Promoting thinking in terms of causal structures: Impact on performance in solving complex problems

Franziska Kessler[1]
Antje Proske[1]
Micah B. Goldwater[2]
Leon Urbas[3]
Samuel Greiff[4]
Susanne Narciss[1]

[1]Department of Psychology, Universität Dresden, Germany
[2]School of Psychology, University of Sydney, Australia
[3]Department of Electrical Engineering, Universität Dresden, Germany
[4]Department of Behavioural and Cognitive Sciences, University of Luxembourg, Luxembourg

Abstract

Goldwater and Gentner (2015) showed that the sensitivity for causal structures can be promoted with an intervention combining explication of causal models and guided structural alignment of situations from disparate fields with the same underlying causal model. We extended this intervention with inference questions and combined it with a subsequent complex problem-solving (CPS) task, in order to investigate whether enhanced sensitivity for causal structures results in better performance in CPS. This study (N = 108) compares the CPS performance indicators knowledge acquisition and knowledge application among three experimental groups (intervention, intervention extended with inference questions, control group) and reveals the following results: 1) The effectiveness of the intervention in increasing the sensitivity for causal structures was replicated. 2) Sensitivity for causal structures and CPS performance indicators were significantly positively correlated. 3) There is no direct effect of the intervention on CPS performance, but an indirect-only effect via enhanced sensitivity.

Keywords: relational categorization, analogical transfer, education, complex problem-solving

Introduction

The emergence of new technologies and the accompanying increased levels of automation have shifted the tasks performed by humans in many workplaces. We see an evolving emphasis of non-routine, dispositive tasks that require active problem solving skills and the ability to draw upon previously acquired knowledge to deal with new situations (Hirsch-Kreinsen & ten Hompel, 2017). One example of such core skills critical to the 21st century workplace is complex problem solving (CPS, Neubert, Mainert, Kretzschmar, & Greiff, 2015). CPS refers to activities necessary for dealing with dynamic non-routine situations and handling of complex systems, i.e. systems composed of many interconnected elements, dynamics and intransparent causal relations. CPS requires the active acquisition and application of knowledge about the system's structure and dynamics (Fischer et al., 2012; Funke, 2001). In order to meet the requirements of modern workplaces, building flexible knowledge that may be applied to novel situations becomes increasingly important in human operator training.

The application of prior knowledge to a new situation is referred to as knowledge transfer. Analogical reasoning constitutes a key process for knowledge transfer (e.g., Gentner, 1983). It involves structurally mapping of a previously encountered situation (i.e., source) to a novel, superficially distinct but structurally equivalent situation (i.e., target) (e.g., Holyoak, 2012). By matching the corresponding elements of the situations, conjectures on future developments of the target may be drawn based on knowledge about the source. A prerequisite for analogical reasoning is the ability to detect the underlying causal structure of the current situation in order to activate a suitable source analogue from memory.

Increased sensitivity for relational, or more specifically, key causal structures in the world, has been shown to develop with increasing expertise. While novices tend to concentrate on salient superficial characteristics, experts have been found to focus on causal structures. This shift of early focus on concrete features to a later focus on relational structures (e.g. causal structure) is referred to as relational shift. It has been reported for cognitive development in general (Rattermann & Gentner, 1998; Richland et al., 2006), but more importantly for our endeavor, it has been shown to accompany development of expertise in different domains. Chi et.al (1981) showed that physics experts were more likely to sort physics problems according to underlying principles, whereas novices tended to sort by superficial feature-based similarity. Shafto and Coley (2003) reported that professional fisherman grouped fish according to behavioral and causal relations, while novices organized them according to similarity in appearance. Rottman, Gentner and Goldwater (2012) designed the Ambiguous Card Sorting Task (ACST), which requires participants to cluster different descriptions of real-world phenomena. The phenomena vary in their underlying causal structures (e.g., common cause or common effect) and their scientific content domain (e.g., biology or environmental sciences). Hence, the sorting procedure can either be performed according to the content domain or according to underlying causal structures of the phenomena. In line with research on relational shift, physical sciences experts were more likely to spontaneously detect and sort phenomena from different physical sciences by causal patterns in the ACST, while novices mainly concentrated on surface features, and hence sorted by the more salient dimension of content domain. Further, experts in environmental sciences showed a higher tendency to detect the key causal models in other domains such as economics, indicating generalized sensitivity beyond their domain of expertise. This tendency was attributed to the fact that experts had acquired abstract mental representations of these causal models. Formation of abstract mental representations constitutes the foundation to draw inferences across domains and facilitates spontaneous transfer from a source to phenomena from disparate fields (e.g., Gentner et al., 2003). This is because for situations that require the same solution strategy, the relational content such as underlying causal models is likely to remain constant across domains, whereas superficial features are likely to lack similarity and may even be misleading.

