Human Interaction under Risk in Cyber-Physical Production Systems

Philipp Brauner, Ralf Philipsen, André Calero Valdez, Martina Ziefle

Human-Computer Interaction Center
RWTH Aachen University
Campus Boulevard 57, 52074 Aachen, Germany
brauner|philipsen|calero-valdez|ziefle@comm.rwth-aachen.de

Abstract. The emergence of cyber-physical production systems poses new challenges for designing the interface between production systems and the human-in-the-loop. In this study, we investigate how human operators interact with risks in a supply chain scenario. We varied the financial magnitude and the expected value of the decisions, the combination of two types of risk (risk in delivery amount and risk in timeliness), as well as three different task displays as within-subject factors. As explanatory user factors we studied the influence of Need for Achievement and the Attitude towards Risk-taking on the dependent variables task speed and accuracy. Results of the user study with 33 participants show that each of the investigated factors either influences decision speed, decision accuracy, or both. Consequently, the human-in-the-loop profits from adequate decision support systems that help to increase decision efficiency and effectiveness and reduce uncertainty and workload. The article concludes with a research agenda to support the human-in-the-loop in production systems.

Keywords: Decision Support Systems, Cyber-Physical Production Systems, Human Factors, Decision under Risk, Socio-Technical System, Risk

1 Introduction

We are at the brink of the 4th industrial revolution and the convergence of information and communication technology with production systems along the value chain promises increased efficiency and effectiveness and overall competitiveness for individual companies, supply chain networks, and societies [1, 2].

Despite increased automation the human-in-the-loop remains a crucial component of cyber-physical production system (CPPS). Bainbridge's *Ironies of Automation* [3] postulate that automation is shifting the role of the human from manual activities to monitoring and controlling tasks. However, if automation fails, the capability to successfully intervene declines through lack of practice, lower knowledge of the underlying processes, or ill-aligned mental models.

In near future, a vast amount of fine-grained, heterogeneous data from different sources will be available in real for automatic machine control, for decision support systems, and the human-in-the-loop interacting with or supervising these systems [4].

To bridge the gap between the human decision maker and the increasingly complex socio-technical cyber-physical production system (STCPPS) we need to understand their specific wants and needs, as well as their emotional and cognitive characteristics. While many generic norms (e.g., *EN ISO 9241*), heuristics [5], and guidelines [6] exist for designing user interfaces, the task-specific requirements for cyber-physical production systems are insufficiently understood. In this work, we focus on the specific application domain of inter-company flow of material and information. One key skill of the human-in-the-loop in this domain is decision making under risk and uncertainty, e.g., the ability to anticipate delivery variances and bottlenecks. To develop user-adaptive decision support systems that augment the decision makers abilities, we first need to identify their abilities, weaknesses, and potential biases by studying their decision-making processes under uncertain and risky conditions.

2 Interacting with Uncertainty and Risk

Knight distinguishes *Uncertainty and Risk* as two distinct types of uncertainty: When the potential outcomes of an event are not known in advance, this is referred to as Uncertainty. In contrast, the term risk is used when the potential outcomes of an event are known beforehand [7]. For example, the two possible outcomes of a coin-toss are known and the chances are quantifiable. Consequently, the outcomes can be maximized by maximizing the expected utility.

Tversky and Kahneman's *Prospect Theory* [8] showed that decisions under risk are prone to irrational behavior: For example, people prefer lower but likely profits over higher but unlikely profits. In contrast, higher but unlikely losses are preferred to lower but likely losses. Examples for decisions with uncertainty include decisions in complex, non-linear, dynamic systems, such as cyber-physical production networks with globally dispersed supply chain networks [9]. Managing uncertainty is more complex and – in the context of cyber-physical production systems – requires concepts such as resilient design of the CPPS, closer cooperation of the stakeholders, supply chain agility, and a culture for handling uncertainty [9].

In general, people are rarely objective in making decisions, but make use of a series of decision heuristics to compensate for abundant, incomplete, or inaccurate information [10]. Even though these heuristics often enable efficient and effective decision making, they still bear the risk of irrational or erroneous decisions that can negatively influence the outcome.

