Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira, Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria

doi: 10.3850/ESREL2026061419_esrel26-p26505-cd

A Framework for Monitoring Group Resilience to Enhance Decision-Making Under Uncertainty

Fabian Fritsch

Department of Mechanical Engineering, Work and Organizational Psychology, University of Kassel, Germany.
E-mail: f.fritsch@uni-kassel.de

Oliver Straeter

Department of Mechanical Engineering, Work and Organizational Psychology, University of Kassel, Germany.
E-mail: straeter@uni-kassel.de

Incidents and accidents in various industrial domains have repeatedly revealed the critical role of the human factor alongside the reliability of technical products and systems. For this purpose, Human Reliability Analysis (HRA) methods have been developed, allowing statements about the reliability of human actions. For example, the reliability of an operator in the control room of a nuclear power plant can be assessed, and corresponding measures can be derived to enhance human reliability.
However, these approaches often fall short when applied to activities in which clearly describable human actions cannot be identified. This applies in particular to tasks carried out in research and development projects or to decision-making under incomplete information, such as during early phases of innovation or industrial transformation processes, that is, decision-making under uncertainty. For these early phases of system development, such as conceptualization or planning activities, suitable methods are required to adequately assess human reliability. Approaches relying solely on project management techniques are insufficient, as they typically neglect essential aspects of human information processing.
This paper presents a group-based assessment framework for measuring and evaluating uncertainties related to human factors as an influence on reliability in projects. The method primarily captures cognitive and group-dynamic aspects and provides an additional means of quality assurance by integrating the evaluation into classical or agile project management contexts. The results indicate a reduction of individual cognitive biases over time, while group-dynamic biases tend to increase, highlighting that group-based decision-making does not automatically lead to higher reliability. This approach strengthens not only the development process but also the reliability and safety of the resulting products and systems, enabling a more holistic consideration of the human factor.

Keywords: Human reliability analysis, decision-making, group assessment.

1. Introduction and Objectives

The design of work is subject to continuous change. Drivers such as digitalization, globalization, and the increasing use of artificial intelligence confront organizations with fundamental transformations in ever shorter cycles. For a structured description and socio-technical analysis of such transformation processes, the Human-Technology-Organization (HTO) approach proposed by Strohm and Ulich (1997) can be applied (also known as Man-Technology-Organization). Figure 1 illustrates this.

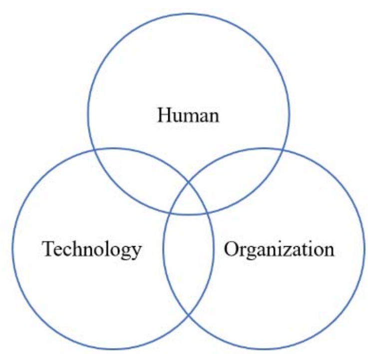

Fig. 1. Human-Technology-Organization (HTO) approach.

Using artificial intelligence as an example, rapid developments on the technical dimension can be observed, accompanied by substantial effects on organizational processes and on interactions between humans and technology.

From the perspective of work design, this creates the need to identify and model challenges at the interfaces of the HTO dimensions in order to avoid corrective approaches to work design (Ulich, 2011). Beyond artificial intelligence, transformation projects are typically characterized by a high degree of novelty, an open outcome, and dynamically changing influencing factors. That organizations often struggle to manage such projects successfully is reflected in recent empirical findings on transformation and digitalization initiatives (Oludapo et al., 2024; Ramesh & Delen, 2021).

At the same time, the human dimension, understood as decision-makers, project managers, and other involved actors, represents a critical component with inherent and largely stable characteristics, particularly with regard to cognitive information processing. In transformation projects, evaluations, trade-offs, and decisions must be made under substantial uncertainty. With respect to human information processing, it is well established that rational judgement cannot be assumed under such conditions (Tversky & Kahneman, 1974). Instead, systematic distortions in judgement, commonly referred to as cognitive biases, may dominate decision-making processes and should therefore be addressed as part of quality assurance in projects and as an integral component of Human Reliability Analysis (HRA).

Examples include confirmation bias, describing the tendency to selectively search for and interpret information in line with an initial assumption or hypothesis, as well as groupthink, where the striving for unanimity in highly cohesive groups may override rational decision-making (Evans, 1990; Janis, 1971). To systematically capture such distortions, a quantitative assessment method for human reliability was developed and validated.

This study demonstrates how this assessment method can be applied as a monitoring instrument in project groups and which results can be obtained through its use. In contrast to the implicit assumption that combining individual judgements within groups inevitably increases reliability, this article explicitly looks at both individual cognitive and group-dynamic factors that can lead to reliability issues. The objective is to provide a structured framework for assessing human reliability in early and uncertain phases of transformation projects, thereby extending existing Human Reliability Analysis approaches to contexts in which clearly observable human actions are not available.