Goldwater and Gentner (2015) investigated what it takes to acquire such abstract mental representations on causal structures in order to perceive causal models across multiple contexts. In their study, they focused on high-level abstracted key causal models that are prevalent in the world (i.e., common cause, common effect, chain, positive feedback). They tested the effect of different learning experiences and found a combination of explication of these causal models and guided structural alignment to be most effective in enhancing the ability to notice the key causal models in novel situations from different fields. The explication phase ensured that participants thoroughly understood the included key causal models. A thorough understanding of the key causal models increased the effectiveness of subsequent guided pair-wise comparisons of analogs, i.e. structural alignment. Analogical comparison promotes the formation of abstract mental representations, i.e., schema abstraction (Gentner, 2010). The structural alignment renders the common underlying causal structure of the analogs more salient, thus fostering schema abstraction which in turn promotes transfer (Gick & Holyoak, 1983; Markman & Gentner, 1993; Reeves & Weisberg, 1994). The effectiveness of analogical comparisons in facilitating learning and transfer of relational content has found considerable evidence in prior studies and has reached “gold-standard“ status for promoting spontaneous transfer (see Alfieri et al., 2013, for review).

Participants who accomplished both tasks, explication and structural alignment, were more likely to sort situation descriptions according to common underlying causal structures, rather than by more salient superficial characteristics in the ambiguous card sorting task (ACST). This relational shift in category understanding was interpreted to be a result of the acquisition of abstract mental representations of the key causal models. The causal systems studied in the intervention can be conceptualized as relational categories. In relational categories, membership is determined by common underlying causal structures such as the key causal models addressed in the intervention. In contrast, feature-based categories are defined in terms of directly observable properties, hence instances usually have high perceivable similarity in feature-based categories (Gentner & Kurtz, 2005).

Relational knowledge is fundamental for many higher cognitive processes (Halford et al., 2010). Understanding of (causal) relations between elements in the world constitutes the basis of mental models which are used as descriptions of the mental representation of phenomena (Vosniadou, 1994) and form the foundation for predictions and inferences about situations (Kokkonen, 2017). Given the importance of relational knowledge, increased sensitivity for key causal structures is likely to have an impact on CPS performance. Key steps in CPS are the acquisition of knowledge of the system’s structure by systematically exploring the system’s behavior and relational links, building an appropriate mental model of the relations between system components and using this representation and knowledge for successfully controlling a complex system (Dörner, 1986). All of those steps may benefit from heightened understanding and increased sensitivity for causal structures.

We want to shed light on the effect that enhanced sensitivity for key causal structures exerts on the process of exploring and controlling a complex problem or system.

The current study

The aim of the current study is to combine the intervention of Goldwater and Gentner (2015) with a subsequent CPS task, in order to investigate whether enhanced sensitivity for key causal structures results in better performance in solving a complex problem. We hypothesize to 1) replicate the effectiveness of the intervention consisting of explication of causal models and guided structural alignment in enhancing sensitivity for causal structures, 2) show a positive relationship between the level of sensitivity for causal structures and CPS performance, 3) demonstrate better performance in CPS for participants who received the intervention. We additionally included a second intervention group (IG 2) that received the same intervention followed by an additional inference question task. The task was designed to induce a mental simulation process of the key causal models. The idea was that detecting the key causal models, i.e. sensitivity, may aid during the knowledge acquisition phase in CPS, but for successful controlling a dynamic system, it is likely that people benefit from a more profound, in-depth understanding of the implications of the relational patterns between elements. Inclusion of this experimental task, permits discerning the supplementary impact of mental simulation processes beyond enhanced sensitivity for those structures.

