assignment and the existing body of research allows controlling for confounding variables sufficiently.

The choice of two trials was motivated by the desire to boost the external validity of the study and to examine whether our experiment yielded consistent results across different models and tasks. Participants were randomly assigned to two groups (A and B). The between-groups factor, coloring, had two levels. To manipulate this factor, we created two sets of process models, one for each trial that contained an identical representation of a domain but varied in their use of coloring. The first variant contained only two types of white (one for the background and one for elements), while the other variant perused different bright colors with high contrast [3]. To simplify wording we will refer to the variants as color and no (bright) colors. Each group received a model with colors and one without colors (Table 1).

The within-group factor, culture, had two levels (Germanic and Confucian). In deciding how to measure culture, we decided not to measure cultural dimensions of individual participants. We had four main reasons:

1) We do not perform a regression based analysis that explores a direct or moderating effect of

cultural values on a particular behavior [65].

2) A group assignment via cultural index scores would not be aligned to our hypotheses as it would change the research design from a cross-country study to a comparison of high and low cultural index scores across cultures.
3) It is not possible to explain every cross-cultural difference with (measurable) cultural dimensions index scores. National culture has evolved over thousands of years and it is an extreme simplification to break cultural differences down to a set of cultural dimension index scores. Additionally, not every cultural dimension has a validated measure. This includes low and high context communication [48].
4) Hofstede's cultural dimensions have not been developed as an individual measure and are known for psychometrically disappointing results at the individual level [66, 67].

As a result we decided to use cultural dimensions only at a national level to justify hypothesis development and country selection. Our group assignment procedure was thus based on cultural identity, which assumes that individuals are socialized by their environment and adopt the cultural identity of the country of birth [68]. For the Confucian group we decided to recruit native participants from China as the GLOBE study suggests that individuals from this country represent the Confucian culture cluster. For the Germanic group we recruited native participants from Germany and Austria who, according to the GLOBE study, represent the Germanic culture cluster [24]. We selected these nations for the manipulations of the culture variable, because they show at a national level high differences regarding Hofstede's [30] cultural dimensions of individualism (China 20; Austria 55; Germany 67), uncertainty avoidance (China 30; Austria 70; Germany 65) and high and low context communication (China high context communication; Austria, Germany low context communication, see [69]). Because the cultural identify of participants could be biased by spending prolonged periods of time in different countries [16, 46], we asked participants to list the countries in which they had lived and the duration of their stay. Participants who lived for more than 12 months in a country other than Germany, Austria or China were excluded.

The choice of a quasi-experiment with an assignment procedure based on the cultural identity meant we needed to place specific emphasis on the identification of relevant control variables to be able to isolate culture as a responsible determinant for the effects we were studying [69, p. 286 ff.]. Based on our understanding of the process model experiment literature, we thus measured *familiarity* with the BPMN grammar [70] and *knowledge* of process modeling concepts [9] as control variables. In this way, we can show that a potential difference between the cultures would not result from differences in BPMN familiarity or modeling knowledge, which are known influence factors [e.g., 9, 52]. We also included measures for demographic variables such as gender and age.

We defined three dependent measures. *Perceived difficulty* refers to the ease of understanding the process models provided [71]. In order to guarantee structural equivalence, we conducted an exploratory factor analysis for the construct of perceived difficulty. The results support that the patterns of correlation were cross-culturally similar.

Second, we measured *understanding accuracy* as the task score for eight comprehension questions per model, similar to other studies [9, 52]. The comprehension questions were developed as follows: First, one researcher created a battery of comprehension questions following examples of existing measures. Second, two process modeling experts with a strong teaching and research focus in business process management completed the comprehension tasks to review each question and to develop benchmark answers. To maintain inter-coder reliability, both experts first individually responded and then met to defend and discuss their coding to generate a final, consensually agreed coding result. After the individual response coding, we calculated a Kappa statistic of 0.86, which suggests excellent inter-coder reliability. After discussion of questions as well as benchmark answers, the agreement was 100 percent.

Third, we measured *task efficiency* as the comprehension task score divided by the task completion time for each model [9].

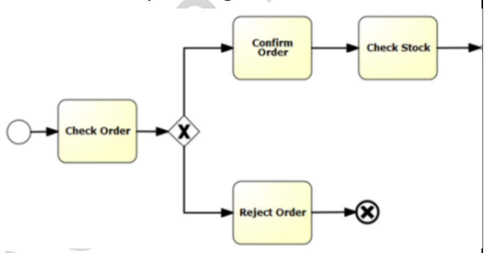
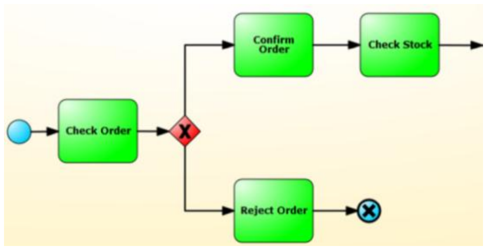
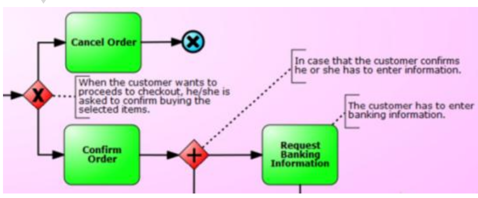
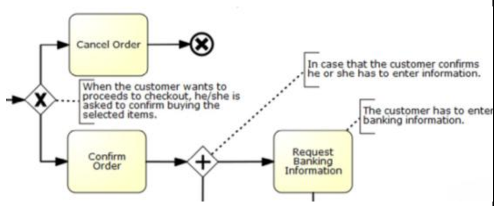
4.2 Materials and Procedures

We used two sets of materials in our experiment. The first set collected demographic information from participants including measures for country of origin and relevant working experience, *familiarity* with the process modeling grammar BPMN [70], and a process modeling *knowledge* quiz [9]. The second set of materials comprised two process model diagrams, one for each trial, alongside with brief textual descriptions of the relevant case. The first showed a model of a manufacturing process, and the second a model of an online shopping process. The process models were created using the BPMN grammar [72] because of its position as the industry standard for process modeling. In order to neutralize model complexity as a potentially confounding variable [15], we kept model size, connection and complex behavior within a narrow, normal range in both trials.

In designing the models for each trial, we decided to introduce a *crossover variation* to minimize results skew due to experimental fatigue, boredom or learning effects that could stem from order effects in multi-trial experimental designs [73]. We changed the order of the trials between group A and B (see Table 1). Additionally, we decided to introduce a second visual design mechanism, textual annotations, into trial 2 in order to explore interaction effects. Textual annotations are a secondary notation element that clarifies instructions about the modeled process in the form of annotations that do not change the semantics of the constructs. In introducing this variation, we needed to ensure that the models remained informationally equivalent to the models in trial 1 because otherwise result difference may stem from participants being provided extra information. To achieve this, we provided an additional case description to the models in trial 1 that explained aspects of process execution. In trial 2, these brief descriptions of relevant business rules were provided as textual annotations in the models. This ensured informational equivalence; however, the materials in trial 2 provided a computational advantage through integration of visual and textual information in the model. Therefore, we expected the second model to be easier to understand as less extraneous cognitive processing is required to process the information than when words and graphics need to visually searched separately [50]. The experiment design decision allows us to examine the influence of culture and coloring on process model understanding under different conditions, that is, a scenario with more cognitive load due to additional visual cues in the process model. Therefore, this multi-trial design allowed to strengthen the internal validity of the design and to more robustly examine the conclusion validity of our results.