2. Evaluating transformation projects using AHRIC

To enable a systematic and comparable assessment of biases, the questionnaire-based method AHRIC (Assessment of Human Reliability in Concept Phases) was developed. Based on self-reports, the method allows for an efficient and investigator-independent data collection using a six-point Likert scale. Respondents indicate their degree of agreement with each item, from which a susceptibility to a specific bias is inferred. This represents a simplified assessment rather than a diagnostic measurement of biases. All items are theoretically derived from established concepts in cognitive psychology and represent observable indicators of bias-related behavior in group settings.

The target group of the instrument comprises individuals involved in the design and implementation of projects characterized by novelty and uncertainty, as typically encountered in transformation contexts, such as decision-makers and project team members. Table 1 presents the assessment procedure based on a short version of the questionnaire (AHRIC-S), which was derived from a longer version of the AHRIC instrument (Fritsch et al., 2025). A central objective of AHRIC-S is high practical applicability and efficiency, enabling repeated and accompanying measurements within project settings.

The instrument is universally applicable across projects and does not require project-specific adaptation. Examples include projects focused on the transformation of mobility as well as environmental projects, such as the safe final disposal of high-level radioactive waste. Given the numerous dependencies and interactions that are typical in such contexts, the use of accompanying safety instruments is of particular importance. The following section describes the study design and statistical approach used to evaluate changes in bias susceptibility across time and experimental conditions.

Table 1. Overview of the items used in AHRIC-S.

Category	Item
Anchoring bias	The initial information had a lasting influence on the outcomes of the meeting. (reverse coded)
Confirmation bias	At the beginning of the meeting, participants already had a specific idea of the results. (reverse coded)
Overconfidence bias	As an expert in my field, I am completely certain that my views and assessments are correct. (reverse coded)
Law of the instrument	In the meeting, familiar solution approaches were predominantly pursued instead of considering new perspectives. (reverse coded)
Group polarization	In our discussion, an existing consensus was reinforced by many additional arguments. (reverse coded)
Group polarization	Within our meeting group, we have very similar views and opinions on the topics discussed. (reverse coded)
Groupthink	The leader in the meeting expressed clear expectations and enforced them. (reverse coded)
Groupthink	Due to the heterogeneous composition of participants, very different views or opinions were discussed in the meeting.
Management objectives	Management imposed requirements and objectives on our meeting that had to be achieved despite substantive contradictions. (reverse coded)

3. Study Design

For the present study, the AHRIC-S assessment procedure was applied as an accompanying measurement instrument in project groups. The objective of these groups was to apply theoretical concepts from work and organizational psychology to a problem context related to the digitalization of public administration. Each group developed conceptual solutions over a period of approximately three months, including analysis, ideation, and evaluation phases. The tasks were deliberately ill-defined to simulate decision-making under uncertainty. The degree of control, autonomy, novelty, and complexity of the task corresponded to the classical definition of project groups (Schaper, 2019). Data collection took place in 2025 at three measurement points:

- T_1: at the beginning of task processing within the project groups
- T_2: intermediate measurement after five weeks
- T_3: final measurement at the last project meeting.

In addition, two different questionnaire conditions were implemented. In one condition, the questionnaire served solely as a measurement instrument and was administered after the group meetings using only the items shown in Table 1. In the second condition, participants were provided with additional prompts after responding to and reviewing the items under specific conditions. These prompts consisted of behavioral instructions aimed at reducing the respective bias in future group work. The prompts were displayed automatically when a predefined response pattern was detected, with the threshold defined as strong or complete agreement with a specific bias category (e.g., confirmation bias). The prompts consisted of short, bias-specific behavioral recommendations aimed at improving decision-making processes in subsequent group interactions. For example, in the case of confirmation bias, participants were encouraged to actively seek disconfirming evidence and consider alternative hypotheses.

As no personal identifiers were collected, the measurement points were treated as independent samples. Due to dropouts, sample sizes varied across measurement points (T_1: $N_1 = 42$, T_2: $N_2 = 24$, T_3: $N_3 = 47$). Consequently, the analyses describe mean differences between measurement points at the group level rather than intra-individual changes.