Methods

We recruited 108 participants (females = 62; males = 45; diverse = 1) from the psychology experiment pool of the TU Dresden. Participants were full-aged (mean = 25.29 years) and had at least C1 level of German language proficiency which corresponds to the ability to use the language proficiently. Participation in the study was compensated with 10€, 15€, or 20€ or 1h, 1.5h, or 2h class credit for participating in the study.

Design, Materials and Procedure

Participants were randomly assigned to one of three conditions in a between-subjects design: Intervention group 1 (IG1) that received the intervention, intervention group 2 (IG2) that received the intervention and the inference questions task, and a control group (CG) that did not receive any sort of intervention and served as a baseline measurement of the sensitivity for causal structures and baseline CPS performance. Participants in the CG started with the ACST and proceeded with the CPS task. The participants in the intervention groups received the intervention first (IG1 & IG2), followed by the inference question task (only IG2), before performing the ACST, and finally the CPS task. The causal structures that were addressed in the intervention (i.e., common effect, common cause, self-enhancing system, and chain) are identical to the ones in the ACST and the CPS task.

Intervention We used an adapted version of the intervention by Goldwater and Gentner (2015) that was designed to efficiently support the acquisition of abstract mental representations of the addressed causal structures. Because it was impossible to implement the positive feedback system from the original intervention into the CPS task, we changed the positive feedback system to a self-enhancing system in the intervention to make it compatible with the requirements in the CPS task.

The explication part of the intervention comprised eight situation descriptions differing in their underlying causal structure and domain (four causal structures from the domain of electrical engineering, four from history, see Figure1). Each situation description was followed by the label of the according underlying causal structure and an explication on how the causal structure fits the particular example situation. Participants were asked to draw a causal diagram for each of the eight situation descriptions.

For the structural alignment task, participants were asked to compare pairs of example descriptions from the explication phase that differed in their content domain (history vs. electrical engineering), but shared the same underlying causal structure. For each of these pairs the structural alignment was guided by providing participants with a two-column table listing the elements of one example description in the first column. Participants were asked to fill in the second column with the corresponding element of the other example, i.e. to structurally align the two examples. Participants did not receive any feedback on the correctness of their causal diagrams or guided structural alignment.

		Common effect	Common cause	Self-enhancing system	Chain
Intervention	Electrical Engineering				
	History				
Assessment (ACST)	Environmental Science	Seed			
	Economics		Seed		
	Mechanical Engineering			Seed	
	Biology				Seed

Figure 1: Matrix of materials used in this study

Inference questions Participants in the IG2 were presented with eight additional phenomena descriptions from different content domains (two examples for each of the four causal structures). For each description, participants were asked to state which of the four causal models applied and received corrective feedback on their response. Subsequently, they had to answer one multiple choice question for each description. The multiple choice questions were designed to elicit mental simulations of the implications (i.e., inferences) of the respective key causal model (see Figure 2).

Situation description:
With advances in knowledge, especially in areas such as medicine and technology, there has ultimately been a reduction in disease and sickness and therefore an increase in the average life expectancy. This means that the population is not dying off as quickly as previously and so the population grows at a higher rate.

Corrective feedback:
This is an example of a chain.

MC question:
If there is a reduction in disease, how would this impact on the development of technology?

a) This would increase the development of technology because disease prevention directly affects the development of technology.
b) This would not affect the development of the technology, because the technology directly impacts disease, not the other way around.
c) The advancement of knowledge and the development of new technology would be reduced, because disease prevention directly affects the technology in a positive way.