The colors were selected according to Asian preferences (see Appendix A). This includes in particular the use of bright and loud background colors such as purple. The different model elements are colored in a consistent manner (e.g., gateways are always colored red, events blue). The colors were assigned to increase the contrast in the diagram but do not provide additional information. For instance, red was deliberately not used as a warning color (see Table 1). All materials are provided in Appendix B.

For each model, participants completed eight process model understanding questions and answered a scale on perceived model difficulty. The model understanding questions were formulated as multiple choice questions to compute *understanding accuracy* as a score. These questions queried participants surface understanding of the business process domain depicted in the model. This type of measure has a long track record in studies of domain understanding generated from viewing conceptual models [15, 52, 56]. To measure task efficiency, we used the model understanding score and divided it by the recorded time taken to complete the understanding questions. *Perceived difficulty* was measured using a six-item Likert scale adopted from Burton-Jones and Meso's [56] ease of understanding scale that captures perceived difficulty of understanding the diagram as well as the domain that is depicted in the diagram.

<table>
<tr><th colspan="3">Table 1: Materials used in experiment</th></tr>
<tr><td>Within-group factor</td><td colspan="2">Germanic / Confucian</td></tr>
<tr><td rowspan="3">Between groups-factor and Trial order</td><td>Group A – trial 1 first</td><td>Group B – trial 2 first</td></tr>
<tr><td>Trial 1 – Manufacturing model

</td><td>Trial 1 – Manufacturing model <u>in color</u>

</td></tr>
<tr><td>Trial 2 – Online shopping annotated model <u>in color</u>

</td><td>Trial 2 – Online shopping annotated model

</td></tr>
</table>

An important aspect of cross cultural research is the equivalence of the constructs in the different cultures [74]. Thus, we followed a back translation procedure [75]. First, we created a German version, which was translated professionally in Chinese ideographs. Next, a second bilingual expert re-translated the questionnaire into German. The back translation covered most of the original wording. Existing differences were discussed and led to minor adjustments in both material variants. Researchers from both cultures with mono and bilingual skills participated in the creation and analysis of the experiment, to ensure that all questions as well as answers were understood appropriately [75].

4.3 Participants

Participants were postgraduate business students who had completed a university course covering process modeling skills and notations. In order to exclude biases due to the teaching style, the course in Germany and China was provided by the same person. The experiment was conducted between October 2011 and February 2012 at three leading universities. 25 students participated in Germany, 40 in Austria and 67 in China. We excluded 5 responses (3 in Germany, 2 in China) due to incomplete answers or for not providing sufficient answers in an adequate time. In every experiment, an incentive with a value of about 40-45 Euro was provided for the student with the highest model understanding score, to incentivize good performance. To gauge whether the sample size was sufficient, we estimated desired statistical power using G*Power 3.0 [76]. Given a sample size of $n = 127$, and expecting a moderate effect size (0.15) and an α error probability of 0.05, our experiment design with 4 groups (between- and within-factors) and 17 response variables in total achieves a statistical power greater than 0.99, which is well above the suggested threshold of 0.8.

5. Results

We started by examining descriptive statistics for validity of our data and conducting relevant manipulation checks. Relevant results are provided in Appendix C. The checks indicated that our hypothesis testing should include BPMN familiarity and process modeling knowledge as covariates, and that experiment fatigue bias was not present. We also examined our results post-hoc for gender differences, finding no evidence for such (Appendix D).

To evaluate our hypotheses, we ran two multiple analyses of covariance (MANCOVA), one for each trial. Independent variables are the between-group factor color and the within-group factor culture. We included as covariates the average total factor score for BPMN familiarity (trial 1 = 3.37; trial 2 = 3.36) and the process modeling knowledge score (trial 1 = 4.39; trial 2 = 4.45). Dependent measures were the understanding accuracy score, the task efficiency score and the average total factor score for perceived difficulty. The tests were computed using IBM SPSS 21.0. The data met the assumption of equal variances in the dependent measures across groups. Levene's test was insignificant for all cases, indicating that the data met this assumption. Descriptive statistics from the two MANCOVA analyses are reported in Table 2. Table 3 provides test results. In summary, we found that across both trials culture and color were significantly related to the dependent variables we examined. We also found that process modeling knowledge was a significant covariate in explaining understanding accuracy, whereas familiarity with BPMN was not. A significant interaction effect between the two factors culture and color occurred in relation to task efficiency and perceived difficulty. Therefore, we conducted additional simple effect analyses to break down the effects that caused the interaction [77]. Effect sizes (partial eta squared) were satisfying with four small, four medium and two large effects based on Cohen's (1988) guidelines (Table 3).

Perceived Difficulty

In Hypothesis 1 we speculated that colors in process models will affect the perceived difficulty. Table 3 shows an interaction effect between culture and color in trial 1 [$F(1, 115) = 9.22$, $p < 0.001$] as well as trial 2 [$F(1, 115) = 4.90$, $p < 0.05$]. As a result, the interpretation of main effects becomes meaningless as the interaction effects overlay main effects and a simple effect analysis is required to identify the origins of the interaction effect [77, 78]. The results are summarized in Table 4. The data in Table 4 shows that in the Confucian cluster the process models with colors were perceived as significantly less difficult in trial 1 (mean difference = -0.50, $p<0.05$) as well as trial 2 (mean difference = -0.48, $p<0.05$). Figure 1 shows the estimated marginal means of perceived difficulty for both trials and illustrates that the colored models are perceived as less difficult in the Confucian group. Therefore, H1a is supported.

In the Germanic group the effects are directed in the opposite direction indicating that colored models are perceived as more difficult. A significant effect emerges only in trial 1 (mean difference = 0.50, $p<0.05$). In trial 2 the effect is not significant (mean difference = 0.24, $p>0.05$). A possible explanation could be that the additional textual annotation cues provided eased the negative effect of color on the perceived difficulty. Given the correct directionality of the effects and the significant results in trial 1, we classify H1b as partially supported.

Table 2: Descriptive Statistics for model understanding task scores							
Trial 1: Manufacturing Case							
Cultural Group		Perceived Difficulty score		Understanding accuracy score		Task Efficiency score	
	Treatment	Mean	Std. D.	Mean	Std. D.	Mean	Std. D.
Germanic	No Color	3.42	1.02	4.53	1.14	1.18	0.57
	Color	3.91	1.00	5.00	1.27	1.67	0.70
	Average	3.67	1.03	4.77	1.22	1.43	0.68
	Difference	-0.49		-0.47		-0.48	
Confucian	No Color	2.59	0.92	4.10	1.60	1.25	0.80
	Color	2.11	0.65	4.31	1.40	1.91	0.87
	Average	2.34	0.82	4.21	1.49	1.59	0.90
	Difference	0.49		-0.21		-0.66	
Total	No Color	3.01	1.05	4.32	1.40	1.22	0.69
	Color	3.01	1.23	4.66	1.37	1.79	0.79
	Average	3.01	1.14	4.49	1.39	1.51	0.79
	Difference	0.00		-0.34		-0.57	
Difference		1.33		0.56		-0.15	
Trial 2: Online Shopping Case							
Cultural Group		Perceived Difficulty score		Understanding accuracy score		Task Efficiency score	
	Treatment	Mean	Std. D.	Mean	Std. D.	Mean	Std. D.
Germanic	No Color	2.81	0.85	6.45	0.96	2.13	1.21
	Color	3.02	0.94	6.11	1.10	2.24	0.62
	Average	2.91	0.89	6.29	1.03	2.18	0.97
	Difference	-0.21		0.34		-0.12	
Confucian	No Color	2.67	0.97	5.72	1.49	1.98	1.01
	Color	2.13	0.75	6.13	1.17	3.21	2.12
	Average	2.41	0.90	5.92	1.35	2.57	1.74
	Difference	0.54		-0.41		-1.23	
Total	No Color	2.74	0.91	6.08	1.30	2.05	1.11
	Color	2.56	0.95	6.12	1.13	2.74	1.64
	Average	2.65	0.93	6.10	1.21	2.38	1.43
	Difference	0.18		-0.04		-0.69	
Difference		0.17		0.21		0.13	