3 Method

With this study, we want to analyze the influence of risk on speed and accuracy of decision in the context of inter-company flow of materials and information. We embedded the study into a novel business simulation game, to both evaluate the factors and their interrelationships mentioned above and make it challenging and captivating for the participants. Section 3.1 presents the game and the decision tasks for the participants. Section 3.2 defines the within-subject variables considered in the study and section 3.4 presents the dependent variables from the experiment. Next, section 3.3 elucidates the user factors studied in this experiment. Finally, section 3.5 presents the hypotheses that guided the study.

3.1 Decision Game

Business simulation games facilitate studying human behavior in non-trivial, sufficiently complex, and experimentally controllable environments [11].

For this study, participants need to work on a series of decision tasks in the context of a purchasing department of manufacturing company. Each task resembles a business offer with a *required price* that must be payed and a *potential profit* that might be realized. The potential profit depends on two types of risk that must be taken into account: A risk in regard to the *quantity* of the delivered goods and a risk in regard to the *punctuality* of the delivery. If only a share of the delivery arrives (quantity risk), only an equal share of the revenues is generated. No profits are generated if the order does not arrive in time (punctuality risk).

For each decision task the participant needs to decide whether to *accept* or to *reject* the offers. Whether an offer has a positive, neutral, or negative expected value (*EV*) is not directly shown, but must be inferred from the presented risks, the price, and the potential profit. If an offer is rejected, nothing is won or lost. If an offer is accepted, its price is deducted, and a revenue is calculated depending on the quantity and punctuality of the delivery.

Finally, the overall *profit* of the participants accumulates through wins and losses across several decision tasks of the game. In the game the participants' task was to maximize their overall profit by *accepting* lucrative orders ($EV>0$) and by *declining* unprofitable orders ($EV<0$).

The business simulation game is designed for use in controlled laboratory environments as well as for widespread online studies. In this case, a laboratory experiment was used to obtain an initial evaluation of the framework.

3.2 Independent Variables

Within the experiment we varied the four within-factor variables time risk, quantity risk, magnitude, and expected value.

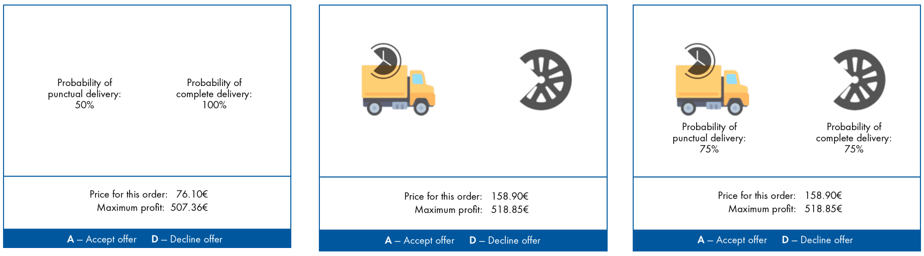
Time Risk: Based on chance, the delivery is either punctual or delayed. The chance for a punctual delivery is either 100%, 75%, or 50%. No revenue can be generated for a delayed delivery.

Quantity Risk: Based on chance, the delivery is made in full (100%) or partially (75% or 50%). For partial deliveries, profits can only be realized proportionately (75% for deliveries with 75% chance and 50% for deliveries with 50% chance).

Magnitude: To study the influence of the magnitude of the decision, the offers were scaled by 100 or 2000 for offers with a lower or higher financial volume.

Expected Value: The expected value of each offer is varied between −20%, −10%, 0%, +10%, +20%. Thus, gains or losses of ±20%, ±10%, or 0% can be realized based on chance and the offers' price.

Visualization: Three different risk visualizations were administered in distinct rounds of the business simulation game. First, a textual representation of the risk, a purely visual representation, and a combined textual and visual representation (Fig. 1).



(a) Text only. (b) Image only. (c) Text & image

Fig. 1. Three task visualizations used in the experiment.