Figure 2: Example of one phenomena description and MC question from the inference question task

Ambiguous Card Sorting Task (ACST) We used an adapted version of the ACST by Rottman et al., (2012) as a measure for sensitivity for causal structures. In order to match it with the causal models in the intervention, we modified the structure of a positive feedback system to the structure of a self-enhancing system as we did for the intervention. The task required participants to cluster descriptions of phenomena. Each description differs in content domain and in underlying causal model and hence allowed for a domain or causal sorting strategy (see Figure 1). Each phenomena description was printed on a file card. Four cards were chosen as “seed cards” which served as the column “header” for the sorting task. The seed cards and an additional card with the label “other” were laid on the table. Participants were instructed to read the seed cards and then sort the remaining 12 cards to one column each which they felt was the best match (see Table 2). The number of cards sorted according to the underlying causal model served as a measure for sensitivity for causal structures. It indicates how frequently participants detected the causal structure and granted more weight to this information than to the more salient domain characteristics.

Seed card (biology, chain): The firing of a neuron is the product of a series of successful steps; electro/chemical propagation, synthesis of the neurotransmitter in the presynaptic terminal, release, binding, and lastly reuptake. The inhibition or increased stimulation of any one of these steps can affect the successful firing of a neuron.	
Matching domain (biology, common effect): Aside from genetics, the four main factors which increase the risk of heart disease are smoking, high blood cholesterol, high blood pressure, and obesity. Doctors advise people with any one, or a combination of these risk factors to improve their health to mitigate the risk of heart attack.	**Matching causal system (chain, environmental sciences):** As a greenhouse gas, CO2 traps heat from sunlight and keeps it inside the earth's atmosphere. Thus, the level of CO2 in the atmosphere has a direct effect on the temperature of the earth. The more CO2 in the atmosphere, the more heat the atmosphere traps and the hotter the earth becomes.

Figure 3: Example of one seed card and a card with domain match and a card with causal system match, taken from the ACST (Rottman et al., 2012)

CPS performance As a measure for CPS, we used the MicroDYN framework (Greiff et al., 2012) which is a well-established measurement tool for CPS with good psychometric characteristics (e.g., Greiff & Wüstenberg, 2014). MicroDYN enables the construction of computer-based microworlds built on linear structural equation systems with input and output variables and opaque relations between them. Causal relations are defined by the experimenter and can exist between input and output variables, between input and input variable, and in form of self-enhancing effects on output variables that add a dynamic character to the microworld (Greiff & Wüstenberg, 2014). We used the 10 items of the standard item battery and kept the semantic embedding of the original items, but altered their internal causal structure in order to match it to the key causal models addressed in the intervention.

Solving a MicroDYN item involves two successive tasks: 1) Exploration of the underlying causal structure of the microworld and 2) controlling of the microworld in order to reach indicated goal states. During the exploration phase, participants can manipulate the input variables. By observing the induced changes in the output variables, they can infer the underlying causal relations within the system and record their assumptions in form of a causal model. Correctness of the model serves as a measure for the CPS indicator *knowledge acquisition* with a dichotomous scoring for each microworld (incorrect: 0, correct: 1). In the control phase, participants are shown the correct underlying causal model and have to reach specified values in the output variables in a maximum of four steps. Success in reaching the indicated goal states serves as a measure for the CPS indicator *knowledge application* with a dichotomous scoring for each MicroDYN item (incorrect: 0, correct: 1). The two performance indicators (*knowledge acquisition*, *knowledge application*) were averaged across the ten items for each participant for further analyses.