Table 3: MANCOVA results for model understanding task scores

Effect type	Factor	Statistic	Trial 1			Trial 2		
			Perceived Difficulty	Understanding Accuracy Score	Task Efficiency	Perceived Difficulty	Understanding Accuracy Score	Task Efficiency
Covariates	Process Modelling Knowledge	F(1,115)	0.16	12.11***	7.95**	1.69	3.17^{+}	0.61
		Partial Eta Squared	0.00	0.09	0.06	0.01	0.03	0.01
	BPMN Familiarity	F(1,115)	0.71	0.20	0.12	0.26	0.33	0.43
		Partial Eta Squared	0.01	0.00	0.00	0.00	0.00	0.00
Main Effects	Culture	F(1,115)	49.61***	0.97	4.77*	12.44***	0.92	3.80*
		Partial Eta Squared	0.30	0.01	0.04	0.10	0.01	0.03
	Color	F(1,115)	0.00	3.68^{+}	22.68***	0.58	0.02	6.27**
		Partial Eta Squared	0.00	0.03	0.16	0.00	0.00	0.05
Interactions	Culture x Color	F(1,115)	9.22***	0.05	0.79	4.90*	2.26	4.77*
		Partial Eta Squared	0.07	0.00	0.01	0.04	0.02	0.04

*** p < 0.001, ** p < 0.01, * p < 0.05, $^{+}$ p < 0.10

Table 4. Simple effects analysis for perceived difficulty				
Variable	**Culture**	**Treatment**	**Mean Difference (I-J)**	**St. Error**
Perceived Difficulty – Trial 1	Confucian	Color (I), No color (J)	-0.50*	0.23
	Germanic	Color (I), No color (J)	0.50*	0.23
Perceived Difficulty – Trial 2 (with additional cues)	Confucian	Color (I), No color (J)	-0.48*	0.23
	Germanic	Color (I), No color (J)	0.24	0.23
Variable	**Color**	**Treatment**	**Mean Difference (I-J)**	**St. Error**
Perceived Difficulty – Trial 1	Color	Confucian (I), Germanic (J)	-1.79***	0.24
	No Color	Confucian (I), Germanic (J)	-0.79**	0.25
Perceived Difficulty – Trial 2 (with additional cues)	Color	Confucian (I), Germanic (J)	-1.00***	0.25
	No Color	Confucian (I), Germanic (J)	-0.28	0.24
*** p < 0.001, ** p < 0.01, * p < 0.05				

Examining the culture factor, Table 4 shows that the culture treatment leads to significant differences in trial 1 for the colored model (mean difference = -1.79, p<0.001) and the non-colored model (mean difference = -0.79, p<0.01). In trial 2 the significant differences between the cultures emerged only for the model with colors (mean difference = -1.00, p<0.001). Again the occurrence of additional cues (text annotations) in trial 2 could be a possible explanation.

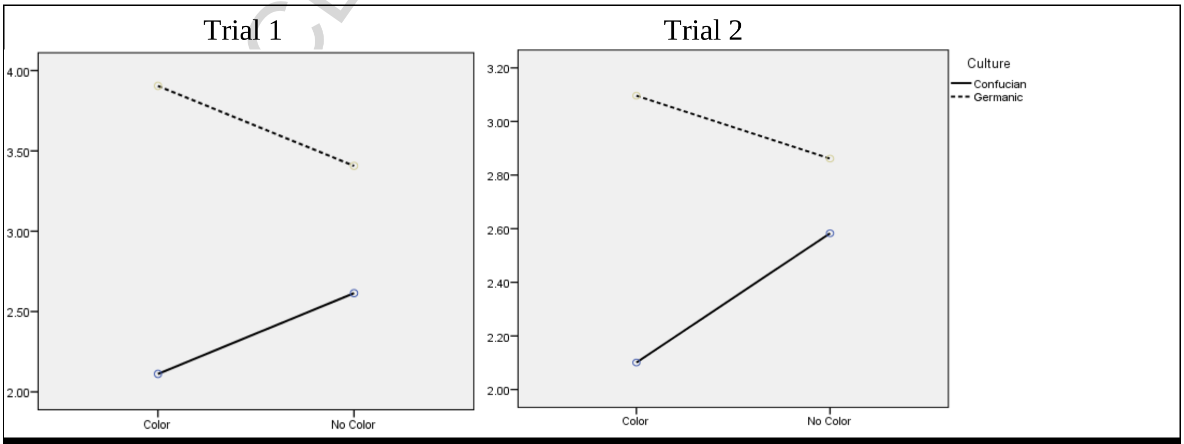


Figure 1. Estimated Marginal Means of Perceived Difficulty across Trials

In H1c we speculated that the Confucian groups will perceive models that use coloring as less difficult than the Germanic Groups. The results show a relation between culture and color. In general process models are perceived as more difficult in the Germanic culture than in the Confucian culture.

Additionally, process models without colors are perceived as less difficult in the Germanic group and more difficult in the Confucian group. The use of colors increases the difference between the cultures. Consequently, we classify H1c as supported.

Understanding accuracy

We speculated that members of the Confucian cluster will have higher understanding accuracy in models using colors (H2a) while members of the Germanic cluster achieve lower understanding accuracy in models using colors (H2b). To explore these results, we again plotted the estimated marginal means (Figure 2). In the Confucian group the participants achieved a greater understanding accuracy in relation to colored models. In contrast, the results in the Germanic group do not indicate a clear pattern. In trial 1 the Germanic participants achieved a higher score in the colored model while they achieved a lower score for the model with colors in trial 2. The results in Table 5 show that the main factor color affected the understanding accuracy score in trial 1 at $F(1,115) = 3.68$, $p<0.10$ and in trail 2 with at $F(1,115) = 0.02$. The simple effects analysis indicates no significant effects as well. In turn, H2a and H2b are rejected.

Table 5. Simple effects analysis for understanding accuracy

Variable	Culture	Treatment	Mean Difference (I-J)	St. Error
Understanding Accuracy – Trial 1	Confucian	Color (I), No color (J)	0.40	0.34
	Germanic	Color (I), No color (J)	0.51	0.33
Understanding Accuracy – Trial 2 (with additional cues)	Confucian	Color (I), No color (J)	0.30	0.31
	Germanic	Color (I), No color (J)	-0.36	0.31
Variable	**Color**	**Treatment**	**Mean Difference (I-J)**	**St. Error**
Understanding Accuracy – Trial 1	Color	Confucian (I), Germanic (J)	-0.31	0.35
	No Color	Confucian (I), Germanic (J)	0.21	0.36
Understanding Accuracy – Trial 2 (with additional cues)	Color	Confucian (I), Germanic (J)	0.09	0.33
	No Color	Confucian (I), Germanic (J)	-0.56	0.33
*** p < 0.001, ** p < 0.01, * p < 0.05				

The visualizations in Figure 2 show that the Confucian group achieved a lower understanding accuracy scores for colored and non-colored models in trial 1. Members of the Germanic cluster, however, achieved a higher score only in relation to the model without colors in trial 2, while the

Confucian group performed better in relation to the model with colors. Table 3 shows that culture is not a significant factor and also the simple effect analysis does not support our hypothesis (p>0.05 for all mean differences). In turn, H2c is rejected.