For data analysis, items were aggregated into two categories: individual cognitive biases (e.g., anchoring, confirmation bias, overconfidence bias) and group-dynamic biases (e.g., groupthink, group polarization). Mean scores were calculated for each category to obtain an overall measure of susceptibility. For data analyses, a 3 × 2 ANOVA (Analysis of Variance) for independent samples was then conducted. The independent variables were measurement time (three levels: T_1, T_2, T_3) and questionnaire condition regarding behavioral prompts (two levels: with vs. without prompts). The dependent variable was the level of human reliability, which was operationalized via agreement with bias-related items. ANOVAs were calculated for the overall score (mean across all items), for individual cognitive and group-dynamic bias dimensions separately, and for each of the nine items individually. For significant results ($\alpha = .05$), Tukey post-hoc tests and descriptive analyses of mean values at each measurement point were conducted, including the calculation of partial eta squared as an effect size measure.

4. Results

Across all items, no significant difference was found between the two questionnaire conditions (with versus without behavioral prompts), ($F(1, 107) = 0.10$, $p = .75$). With respect to measurement time, no significant difference was observed for the overall score across all items, ($F(2, 107) = 2.07$, $p = .13$).

When individual cognitive and group-dynamic biases were analyzed separately, a significant main effect of measurement time emerged for individual cognitive biases ($F(2, 107) = 9.39$, $p < .001$, $\eta_p^2 = .15$). Tukey post-hoc tests indicated a significant decrease between T_1 and T_3. Descriptively, this corresponds to a reduction of 17.7% from T_1 to T_3.

At the group-dynamic level, no significant main effect of measurement time was observed, ($F(2, 107) = 0.83$, $p = .44$). Descriptively, an increase of 8% between T_1 and T_3 was found. Mean values for individual cognitive biases and group-dynamic biases across measurement points are reported separately in Table 2.

Table 2. Means across measurement points for individual cognitive biases and group-dynamic biases.

Measurement point	Individual cognitive biases		Group-dynamic biases	
	M	*SD*	*M*	*SD*
T_1	2.94	0.56	2.21	0.55
T_2	2.53	0.61	2.31	0.57
T_3	2.42	0.57	2.39	0.56

Note. *M* = mean; *SD* = standard deviation. Measurement points were treated as independent samples.

5. Discussion

The aim of this study was to introduce a structured assessment framework for transformation projects and to demonstrate its practical application. The results provide initial empirical indications regarding individual cognitive and group-dynamic biases. No effect of behavioral prompts was observed. One possible explanation is that prompts were only displayed when a certain level of agreement with a bias was reached, meaning that some participants in this condition did not receive any prompts at all.

Furthermore, the results indicate a reduction in individual cognitive biases over time. This corresponds to a decrease of approximately 0.5 points on the six-point Likert scale, indicating a moderate practical effect. This interpretation is consistent with the observed effect size ($\eta_p^2 = .15$), which can be interpreted as a medium effect.

In contrast, the results indicate an increase in group-dynamic biases. Although this increase was not statistically significant, it reflects a small upward shift in mean values, suggesting a tendency toward stronger group-dynamic effects over time. These findings are consistent with existing assumptions. Simply answering questions about a specific behavior can already influence that behavior, a phenomenon known as the question–behavior effect (Rodrigues et al., 2015). In addition, the increase of polarizing effects in groups over time represents a well-documented pattern in social psychology (Hewstone & Martin, 2023).

From a human reliability perspective, these findings underline a critical point: group-based decision-making does not inherently result in higher reliability than individual judgement.

Figure 2 presents the biases identified with the applied method and models the study's findings in an integrated representation, following a classical HRA approach. The three phases represent the three measurement points (T_1, T_2, T_3). At T_1, corresponding to the initial phase of problem definition and requirement analysis, three out of the seven examined biases exceeded the defined threshold. Based on this ratio, a probability of success of $p(\text{success}) = 0.429$ can be derived. Here, success is defined as the absence of elevated bias levels, assuming an equal contribution of each bias to overall reliability.

Extending this calculation to the remaining measurement points (T_2 and T_3) results in the probabilities illustrated on the right-hand side of Figure 2. While individual cognitive biases may decrease through reflection, experience, or repeated measurement, groups introduce their own failure modes, such as groupthink or polarization. The aggregation of individual judgements therefore does not automatically compensate the limitations of human information processing, but can instead shift the issues of human reliability from the individual to the collective level.

This has direct implications for decision-making under uncertainty. In early phases of transformation projects, where objective feedback is limited and outcomes are open, groups may develop a false sense of reliability based on consensus or shared confidence. Monitoring group-dynamic biases is therefore essential to avoid overestimating the robustness and resilience of collective decisions.