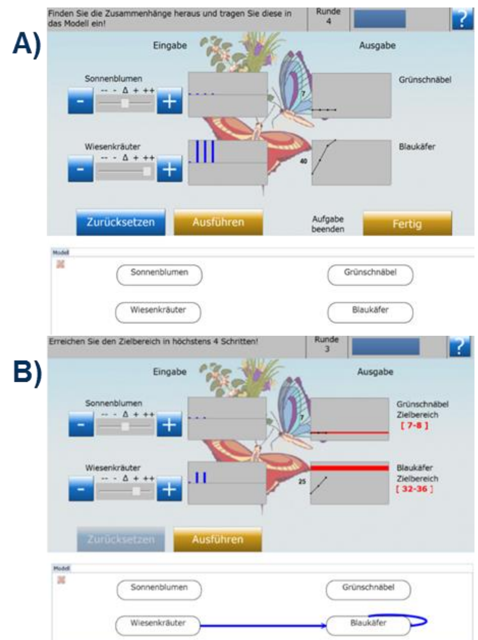

Figure 4: Screenshot of the MicroDYN microworld "Butterflies". The knowledge acquisition phase is shown in panel A), the knowledge application phase in panel B).

Results

The ACST scores revealed that the adapted version of the intervention was effective in enhancing sensitivity for the key causal structures (see Figure 2) Comparing the number of causal sorts between the experimental groups, revealed a significant main effect $F(2, 105) = 27.80$, $p < .001$, $\eta^2 = .35$. Post-hoc tests (bonferroni-corrected) showed significantly higher levels of sensitivity in IG1 compared to the CG, $t = 6.26$, $p < .001$, $d = 1.47$, and IG2 compared to the CG, $t = 6.63$, $p < .001$, $d = 1.55$, but no differences between the two intervention groups $t = -.37$, $p = 1.0$, $d = -.09$ (see Figure 5).

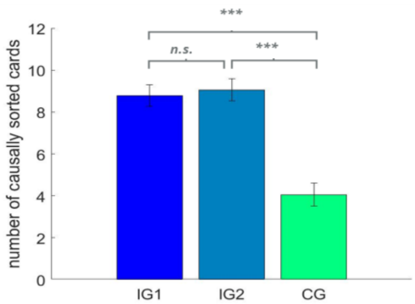

Figure 5: Causal structure sensitivity scores by groups

In order to discern the effectiveness of the intervention on an individual level, we assessed to what degree participants sorted according to causal model. We were particularly interested in the proportion of participants who adopted a mainly causal sorting strategy (i.e. causal sorter), because we assumed that the effects in CPS performance will become apparent only if the relational shift is strong. We set the cut-off for classification as a causal sorter to at least 9 (out of 12) causally sorted cards in the ACST. This cut-off entails that participants detected at least three out of four causal models reliably. About two thirds of the participants in the intervention groups were classified as causal sorters (IG1 = 63.9 %, IG2 = 66.7 %), compared to only 4 subjects (11%) in the CG. The difference in distribution between groups was statistically significant $\chi^2(2) = 28.31$, $p < .001$, Cramer's $V = .51$.

In order to test the assumption that sensitivity for causal structures is positively associated with CPS performance, we correlated the CPS indicators with the number of causal sorts in the ACST. Results showed significant positive correlations for both CPS indicators (*knowledge acquisition:* $r = .34$, $p < .001$; *knowledge application:* $r = .22$, $p < .05$). Further, we wanted to test whether those participants who adopted a mainly causal sorting style, i.e. causal sorters, performed better in the CPS task. A one-way MANOVA with the two CPS performance indicators as the dependent variables and sorting style as the independent variable revealed significantly better performance for causal sorters Wilk's $\Lambda = .88$, $F(2,105) = 6.88$, $p = .002$. Post-hoc ANOVAs showed significant main effects for both CPS performance indicators: *knowledge acquisition* $F(1,106) = 13.88$, $p < .001$, *knowledge application* $F(1,106) = 4.14$, $p < .05$ (see Figure 6).