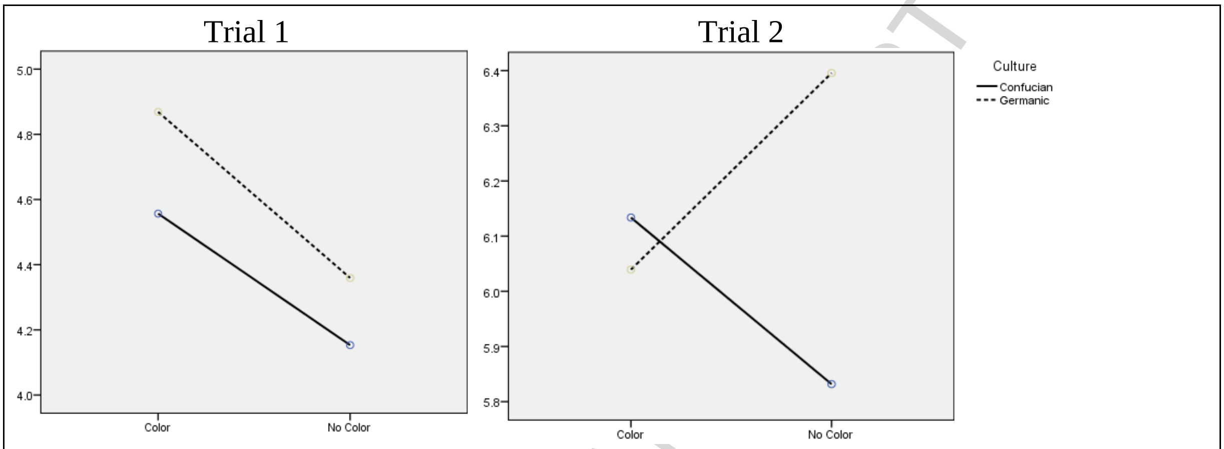


Figure 2. Estimated marginal means of understanding accuracy

Understanding efficiency

We expected that understanding task efficiency is higher in the Confucian groups when models are colored (H3a). Again, a contrary effect was assumed in Germanic group (H3b). Table 3 shows that the main effect color is significant in trial 1 [F(1,115), 22.68, p<0.001]. However, in trial 2 an interaction effect occurs. Therefore, we performed a simple effect analysis that is summarized in Table 6.

Table 6. Simple effects analysis for understanding efficiency (Trial 2)

Variable	Culture	Treatment	Mean Difference (I-J)	St. Error
Understanding Efficiency – Trial 2 (with additional cues)	Confucian	Color (I), No color (J)	1.18***	0.35
	Germanic	Color (I), No color (J)	0.08	0.36
Variable	**Color**	**Treatment**	**Mean Difference (I-J)**	**St. Error**
Understanding Efficiency – Trial 2 (with additional cues)	Color	Confucian (I), Germanic (J)	-1.09**	0.38
	No Color	Confucian (I), Germanic (J)	-0.00	0.37

The difference between the model with colors and the model without colors are significant in the Confucian group (mean difference: 1.18, p<0.001). However, in the trial 2 the effect is not significant in the Germanic group (mean difference: 0.08). The graphs in Figure 3 illustrate the direction of the effects. In the Confucian group the direction of the effect is in line with H3a as a greater efficiency is achieved in the models with colors. H3a is therefore supported. In the Germanic group we postulated a lower understanding efficiency in colorful models. However, the graph shows that the efficiency across cases using colors is higher. This effect is significant in trial 1 for the Germanic group. In the Germanic group no significant effect occurred in trial 2 in relation to the task efficiency. A possible explanation could be that the textual annotation cues neutralize the color effect. H3b is rejected.

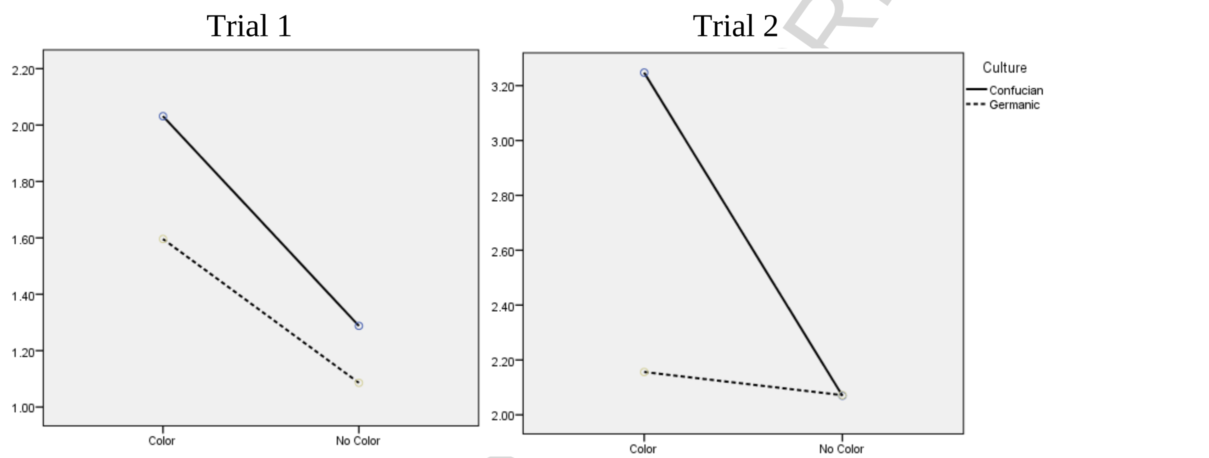


Figure 3. Estimated marginal means of task efficiency

Finally, the main effect culture is significant in trial 1 [F(1,115), 4.77, $p<0.05$] and the mean difference for trial 2 between both cultures is significant for the colored model (mean difference: -1.09, $p<0.001$). An interaction effect in trial 2 emerges as the difference between the two groups exists only in relation to the colored models. The Confucian group achieves in both trials higher task efficiency for models that use coloring than the Germanic group. H3c is supported.

6. Discussion

6.1 Summary of Results

We set out to study nine hypotheses about the impact of culture and the use of coloring in process models on users' ability to understand the process model. Table 7 summarizes the insights gained on our hypotheses.

The existing literature has only found spurious and inconsistent results regarding process model understanding efficiency [9, 15, 52]. Our results, however, suggest that colors can reduce the perception of cognitive load in form of perceived difficulty and they can improve task efficiency. The effect appears to be strongest for color differences within the Confucian group and between the

Germanic and Confucian group. The color stimulus aided Confucian individuals to increase the efficiency and to reduce the perceived difficulty, both for models with and without additional textual annotations. The usage of colors in the Germanic culture, however, seems to be more complex. Our results suggest that colors can increase the perceived difficulty but additional visual cues such as annotations compensate for this effect. Contrary to our speculation, colors also increased the task efficiency in the Germanic group, yet also this effect was neutralized in the model with annotations. Thus, bright colors can have both positive and negative effects in Germanic cultures. Overall, the effects in the Germanic group seem to be weaker as the cross-country comparison suggests.