All factors are taken into account for calculating the required price and the potential profit of an offer. Each offer (i.e., decision task) then presents its price, potential profit, the time and quantity risk, as well as the magnitude of the decision. Apparently, the expected value is not presented and must be inferred from the other numbers. For each of the combinations decision tasks were constructed (uniformly distributed), shuffled, and presented to the participants.

3.3 Explanatory User Factors

We administered a survey before the experiment in the business simulation game to understand if individual user factors influence the interaction with risk in cyber-physical production systems. Besides the participant's age and gender the following latent constructs were queried:

Risk Attitude: An individuals' attitude towards taking risk may shape the performance in the decision tasks. In this explanatory study we used Beierlein et al.'s validated single item scale to capture this risk attitude [12].

Need for Achievement (NfA): Motivation plays a central role for the choice of actions and the performance in these actions. The Need for Achievement inventory by Schuler and Prochaska [13] is a psychometric instrument for assessing an individuals' need for achievement. We used a short scale with 6 items and an outstanding internal reliability of α=.977.

3.4 Dependent Variables

As dependent variable we measured the accuracy, as well as the speed of the decision. The accuracy is based only on the expected value of the decision under risk and not the actual outcome: A decision is considered as accurate, if the $EV>0$ and as inaccurate for $EV<0$.

The participants were instructed to aim for highest profit. Speed was neither instructed, nor used for feedback. Consequently, the following evaluation focuses on the accuracy and speed is only reported for information.

3.5 Hypothesis

The experiment is guided by the following research questions, derived from literature and qualitative preliminary studies:

H1: Tasks with higher risks yield lower speed and lower accuracy. Tasks with combined risks are particular difficult for participants (lowest speed, lowest accuracy).
H2: A task's expected value influences both speed and accuracy of the decision.
H3: A decision's financial magnitude has a significant influence on accuracy.
H4: Double coding of risks through text *and* images increases speed compared to single coding.
H5: Higher need for achievement relates to higher accuracy and lower speed.
H6: Higher attitude towards risk relates to higher speed and lower accuracy.

The testing of these fundamental hypotheses is intended on the one hand to evaluate the relationships between risk and decision-making behavior in the context of inter-company flow and on the other hand to evaluate the research framework developed.

3.6 Statistical procedures

We used parametrical and non-parametrical methods (Person's r, Spearman's ρ correlations, single and multivariate (repeated measures) analyses of variance (RM-M/ANOVA). The level of significance is set to α=.05 and Pillai's V is used for multivariate tests. if sphericity is not met, Greenhouse-Geisser corrected values are used but uncorrected *df*s are reported or legibility. Arithmetic means are reported with standard deviation denoted as ±. Only trials with an expected value ≠0% are considered, as otherwise accuracy is not well defined.

3.7 Description of the Sample

33 subjects in the age from 22 to 45 years (M=26.4, Md=26, SD=2.3) have participated in the study, 28 were female and 5 were male. Within the sample age and gender is not correlated [$\rho(n=33-2)=.069, p>.05$].

4 Results

This section presents the results for each of the investigated factors. First, the influence of the system factors are presented, followed by a presentation of the interface factors and concluded with an analysis of the individual user factors. Overall, the mean median reaction time is 2.851±1.404s and the average mean accuracy is 73.1±7.1%.

4.1 Effects of the System Factors

The **Magnitude** of the orders has a significant overall effect on the participant's decision as affirmed by a RM-MANOVA [V=.416, $F_{2,31}$=11.050, p<.001, η^2=.416]. Specifically, **Magnitude** influences the reaction time [p<.001, η^2=.388], but not the accuracy of the decisions (p=.272>.05). Decision tasks with a higher financial volume were carried out slower than tasks with lower volume, but with about the same accuracy (see Tab. 1a).

The **Expected Value** of the tasks has an overall effect on the decisions [V=.780, p<.001, η^2=.780] and it influences the accuracy [p=.014, η^2=.159], but not the speed of the decisions [p=.295>.05]. Decisions with a lower absolute expected value (i.e., ±10%) had a lower accuracy than decisions with higher expected values (i.e., ±20%). Whether the expected value of the decision task had a positive or negative outcome had no significant influence on the accuracy (see Tab. 1b).