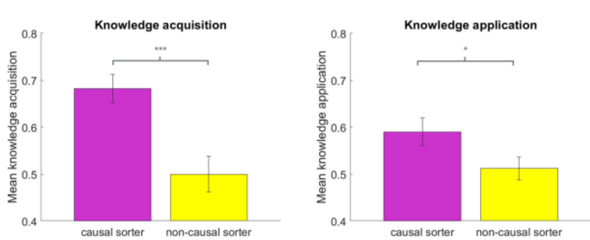

Figure 6: Differences between causal and non-causal sorters

To test the hypothesis that the intervention groups outperformed the control group, we conducted a one-way MANOVA with the CPS performance indicators as the dependent variables and experimental condition as the independent variable which revealed no effect of condition, Wilk's $\Lambda = .97$, $F(4,210) = .71$, $p = .59$ (see Figure 7).

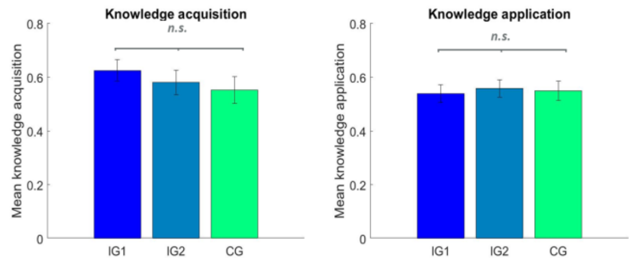

Figure 7: CPS performance scores by experimental group

This lack of difference between the experimental groups may be attributed to the fact that about two thirds of the participants responded to the intervention in the intended way, while others only partly shifted to a more causal sorting pattern. The rationale for an expected effect of condition was that the intervention increases the sensitivity for causal structures, which would then boost performance in CPS. To test whether this holds true, we conducted mediation analyses to examine whether the effect of the intervention on CPS performance is indirect and mediated by increased sensitivity for causal structures. According to the classification suggested by Zhao et.al (2010) these mediation analyses revealed indirect-only effects of intervention via enhanced sensitivity for causal structures on each CPS performance indicator. Bootstrapping with 5000 samples was employed to compute confidence intervals, indicating significant results for both indicators: *knowledge acquisition* = .144, 95%-CI[0.054-0.254], *knowledge application* = .084, 95%-CI[0.033-0.145] (see Figure 8).

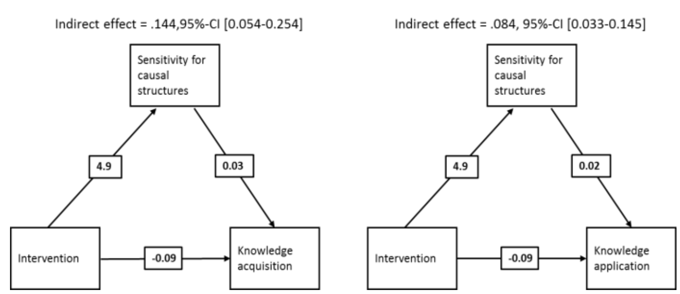

Figure 8: Mediation paths analyses

Discussion

Expertise is commonly accompanied by increased sensitivity for recurring causal structures in the world. This sensitivity is thought to be a result of having acquired abstract mental representations which facilitates their recognition across various contents and aids understanding and prediction of key phenomena. Intriguingly, Goldwater and Gentner (2015) found that causal structure sensitivity can be promoted through a short intervention combining (a) the explication of key causal models with (b) subsequent guided structural alignment. In order to investigate to what extent the enhanced sensitivity for causal structures can translate into a more expert-like approach of dealing with complex problem situations, we studied the effect of the intervention on subsequent CPS performance.

The key findings of this study include firstly the replication of the effectiveness of the intervention in enhancing sensitivity for the included key causal models. Secondly, we found that higher levels of sensitivity for causal models was associated with better performance in CPS, which held true for both CPS indicators *knowledge acquisition* and *knowledge application*. Participants who exerted a mainly causal sorting strategy significantly outperformed participants with a mixed or mainly domain sorting strategy in a CPS task. Thirdly, the intervention did not increase performance in CPS for all participants in the intervention groups. Instead, the effect was indirect-only and mediated by enhanced sensitivity for causal structures. Finally, the inference questions that were meant to induce the mental simulation of the causal structures had no added impact on CPS performance.