6.2 Implications for Research

Our study is unique in that it uncovers the relevance of colors *in different cultures* to enhance the performance of visual communication via process models. The fact that most of our hypotheses were supported emphasizes the importance of this research, and demonstrates that cultural differences should not be ignored in process modeling.

We observe that color schemes that match cultural preferences can be applied mindfully to support visual communication. Yet, perceived difficulty, understanding accuracy and understanding efficiency *responded to a different degree* to our treatment. Importantly, we found only minor differences in the accuracy across treatments. Additionally, the covariate *process modelling knowledge* significantly affected the dependent variables. Interestingly, BPMN familiarity did not influence process model understanding. Appendix D contains an alternative model to explore if users with less BPMN familiarity benefit from colors. However, the results indicate that this is not the case.

This means that the explanation for the observed effects are likely related to efficient sensory-cognitive processing as opposed to intellectual potential. The difference in cognitive load might provide evidence for a more efficient usage of short-term memory capacity. According to this argument, the colors would guide the Confucian participants and let them use the short-term memory more efficiently, while the Germanic group was distracted. The actual understanding accuracy, however, seems not to be influenced by colors, indicating that the participants are able to handle the additional cognitive load.

Table 7: Summarized Interpretation of Hypotheses			
No	Hypothesis	Result from Tests	Interpretation
H1a	Members of the Confucian culture cluster will perceive models that contain coloring as less *difficult* to understand than models that do not contain coloring.	Supported	Colors support the visual channel and apparently reduce the cognitive load in process models for Confucian participants.
H1b	Members of the Germanic culture cluster will perceive models that contain coloring as more *difficult* to understand than models that do not contain coloring.	Partially Supported	Conceptual process models with colors add cognitive load in the content presentation format. The effect can apparently be compensated with additional textual cues (annotations).
H1c	Members of the Confucian culture cluster will perceive models that use coloring as less *difficult* than members of the Germanic culture cluster.	Supported	The results suggest a difference in the perception of difficulty in relation to process models. The use of colors leads to differently directed effects in the two groups and enhances cross-cultural differences.
H2a	Members of the Confucian culture cluster will achieve higher levels of *understanding* accuracy if the model contains coloring than if the model does not contain coloring.	Not supported	The Confucian group with colors achieved a higher score of understanding accuracy but the difference was not significant. No effects of colors on the extraneous load in relation to understanding accuracy were identified.
H2b	Members of the Germanic culture cluster will achieve higher levels of *understanding* accuracy if the model does not contain coloring than if the model contains coloring.	Not supported	The findings are inconsistent. In trial 1 the understanding accuracy was greater in the group with colors. In contrast, the model without colors achieved a higher score in trial 2. The difference between colored and non-colored models was not significant. No effects of colors on the extraneous load in relation to understanding accuracy were identified.
H2c	Members of the Confucian culture cluster will achieve higher levels of *understanding* accuracy for colored models than members of the Germanic culture cluster.	Not supported	The main effect of culture is not significant indicating that no cultural differences in relation to understanding accuracy exist.
H3a	Members of the Confucian culture cluster will achieve higher levels of task *efficiency* in tasks with models that contain coloring than for tasks with models that do not contain coloring.	Supported	The main effect color is significant and the direction of the effect is in-line with our hypothesis. Colors apparently reduce the extraneous load of process models and increase the task efficiency with Confucian participants.
H3b	Members of the Germanic culture cluster will achieve higher levels of task *efficiency* in tasks with models that do not contain coloring than for tasks with models that contain coloring.	Not supported	The main effect color is significant in trial 1. However, the effect has the same direction as in the Confucian group. Colors in process models increase task efficiency significantly. Additional textual cues seem to neutralize the effect.
H3c	Members of the Confucian culture cluster will achieve higher task *efficiency* in tasks with models that contain coloring than members of the Germanic culture cluster.	Supported	The reduction of extraneous load is greater in the Confucian group leading to a substantially increased efficiency compared to the Germanic group.

This finding is particularly relevant as it indicates that the effect of colors is limited. Apparently, cultural differences imply differences in cognitive processing but participants are able to compensate the additional load if they have the required additional time to reach similar levels of outcomes (i.e., understanding accuracy). These findings suggest that cultural differences may not necessarily imply a lack of accuracy of understanding, but culturally inappropriate modeling may lead to unnecessary additional cognitive effort on the model viewers, who feel that they need to invest more cognitive resources into developing an appropriate understanding of the modeled domain.

Also, our results add more understanding about second notation in process models. Colors provide a mechanism to perceptual discriminability of modeling elements without changing the semantics of the model itself [13]. Previous results derived from empirical data in the western world suggested that colors can reduce visual search in a diagram and can lead to a reduction in cognitive load, which translates into better model understanding [79]. Our results indicate that this is not necessarily always the case. Only if the colors match the receiver's preferences positive effects emerge while heterogeneous effects occur if the color scheme does not match. The realization of *cultural biases* in working with process models also poses additional questions around the acceptance, intensity and appropriation of modeling in different countries or cultural regions. For instance, it is likely that the additional effort caused by culturally inappropriate modeling might act as a barrier to acceptance of the models and thus to implications regarding systems analysis or design work based on models being rejected. Consequently, users may not fully utilize process models for decision making and conclusions drawn from the process models are potentially biased because some users regard them as too difficult to understand and are not willing to invest the required cognitive effort.

6.3 Implications for Practice

Our work predominantly has implications for design choices made in the process of process modeling in dependence of the cultural setting in which this work takes place. We studied the secondary notation, and thus our results provide an instrument to increase performance without changing process semantics. Cross-cultural development teams, for instance, provided with both colored and non-colored models, and given the possibility to choose between them, are probably likely to perform better in model understanding efficiency. Process models provide manifold opportunities to support organizational decisions including system implementation, process redesign or organizational innovation initiatives [1-3]. They further facilitate communication among team members and thus aid decision making within groups [4, 5]. Consequently, our results can be used to leverage process models more efficiently for decision support. To ease the use of process models by model viewers, culturally appropriate customization mechanisms could also be applied in contexts where model use is expected to occur in different cultural settings (e.g., in multi-national process standardization projects). A possible scenario would be a customized model design that is automatically used to represent the information in the best way for a specific cultural audience. Such customizations could be implemented in process modeling tool suites by offering views that either include or hide additional colors. Culturally sensitive color palettes such as those designed by Reinecke and Bernstein [16] could be used in this context. As a general guideline, our results suggest that it is always beneficial to use loud and bright colors for members of a Confucian culture. However, the country selection was based on average cultural dimension scores (see section 4.1). Several other countries such as Vietnam, Nepal, India and Indonesia have similar characteristics (see Appendix E). Further research should explore whether the identified color effects occur also in these countries. For members of a Germanic culture the application of an Asian color scheme could have potentially negative effects. Again it is possible that the same effects occur in countries with similar cultural dimension scores such as Belgium and France (Appendix E). Alternative color schemes in accordance to the Germanic culture are more likely beneficial and could further stimulate the performance. This includes less bright colors such as dark blue and grey. Unfortunately, compared to the Asian color preferences little is known about Germanic preferences as the literature compares usually Asia with the "western world" (see Appendix A). Therefore, it is highly relevant to determine whether differences between western cultures (e.g., House et al.'s [24] Anglo, Latin Europe and Germanic culture cluster) exist and how far they influence understanding performance. However, in mixed teams that share the same process model (e.g., in a face to face meeting) the communication via process models can be supported by the application of a Confucian color scheme in combination with additional visual cues to neutralize the negative effects for Germanic members.