Table 1. Speed and Accuracy of decisions depending on…

(a) Level (η^2=.416)

	Speed [ms]	Accuracy [%]
100€	2739 ± 1353	$72.7 \pm 6.9\%$
2000€	2972 ± 1473	$73.6 \pm 8.0\%$

(b) Expected Value (η^2=.780)

	Speed [ms]	Accuracy [%]
-20%	2793 ± 1376	70.4 ± 18.5 %
-10%	2881 ± 1405	59.1 ± 16.7 %
+10%	2888 ± 1428	64.7 ± 10.9 %
+20%	2920 ± 1592	71.5 ± 10.8 %

A multivariate repeated measures ANOVA with the two factors **risk time** and **risk quantity** yield a significant effect for both main factors [risk time V=.772, $F_{4,29}$=24.491, p<.001, η^2=.772; risk quantity V=.708, $F_{4,29}$=17.606, p<.001, η^2=.708] and the interaction of both within-subject factors [V=.789, $F_{8,25}$=11.696, p<.001, η^2=.789]. A further analysis shows that risk in time, quantity, and their combination influences decision accuracy, whereas decision speed is only affected by risk in quantity. As expected, the participants were fastest and achieved highest accuracy scores (91.1±9.3%) for tasks with no risk (see Table 3). They were equally fast but achieved lowest accuracy scores (68.1±5.8%) for tasks with the highest risk.

Table 2. Speed and Accuracy of decisions depending on…

(a) Task visualization (η^2=.397)

	Speed [ms]	Accuracy [%]
Text	3066 ± 1956	$72.3 \pm 7.9\%$
Image	2676 ± 1052	$75.0 \pm 8.1\%$
Text/Image	2879 ± 1531	$72.2 \pm 7.7\%$

(b) Repetition (η^2=.400)

	Speed [ms]	Accuracy [%]
1st round	3351 ± 1850	73.9 ± 7.2 %
2nd round	2785 ± 1552	72.9 ± 8.6 %
3rd round	2484 ± 1060	72.8 ± 8.2 %

Table 3. Effect of Risk in Delivery Time (η^2=.772) and Quantity (η^2=.708).

Risk Time	Risk Quantity	Speed [ms]	Accuracy [%]
50%	50%	2514 ± 1210	$68.1 \pm 5.8\%$
	75%	2814 ± 1459	$70.0 \pm 9.8\%$
	100%	3005 ± 1586	$71.8 \pm 10.7\%$
75%	50%	3208 ± 1921	$68.5 \pm 10.6\%$
	75%	2936 ± 1444	$68.8 \pm 8.8\%$
	100%	3019 ± 1699	$69.6 \pm 9.1\%$
100%	50%	3044 ± 1339	$70.9 \pm 11.2\%$
	75%	2848 ± 1489	$79.7 \pm 10.9\%$
	100%	2450 ± 1104	$91.1 \pm 9.3\%$

4.2 Effects of the User Interface

Presentation of the task (image, text, or combined) had a medium overall effect on the participants decision performance [V=.397, $F_{4,29}$=4.778, p=.004<.05, η^2=.397]. The average **Speed** of the decisions was not significantly different for the three task visualizations [$F_{2,64}$=2.102, p=.139>.05, η^2=.062], but task **Accuracy** was significantly higher for the image condition than for textual or combined task presentation [$F_{2,64}$=4.885, p=.012<.05, η^2=.132] (see Table 2a).

4.3 Effects of Personality Factors

First, we checked for the influence of an individual's **Need for Achievement** on the decisions using an ANCOVA and **Need for Achievement** as covariate and speed and accuracy as dependent variables. Both variables were significantly influenced [F=5.213, p=.011<.05, η^2=.258] and mean median correctness was significantly higher [p=.009<.05, η^2=.202] whereas decision speed [p=.014<.05, η^2=.179] were significantly lower for participants with a higher **Need for Achievement**.