The results of this study add to the growing body of research highlighting the power of construction of relational categories in supporting transfer of knowledge across disparate fields. In essence, relational categories subsume analogue instances that share little superficial similarity but are based on the same causal structure. That is, relational categories take the underlying causal structure or principle as a taxonomy for classification which is often more informative for deciding upon an adequate solution strategy and hence aids transfer (Goldwater & Schalk, 2016). In this study, the phenomena descriptions in the ACST that are based on the same underling causal structures, can therefore be conceptualized as members of the four relational categories common cause, common effect, chain and self-enhancing system.

Organizing knowledge in relational categories, increases the propensity to classify new phenomena descriptions as instances of a particular relational category. In our study, the tendency to organize newly encountered situations either in terms of causal structures or in terms of superficial features, was measured with the ACST. A mainly causal sorting strategy employed in the ACST, i.e. organizing new situations in terms of relational categories, was in fact positively associated with CPS performance. According to the *category status hypothesis* the construction of relational categories may promote spontaneous transfer by paving the way for the activation of structurally similar cases from memory that are usually difficult to access by mere superficial, feature-based similarity (Kurtz & Honke, 2020; Snoddy & Kurtz, 2020). Further, perceiving the underlying causal structure and activation of the corresponding relational category, may not only ease the access of previously encountered analogous cases, but also allow for activation of previously successful problem solving strategy for this particular problem type. Based on these considerations and the findings of our current study, it would be worthwhile to combine the sensitivity intervention with training components addressing problem solving strategies for successful manipulation of each of the causal structures, and investigate if this increases the benefits in CPS performance.

To our knowledge, this is the first study that related sensitivity for particular causal structures with performance in CPS based on these particular structures. Though these are just initial results, they point towards a new and promising avenue for further research. Bridging of insights from cognitive psychology on relational categorization and the concept of CPS that taps into 21st century skills, highlights the scope of applicability and relevance of the concept of relational categorization. In educational settings, many of which seeking to impart flexible knowledge ready to be transferred, cross-domain comparisons and the set-up of relational categories should be encouraged, rather than isolated teaching of phenomena in separated content domains (Goldwater & Schalk, 2016). This may support adequate performance in future workplaces. Creating more opportunities and providing support for the successful acquisition of abstract mental representations and according relational categories, should be granted high priority.

Acknowledgements

This work was funded by "Deutsche Forschungsgemeinschaft" (DFG, German Research Foundation) under grant number 319919706/RTG2323 "Conducive Design of Cyber-Physical Production Systems".

References

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning Through Case Comparisons: A Meta-Analytic Review. *Educational Psychologist*, *48*(2), 87–113. https://doi.org/10.1080/00461520.2013.775712

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices*. *Cognitive Science*, *5*(2), 121–152. https://doi.org/10.1207/s15516709cog0502_2

Dörner, D. (1986). Diagnostik der operativen Intelligenz. *Diagnostica*, *32*, 290–308. https://psycnet.apa.org/record/1989-71230-001

Fischer, A., Greiff, S., & Funke, J. (2012). The Process of Solving Complex Problems. *The Journal of Problem Solving*, *4*(1). https://doi.org/10.7771/1932-6246.1118

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking & Reasoning*, *7*(1), 69–89. https://doi.org/10.1080/13546780042000046

Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy. *Cognitve Science*, *7*(2), 155–170. https://doi.org/10.1016/S0364-0213(83)80009-3