Cultural customizations can also be included in modeling convention documents that stipulate best practice regulations for process modeling and their use. Furthermore, cultural biases could be included in training and education programs to raise awareness and suggest appropriate mitigation strategies when working with individuals from different backgrounds. In process modeling education, for instance, our findings raise awareness for potential struggles of students from different cultural backgrounds, not because of challenges in understanding essential modeling concepts such as concurrency, repetition and deadlocks but rather because of biases in the explicit representation of such concepts in graphical models.

6.4 Limitations

Our findings and implications are bounded by several limitations pertaining to our research. First, participants of our experiment were postgraduate students. This limits the external validity of our results. At best, the participants are comparable to novice organizational analysts. However, the results may not generalize to highly experienced practitioners. Additionally, our data collection assumes that individuals from China represent the Confucian culture cluster while individuals from Germany or Austria represent the Germanic culture cluster. The assignment of individuals from countries to cultural clusters is based on the GLOBE study [80]. We reason that an individual who has spent most of its life in one of these countries adopts the related culture of the cluster. Due to the dynamic character of culture it is possible that this country classification is not valid anymore. However, GLOBE is the most recent large-scale study in cross-cultural research and compared to previous studies (e.g., [81]) relatively young. Furthermore, the participants are not representative for the cultures as they cover only a specific age group (between 20 and 30) in selected locations and the sample size is too small to derive results for the basic population. Additionally, due to our experimental design not each member of the population had an equal probability of selection. However, it is a common practice to include students in cross-cultural experiments in order to discover behavior differences [82, 83]. As we did not focus on specific traits of expertise of senior analysts, we deem the choice of postgraduate student participants appropriate for our objectives.

Second, we did not measure cultural index scores in our study. We collected information about the country in which the participant was born as well as all countries the person has lived in. Based on the information we assigned the culture (Germanic and Confucian) and excluded participants that spent a substantial periods of time in other cultures (see section 4.1). This approach is based on the assumption that shared values exist among all individuals in a particular culture. However, we neither propose nor test a direct or moderating effect of culture dimensions on understanding performance. This would also require a correlation based study while we conducted a quasi-experiment (see section 4.1). Measuring culture on the individual level is not trivial [66, 67] but future research should explore if direct relations between culture dimensions and color preferences exist.

Third, our findings may be limited based on the selection of model cases. We used two models with similar complexity. Still, both considered models can be regarded as rather simple when compared to industry-sized models. Further research is required to explore if colors can increase understanding performance also in more complex models. Additionally, the color assignment in our setting was random. It might be possible that the application of a color scheme influences the performance as well. Thus, future research should explore whether the positive effects of colors can be further improved by supporting the semantic meaning (e.g., red for gateways) and how far cultural differences occur in relation to color assignment.

Fourth, differences in the understanding accuracy between trial 1 and trial 2 indicate, as expected, that the second model was easier than the first one. Still, a learning effect cannot be fully ruled out, and additionally we suspect that the domain of the model might also partially explain the differences. Students could have been more familiar with the online shopping scenario. These differences in the difficulty could influence the results of the color treatment. In relation to H1b, for instance, we identified only in relation to trial 1 that members of the Germanic cluster perceived models with colors as more difficult. It might be that trail 2 was too easy to detect this effect. Thus, the influence of process difficulty on the color perception should be further investigated. Our results further only indicate that differences exist initially. Nevertheless, it could be possible that the Germanic group is simply not used to these colors and over time, the differences could fade. A repeated measures approach is required to explore the consistency of cultural differences over time.

Finally, we took specific measures for avoiding a bias in the experiment. Item translation is critical, which we addressed with back-translation. Additional cultural factors differential in social desirability or in tester effects might play a role though not explicitly modeled in the experiment [69, p. 309]. Still, hypotheses H1a and H1b as well as H3a and H3b find differences in how far the treatments affect each of the two groups (A and B). These comparisons should be culturally unbiased by definition as they refer to effects within each culture separately. Furthermore, in every quasi-experiment biases could emerge in relation to confounding variables. While we can rule out related effects in relation to BPMN familiarity, process management knowledge, and gender (see Appendix D), it is possible that further confounding variables exist. For example, it may be that participants varied in their attitude towards process modeling in general or BPMN in particular.

7. Conclusion

Process models can facilitate the communication of technical information and related decisions making among organizational stakeholders from different countries. Secondary notation can be used to provide additional, culture-dependent visual cues without impeding the semantics provided by process modeling grammars.

The results of our study provide evidence for the presence of cross-cultural differences, on the basis of which we developed first normative advice about how models should be designed in order to allow users in different cultural settings to increase understanding efficiency and reduce the perceived difficulty of this form of communication. Our work also shows that in particular Confucian model user’s benefit from this color scheme. In contrast, the results from the Germanic group indicate that even if the colors do not match cultural preferences, the understanding performance does not decline.

Overall, our work adds to the growing body of knowledge on design choices in process modeling and their outcomes, and thus paves the way for a cultural adaptive design of modeling grammars that supports decision making in the analysis and design of information systems.

References

[1] I. Davies, P. Green, M. Rosemann, M. Indulska, S. Gallo, How do practitioners use conceptual modeling in practice?, Data & Knowledge Engineering, 58 (2006) 358-380.
[2] K. Figl, J. Recker, J. Mendling, A Study on the Effects of Routing Symbol Design on Process Model Comprehension, Decision Support Systems, 54 (2013) 1104-1118.
[3] H.A. Reijers, T. Freytag, J. Mendling, A. Eckleder, Syntax Highlighting in Business Process Models, Decision Support Systems (DSS), 51 (2011) 339-349.
[4] M. Dumas, M. La Rosa, J. Mendling, H.A. Reijers, Fundamentals of Business Process Management, Springer Heidelberg, 2013.