A RM-MANOVA with **Round** as independent within-subject variable and **Speed** and **Accuracy** as dependent variables yield a significant model. Thus, practice has a significant effect on the overall decisions [V=.400, $F_{4,29}$=4.837, p<.001, η^2=.400]. speed of the decisions [p<.001, η^2=.314], but not on their accuracy [p>.564]. Consequently, the decision speed increases with practice, while the decision accuracy remains constant.

5 Discussion

We developed a novel evaluation framework for studying the influence of personality (*user factors*), the interface between the human operator and (*interface factors*), and parameters of the underlying production system (*system factors*) on decisions in cyber-physical production systems.

Within this explanatory study, we could show that *all* investigated factor domains influence decision efficiency and/or effectiveness in our material disposition scenario.

Our first hypothesis (H1)—higher risk have a negative effect on decisions speed and accuracy—has been confirmed. Our second assumption—the combination of different risks has a negative impact—could not be shown in this study, as even single risks had a disastrous effect on performance. A more granular gradation of the risk levels in future experiments may still corroborate the presumed effect of the combined risks.

Our study partially falsifies the second hypothesis (H2) that a task's expected value has an influence on performance. Apparently, decision tasks with a lower expected value (i.e., ±10%) are more difficult and yield higher error rates in contrast to tasks with higher expected values (i.e., ±20%). Surprisingly, speed was not affected by the expected value.

In contrast, we found that the financial volume of a decision had a significant role on performance (speed, but not accuracy). Consequently, H3 is falsified. The study shows that the participants invested more times on decisions with a larger volume. It can be assumed that larger amounts of money and thus larger possible losses could lead to greater care in the sense of a higher time-investment in the decision. However, if the difficulty of the task remains the same, longer processing only does not automatically lead to higher accuracy in the present context.

Within this study, we focused on the visualization of the risk as one of the many imaginable interface factors. To our surprise, the presentation had no effect on the decision speed, but on the accuracy of the decisions. We have assumed that a double coding of the task (*text & image*) would yield highest performance (H4). However, pure image coding had the highest accuracy within this study. Why this is the case cannot yet be finally explained on the basis of the available data. Limitingly, it has to be considered that the pure image representation has led to significantly higher accuracy values, but the absolute accuracy values of all forms of representation differ only slightly.

Looking at user factors in terms of personality traits, it turned out that both H5 and H6 have been confirmed. As expected, a high need for achievement leads to longer processing times and more accurate results (H5). Conversely, a higher personal willingness to take risks leads to faster decisions, but also to less accurate results (H6).

In conclusion, the study shows that decision speed and most importantly decision quality is influenced by the personality of the participants as well as by the type of decision task. While the present study provides only a first glimpse at some of the numerous factors that influence human performance, it provides a viable experimental framework to study the interaction of personality, and interface design, and the various components of the underlying production system.

6 Limitations and Outlook

Obviously, the task in this baseline experiment is trivial to automate. But as Bainbridge's *ironies of automation* [3,14] postulate, automation will not make the human actor in CPPS superfluous, but will rather shift its role to supervisory tasks. In case of failing automation or when the automation need to be supervised, the human agent will still need to process the available information, evaluate its meaning, and make a correct decision. In these cases, decision biases, the handling of incomplete information, personality, as well as interface influence the overall decision quality, the perceived workload, and also job satisfaction.

However, the number of investigated factors and the limited number of experimental trials forced us to analyze the data using a sequence of singular tests. Future work should integrate more system, interface, and human factors in a common statistical model that facilitates their prioritization in regard to their influence on efficiency and effectiveness. For example, the study was not controlled for perceptual speed, which was found to be important in similar studies on decision performance without risk. Likewise, the task complexity was rather low as only singular decisions had to be made at any given time. We expect that different and more pronounced effects will emerge in even more complex decision situations, when additional and more complex parameters have to be perceived, interpreted, integrated with prior knowledge and experience and correct decisions need to be inferred and communicated to the system. To fully understand how human operators interact with cyber-physical production systems this vast factor space must be described. Future research must therefore identify, quantify and weigh the components and their interactions.