Overall, the presented assessment framework using AHRIC provides a structured approach for measuring and evaluating psychological influences in project settings. Due to its self-assessment-based and human-centered design, project groups, semi-autonomous work groups, or decision-making bodies such as steering committees can independently evaluate their susceptibility to biases. This supports reliable transformation design at the project

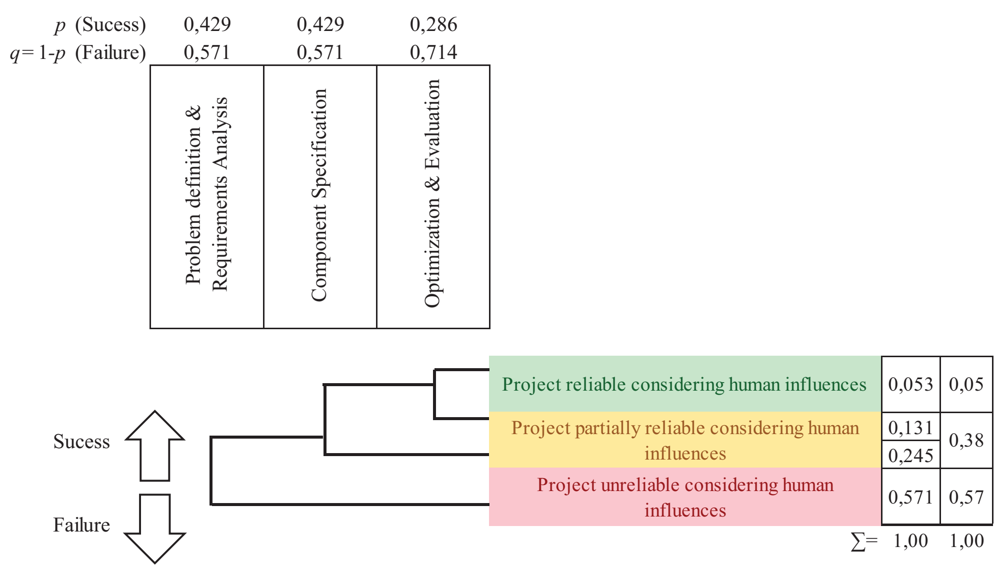

Fig. 2. Event sequence diagram with study results using AHRIC-S

level. At the same time, the efficiency of the instrument, with an application time of approximately three minutes, allows for implementation at the organizational level, enabling continuous monitoring of human reliability in transformation processes.

Within the framework of the HTO approach, this study primarily addresses the interface between humans and organizations, focusing on cognitive and group-dynamic biases in the organizational project context. In addition, Straeter (2026) demonstrates the interplay between humans and technology from the perspective of HRA. Together, these contributions provide a more comprehensive perspective on human reliability in socio-technical systems.

References

Evans, J. S. B. T. (1990). Bias in Human Reasoning: Causes and Consequences. Psychology Press.

Fritsch, F., Sträter, O., Röhlig, K.-J., & Heiermann, M. (2025). ENSURE „Endlagersicherheit: Ungewissheiten und Regulatorische Aspekte“. KOBRA Universität Kassel. https://doi.org/10.17170/kobra-2025081211387

Hewstone, M., & Martin, R. (2023). Sozialer Einfluss. In Sozialpsychologie (S. 265–309). Springer.

Janis, I. (1971). Groupthink Among Policy Makers. In N. Sanford & C. Comstock (Hrsg.), Sanctions for Evil (S. 71–89). Boston: Beacon Press.

Oludapo, S., Carroll, N., & Helfert, M. (2024). Why do so many digital transformations fail? A bibliometric analysis and future research agenda. Journal of Business Research, 174, 114528. https://doi.org/10.1016/j.jbusres.2024.114528

Ramesh, N., & Delen, D. (2021). Digital Transformation: How to Beat the 90% Failure Rate? IEEE Engineering Management Review, 49(3), 22–25. https://doi.org/10.1109/EMR.2021.3070139

Rodrigues, A. M., O'Brien, N., French, D. P., Glidewell, L., & Sniehotta, F. F. (2015). The question–behavior effect: Genuine effect or spurious phenomenon? A systematic review of randomized controlled trials with meta-analyses. Health Psychology, 34(1), 61–78. https://doi.org/10.1037/hea0000104

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. Science, 185(4157), 1124–1131.

Schaper, N. (2019). Gruppenarbeit in der Produktion. In: Arbeits- und Organisationspsychologie. Springer-Lehrbuch. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-56666-4_23

Straeter, O. (2026). The importance of Human Reliability for assessing the Risk of Artificial-Intelligence-Systems according to the EU AI-Act. 36th European Safety and Reliability Conference (ESREL 2026).

Strohm, O., & Ulich, E. (Eds.). (1997). Unternehmen arbeitspsychologisch bewerten: Ein Mehr-Ebenen-Ansatz unter besonderer Berücksichtigung von Mensch, Technik und Organisation (Vol. 10). vdf Hochschulverlag AG.

Ulich, E. (2011). Arbeitspsychologie. vdf Hochschulverlag AG.