[5] N. Kock, J. Verville, A. Danesh-Pajou, D. DeLuca, Communication Flow Orientation in Business Process Modeling and Its Effect on Redesign Success: Results from a Field Study, Decision Support Systems, 46 (2009) 562-575.
[6] A.-M. Ozimec, M. Natter, T. Reutterer, Geographical information systems-based marketing decisions: Effects of alternative visualizations on decision quality, Journal of Marketing, 74 (2010) 94-110.
[7] B. Efrim, J., C. Carnaghan, P.S. Alencar, Business Modeling to Improve Auditor Risk Assessment: An Investigation of Alternative Representations, Journal of Information Systems, 28 (2014) 231-256.
[8] J.S. Giboney, S.A. Brown, P.B. Lowry, J.F. Nunamaker Jr., User acceptance of knowledge-based system recommendations: Explanations, arguments, and fit, Decision Support Systems, 72 (2015) 1-10.
[9] J. Mendling, M. Strembeck, J. Recker, Factors of Process Model Comprehension - Findings from a Series of Experiments, Decision Support Systems, 53 (2012) 195-206.
[10] K. Figl, J. Mendling, M. Strembeck, The Influence of Notational Deficiencies on Process Model Comprehension, J Assoc Inf Syst, 14 (2013) 312-338.
[11] M. La Rosa, A.H.M. ter Hofstede, P. Wohed, H.A. Reijers, W.M.P. van der Aalst, Managing Process Model Complexity via Concrete Syntax Modifications, IEEE Transactions on Industrial Informatics, 7 (2011) 255-265.
[12] T.R.G. Green, M. Petre, Usability Analysis of Visual Programming Environments: A 'Cognitive Dimensions' Framework, Journal of Visual Languages and Computing, 7 (1996) 131-174.
[13] D.L. Moody, The “Physics” of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering, IEEE Trans. Software Eng. (TSE), 35 (2009) 756-779.
[14] P. Bera, Does Cognitive Overload Matter in Understanding BPMN Models?, Journal of Computer Information Systems, 52 (2012) 59-69.
[15] J. Recker, Empirical Investigation of the Usefulness of Gateway Constructs in Process Models, Eur J Inf Syst, 22 (2013) 673-689.
[16] K. Reinecke, A. Bernstein, Knowing What a User Likes: A Design Science Approach to Interfaces that Automatically Adapt to Culture, MIS Quarterly, 37 (2013) 427-453.
[17] W. Qiuye, A Cross-cultural Comparison of the Use of Graphics in Scientific and Technical Communication, Technical Communication, 47 (2000) 553-560.
[18] T.J. Madden, K. Hewett, M.S. Roth, Managing Images in Different Cultures: A Cross-National Study of Color Meanings and Preferences, Journal of International Marketing, 8 (2000) 90-107.
[19] T. Schmiedel, J. vom Brocke, J. Recker, Which Cultural Values Matter to Business Process Management? Results from a Global Delphi Study, Business Process Management Journal, 19 (2013) 292-317.
[20] K. Jensen, Coloured Petri Nets: Basic Concepts, Analysis Methods and Practical Use, Springer, Berlin, Germany, 1992.
[21] J. Noiwan, A.F. Norcio, Cultural differences on attention and perceived usability: Investigating color combinations of animated graphics, Int. J. Human-Computer Studies, 64 (2006) 103–122.
[22] B.A.C. Saunders, J.v. Brakel, Are there nontrivial constraints on colour categorization?, Behavioral and Brain Sciences, 20 (1997) 167-179.
[23] D. Roberson, J. Davidoff, I.R.L. Davies, L.R. Shapiro, Color categories: Evidence for the cultural relativity hypothesis, Cognitive Psychol, 50 (2005) 378-411.
[24] R.J. House, P.J. Hanges, M. Javidan, P.W. Dorfman, V. Gupta, (Eds.), Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies, Sage, Thousand Oaks, 2004.
[25] J. Mendling, H.A. Reijers, J. Recker, Activity Labeling in Process Modeling: Empirical Insights and Recommendations, Information Systems (IS), 35 (2010) 467-482.
[26] J.V. White, Using Charts and Graphs: 1000 Ideas for Visual Persuasion, R. R. Bowker Company, New York and London, 1984.
[27] L.E. Brasseur, Visualizing Technical Information: A Cultural Critique, Baywood Publishing, Amityville, NY, 2003.
[28] S.M. Kosslyn, Elements of Graph Design, W. H. Freeman, New York, 1994.
[29] B. Tversky, Visualizing Thought, Topics in Cognitive Science, 3 (2011) 499-535.
[30] G.H. Hofstede, Culture’s Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations, Sage, Beverly Hills, 2001.

[31] D.E. Leidner, T. Kayworth, A review of culture in information systems research: Towards a theory of information technology culture conflict, MIS Quarterly, 30 (2006) 357–399.
[32] P. Shachaf, Cultural Diversity and Information and Communication Technology Impacts on Global Virtual Teams: An Exploratory Study, Information & Management, 45 (2008) 131-142.
[33] M. Aiken, J. Martin, A. Shirani, T. Singleton, A group decision support system for multicultural and multilingual communication Decision Support Systems, 12 (1994) 93-96.
[34] S. McCoy, D.F. Galletta, W.R. King, Applying TAM Across Cultures: The Need for Caution, Eur J Inf Syst, 16 (2007) 81-90.
[35] L.L. Tung, M.A. Quaddusb, Cultural differences explaining the differences in results in GSS: implications for the next decade, Decision Support Systems, 33 (2002) 177–199.
[36] A. Kankanhalli, B.C.Y. Tan, K.-K. Wei, M.C. Holmes, Cross-cultural differences and information systems developer values, Decision Support Systems, 38 (2004) 183-195.
[37] G. Dietz, M. Juhrisch, Negotiating language barriers – a methodology for cross-organisational conceptual modelling, Eur J Inf Syst, 21 (2012) 229-254.
[38] Y. Wang, D. Wang, Cultural Contexts in Technical Communication: A Study of Chinese and German Automobile Literature, Technical Communication, 23 (2009) 39-50.
[39] L. Jacobs, C. Keown, R. Worthley, K.-I. Ghymn, Cross-cultural Colour Comparisons: Global Marketers Beware!, International Marketing Review, 8 (1991) 21-30.
[40] P. Sable, O. Akcay, Color: Cross Cultural Marketing Perspectives As To What Governs Our Response To It, American Society of Business and Behavioral Sciences, 17 (2010) 950-954.
[41] W. Barber, A. Badre, Culturability: The merging of culture and usability, Proceedings of the Conference on Human Factors and the Web, 1998.
[42] E. Würtz, A cross-cultural analysis of websites from high-context cultures and low-context cultures, J Comput-Mediat Comm, 11 (2005).
[43] M.A. Webster, P. Kay, Individual and population differences in focal colors, in: R.E. MacLaury, G.V. Paramei, D. Dedrick (Eds.) Anthropology of Color: Interdisciplinary multilevel modeling, John Benjamins, Amsterdam, 2007, pp. 29-53.
[44] A. Franklin, A. Clifford, E. Williamson, I. Davies, Color term knowledge does not affect categorical perception of color in toddlers, Journal of Experimental Child Psychology, 90 (2005) 114-141.
[45] P. Athanasopoulos, L. Damjanovic, A. Krajciova, M. Sasaki, Representation of Colour Concepts in Bilingual Cognition: The Case of Japanese Blues, Bilingualism: Language and Cognition, 14 (2011) 9-17.
[46] K. Reinecke, A. Bernstein, Improving Performance, Perceived Usability, and Aesthetics with Culturally Adaptive User Interfaces, ACM Trans. Comput.-Hum. Interact., 18 (2011) Article 8.
[47] A. Marcus, E.W. Gould, Cultural dimensions and global web design: What? So what? Now what?, Proceedings of the Conference on Human Factors and the Web, 2001.
[48] E.T. Hall, M.R. Hall, Understanding cultural differences, Intercultural Press, Yarmouth, 1990.
[49] D.L. Moody, Why a Diagram is Only Sometimes Worth a Thousand Words: An Analysis of the BPMN 2.0 Visual Notation, 2011.
[50] R.E. Mayer, Multimedia Learning, 2. ed. ed., Cambridge University Press2009.
[51] J. Recker, H.A. Reijers, S.G. van de Wouw, Process Model Comprehension: The Effects of Cognitive Abilities, Learning Style, and Strategy, Communications of the Association for Information Systems, 34 (2014) 199-222.
[52] J. Recker, A. Dreiling, The Effects of Content Presentation Format and User Characteristics on Novice Developers' Understanding of Process Models, Communications of the Association for Information Systems, 28 (2011) 65–84.
[53] K.E. DeLeeuw, R.E. Mayer, A Comparison of Three Measures of Cognitive Load: Evidence for Separable Measures of Intrinsic, Extraneous, and Germane Load, J Educ Psychol, 100 (2008) 223–234.
[54] W. Huang, P. Eades, S.-H. Hong, Measuring Effectiveness of Graph Visualizations: A Cognitive Load Perspective, Information Visualization, 8 (2009) 139-152.
[55] J.L. Plass, D.M. Chun, R.E. Mayer, D. Leutner, Cognitive Load in Reading a Foreign Language Text with Multimedia Aids and the Influence of Verbal and Spatial Abilities, Computers in Human Behavior, 19 (2003) 221-243.