The developed research framework can be the basis for this future research as it enables a thorough and systematic investigation of possible influencing factors and their interactions in the context of socio-technical cyber-physical production systems [15]. In addition to investigating the influence of user factors on decision-making, the research agenda focuses the development of suitable and user-adaptive decision support systems for both the inter-company flow of materials and information and the underlying production systems. In particular, trust in decision support systems, (blind) compliance and the restoration of trust in automation after a failure in different risk contexts are research topics to be addressed in order to bridge the gap between the human decision maker and the increasingly complex socio-technical cyber-physical production systems.

Acknowledgements. The work is funded by the German Research Foundation (DFG) as part of the German Excellence Initiative and the Cluster of Excellence "Integrative Production Technology for High Wage Countries" (EXC 128). We thank all participants for their willingness to support our study and Lara Mhetawi, Katharina Merkel, Fabian Comanns, and Wiktoria Wilkowska for research support.

References

1. Evans, P.C., Annunziata, M.: Industrial Internet: Pushing the Boundaries of Minds and Machines. Technical report, General Electric (2012)
2. Bruner, J.: Industrial Internet - The machines are talking. Reilly Media, Inc. (2013)
3. Bainbridge, L.: Ironies of Automation. Automatica 19(6) (1983) 775–779
4. Schlick, C., Stich, V., Schmitt, R., Schuh, G., Ziefle, M., Brecher, C., Blum, M., Mertens, A., Faber, M., Kuz, S., Petruck, H., Fuhrmann, M., Luckert, M., Brambring, F., Reuter, C., Hering, N., Groten, M., Korall, S., Pause, D., Brauner, P., Herfs, W., Odenbusch, M., Wein, S., Stiller, S., Berthold, M.: Cognition-Enhanced, Self-optimizing Production Networks. In Brecher, C., Özdemir, D., eds.: Integrative Production Technology - Theory and Applications. Springer International Publishing, Berlin, Heidelberg (2017) 645–743
5. Nielsen, J.: Usability Engineering. Morgan Kaufmann Publishers (1994)
6. Gould, J.D., Lewis, C.: Designing for Usability: Key Principles and what Designers Think. Communications of the ACM 28 (1985) 300–311
7. Knight, F.H.: Risk, Uncertainty and Profit. Houghton Mifflin Co, Boston, MA (1921)
8. Tversky, A., Kahneman, D.: Availability: A heuristic for judging frequency and probability. Cognitive Psychology 5(2) (1973) 207–232
9. Christopher, M., Peck, H.: Building the Resilient Supply Chain. The International Journal of Logistics Management 15(2) (2004) 1–14
10. Gilovich, T., Griffi, D., Kahneman, D.: Heuristics and Biases: The Psychology of Intuitive Judgment. 8th edn. Cambridge University Press (2002)
11. Brauner, P., Ziefle, M.: Game-Based Learning in Manufacturing and Business. In: 6th International Conference on Competitive Manufacturing (COMA 2016). (2016)
12. Beierlein, C., Kovaleva, A., Kemper, C.J., Rammstedt, B.: Eine Single-Item-Skala zur Erfassung von Risikobereitschaft: Die Kurzsskala Risikobereitschaft-1 (R-1) [A Single Item scale for measuring risk-taking attitude]. (2014)
13. Schuler, H., Prochaska, M.: Leistungsmotivationsinventar [Need for Achievement Inventory]. Hogrefe (2001)
14. Baxter, G., Rooksby, J., Wang, Y., Khajeh-Hosseini, A.: The ironies of automation – Still going strong at 30? In: Proceedings of ECCE 2012 Conference, 29th-31st August, Edinburgh. (2012) 65–71
15. Brauner, P., Valdez, A.C., Philipsen, R., Ziefle, M.: On Studying Human Factors in Complex Cyber-Physical Systems. In: Workshop Human Factors in Information Visualization and Decision Support Systems Held as Part of the Mensch und Computer 2016, Gesellschaft für Informatik (2016)