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science*, *34*(5),

752–775. https://doi.org/10.1111/j.1551-6709.2010.01114.x
Gentner, D., & Kurtz, K. J. (2005). Relational categories. In *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin.* (pp. 151–175). https://doi.org/10.1037/11156-009
Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, *95*(2), 393–408. https://doi.org/10.1037/0022-0663.95.2.393
Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38. https://doi.org/10.1016/0010-0285(83)90002-6
Goldwater, M. B., & Gentner, D. (2015). On the acquisition of abstract knowledge: Structural alignment and explication in learning causal system categories. *Cognition*, *137*, 137–153. https://doi.org/10.1016/j.cognition.2014.12.001
Goldwater, M. B., & Schalk, L. (2016). Relational categories as a bridge between cognitive and educational research. *Psychological Bulletin*, *142*(7), 729–757. https://doi.org/10.1037/bul0000043
Greiff, S., & Wüstenberg, S. (2014). Assessment with microworlds using MicroDYN: Measurement invariance and latent mean comparisons - Psychometric properties across several student samples and blue-collar workers. *European Journal of Psychological Assessment*, *30*(4), 304–314. https://doi.org/10.1027/1015-5759/a000194
Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic Problem Solving: A New Assessment Perspective. *Applied Psychological Measurement*, *36*(3), 189–213. https://doi.org/10.1177/0146621612439620
Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. In *Trends in Cognitive Sciences* (Vol. 14, Issue 11, pp. 497–505). https://doi.org/10.1016/j.tics.2010.08.005
Hirsch-Kreinsen, H., & ten Hompel, M. (2017). Digitalisierung industrieller Arbeit: Entwicklungsperspektiven und Gestaltungsansätze. In *Handbuch Industrie 4.0 Bd.3* (pp. 357–376). https://doi.org/10.1007/978-3-662-53251-5_21
Holyoak, K. J. (2012). Analogy and Relational Reasoning. In *The Oxford Handbook of Thinking and Reasoning*. https://doi.org/10.1093/oxfordhb/9780199734689.013.0013
Kokkonen, T. (2017). Models as Relational Categories. *Science & Education*, *26*(7–9), 777–798. https://doi.org/10.1007/s11191-017-9928-9
Kurtz, K. J., & Honke, G. (2020). Sorting out the problem of inert knowledge: Category construction to promote spontaneous transfer. *Journal of Experimental Psychology: Learning Memory and Cognition*, *46*(5), 803–821. https://doi.org/10.1037/xlm0000750
Markman, A. B., & Gentner, D. (1993). Structural Alignment during Similarity Comparisons. *Cognitive Psychology*, *25*(4), 431–467. https://doi.org/10.1006/cogp.1993.1011
Neubert, J. C., Mainert, J., Kretzschmar, A., & Greiff, S. (2015). The assessment of 21st century skills in industrial and organizational psychology: Complex and collaborative problem solving. In *Industrial and Organizational Psychology* (Vol. 8, Issue 2, pp. 238–268). https://doi.org/10.1017/iop.2015.14
Rattermann, M. J., & Gentner, D. (1998). More evidence for a relational shift in the development of analogy: Children's performance on a causal-mapping task. *Cognitive Development*, *13*(4), 453–478. https://doi.org/10.1016/S0885-2014(98)90003-X
Reeves, L. M., & Weisberg, R. W. (1994). The Role of Content and Abstract Information in Analogical Transfer. *Psychological Bulletin*, *115*(3), 381–400. https://doi.org/10.1037/0033-2909.115.3.381
Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, *94*(3), 249–273. https://doi.org/10.1016/j.jecp.2006.02.002
Rottman, B. M., Gentner, D., & Goldwater, M. B. (2012). Causal Systems Categories: Differences in Novice and Expert Categorization of Causal Phenomena. *Cognitive Science*, *36*(5), 919–932. https://doi.org/10.1111/j.1551-6709.2012.01253.x
Shafto, P., & Coley, J. D. (2003). Development of Categorization and Reasoning in the Natural World: Novices to Experts, Naive Similarity to Ecological Knowledge. *Journal of Experimental Psychology: Learning Memory and Cognition*, *29*(4), 641–649. https://doi.org/10.1037/0278-7393.29.4.641
Snoddy, S., & Kurtz, K. J. (2020). Preventing inert knowledge: Category status promotes spontaneous structure-based retrieval of prior knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0000974
Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, *4*(1), 45–69. https://doi.org/10.1016/0959-4752(94)90018-3
Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, *37*(2), 197–206. https://doi.org/10.1086/651257