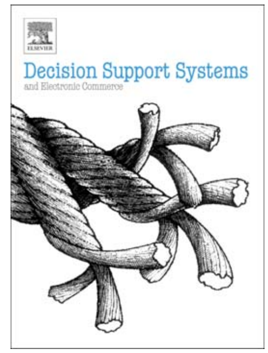
[56] A. Burton-Jones, P.N. Meso, The Effects of Decomposition Quality and Multiple Forms of Information on Novices' Understanding of a Domain from a Conceptual Model, J Assoc Inf Syst, 9 (2008) 784-802.
[57] D. Lewis, A. Barron, Animated Demonstrations: Evidence of Improved Performance Efficiency and the Worked Example Effect, Proceedings of the 1st International Conference on Human Centered Design: Held as Part of HCI International 2009, Springer, 2009, pp. 247-255.
[58] P.W.M. Van Gerven, F. Paas, J.J.G. Van Merrienboer, H.G. Schmidt, Cognitive load theory and aging: Effects of worked examples on training efficiency, Learning and Instruction, 12 (2002) 87-105.
[59] A. Burton-Jones, Y. Wand, R. Weber, Guidelines for Empirical Evaluations of Conceptual Modeling Grammars, J Assoc Inf Syst, 10 (2009) 495-532.
[60] T.D. Cook, D.T. Campbell, Quasi-Experimentation: Design and Analysis Issues for Field Settings, Houghton Mifflin, Boston, MA, 1979.
[61] H. Fehr-Duda, M. De Gennaro, R. Schubert, Gender, financial risk, and probability weights, Theory and Decision, 60 (2006) 283-313.
[62] M.F. Rivas, An Experiment on Corruption and Gender, Bulletin of Economic Research, 65 (2013) 10-42.
[63] B.A. Thyer, Quasi-experimental research designs, Oxford University Press, New York, 2012.
[64] B.E. Huitema, The analysis of covariance and alternatives : statistical methods for experiments, quasi-experiments, and single-case studies, 2nd ed., Wiley, Hoboken, N.J., 2011.
[65] T.-F. Kummer, T. Schmiedel, Reviewing the Role of Culture in Strategic Information Systems Research: A call for Prescriptive Theorizing on Culture Management, Communications of the Association for Information Systems, 38 (2016) article 5.
[66] P.E. Spector, C.L. Cooper, K. Sparks, An International Study of the Psychometric Properties of the Hofstede Values Survey Module 1994: A Comparison of Individual and Country/Province Level Results., Applied Psychology, 50 (2001) 269-281.
[67] B. Yoo, N. Donthu, T. Lenartowicz, Measuring Hofstede's five dimensions of cultural values at the individual level: Development and validation of CVSCALE, Journal of International Consumer Marketing, 23 (2011) 193-210.
[68] N. Dawar, P. Parker, Marketing Universals: Consumers' Use of Brand Name, Price, Physical Appearance, and Retailer Reputation as Signals of Product Quality, Journal of Marketing, 58 (1994) 81-95.
[69] L. Copeland, L. Griggs, Going International: How to make friends and deal effectively in the global marketplace, Random House, New York, 1985.
[70] J. Recker, Continued Use of Process Modeling Grammars: The Impact of Individual Difference Factors, Eur J Inf Syst, 19 (2010) 76-92.
[71] A. Burton-Jones, P.N. Meso, Conceptualizing Systems for Understanding: An Empirical Test of Decomposition Principles in Object-Oriented Analysis, Information Systems Research, 17 (2006) 38-60.
[72] OMG, Business Process Model and Notation, Object Management Group, 2012.
[73] D.C. Howell, Statistical Methods for Psychology, 7th ed., Wadsworth Cengage Learning, Belmont, California, 2009.
[74] F.J.R. van de Vijver, N.K. Tanzer, Bias and equivalence in cross-cultural assessment: An overview., Eur Rev Appl Psychol, 47 (1997) 263-280.
[75] G.S. Choi, R. Oehlmann, H. Dalke, D. Cottington, Cross-Cultural Comparison of Color Preferences between English and Korean Subjects, Proceedings of 6th Social Intelligence Design Conference, 2007, pp. 169-178.
[76] F. Faul, E. Erdfelder, A.-G. Lang, B. Axel, G*Power 3: A Flexible Statistical Power Analysis for the Social, Behavioral, and Biomedical Sciences, Behavior Research Methods, 39 (2007) 175-191.
[77] A.C. Elliott, W.A. Woodward, Statistical Analysis Quick Reference Guidebook: With SPSS Examples, SAGE, Thousand Oaks, 2007.
[78] R.M. Heiberger, B. Holland, Statistical Analysis and Data Display: An Intermediate Course with Examples in S-PLUS, R, and SAS, Springer, New York, 2004.
[79] H.A. Reijers, J. Mendling, A Study into the Factors that Influence the Understandability of Business Process Models, IEEE Transactions on Systems Man & Cybernetics, Part A (SMCA), 41 (2011) 449-462.

[80] R.J. House, M. Javidan, Overview of GLOBE, in: R.J. House, P.J. Hanges, M. Javidan, P.W. Dorfman, V. Gupta (Eds.) Culture, leadership and organizations: The GLOBE study in 62 societies, Sage, Thousand Oaks, 2004, pp. 9–18.
[81] G.H. Hofstede, Culture’s consequences: International differences in work related values, Sage, Beverly Hills, 1980.
[82] M. Minkov, Cross-Cultural Analysis: The Science and Art of Comparing the World's Modern Societies and Their Cultures, SAGE, Thousand Oaks, 2013.
[83] R. Fischer, About chicken and eggs: four methods for investigating culture–behaviour links, in: F.J.R. van de Vijver, A. Chasiotis, S.M. Breugelmans (Eds.) Fundamental Questions in Cross-Cultural Psychology, Cambridge University Press, New York, 2011, pp. 190-213.

Highlights:

- We explore how cross-cultural differences affect business process modelling
- We use cultural-adjusted color schemes to explore effects on understandability
- The results indicate that colors increase efficiency and reduce perceived difficulty
- We develop first normative advice about cultural adaptive model design

Accepted Manuscript



Enhancing understandability of process models through cultural-dependent color adjustments

Tyge-F. Kummer, Jan Recker, Jan Mendling

PII: S0167-9236(16)30057-4
DOI: doi: 10.1016/j.dss.2016.04.004
Reference: DECSUP 12709

To appear in: *Decision Support Systems*

Received date: 22 December 2015
Revised date: 19 April 2016
Accepted date: 19 April 2016

Please cite this article as: Tyge-F. Kummer, Jan Recker, Jan Mendling, Enhancing understandability of process models through cultural-dependent color adjustments, *Decision Support Systems* (2016), doi: 10.1016/j.dss.2016.04.004