DOI: 10.20472/ES.2019.8.2.008

OVERCOMING ROUTINE: A 21ST CENTURY SKILL FOR A 21ST CENTURY ECONOMY

ULRICH STRUNZ, CHRISTIAN CHLUPSA

Abstract:

The aim of this study is to show that overcoming routine holds efficiency potential in complex decision environments and to find an indication on what percentage of decision-makers will successfully overcome routine, when thinking time is not incentivized. Curiosity might favor non-routine behavior, as it is the recognition, pursuit and desire to explore uncertain and ambiguous events. Two hundred sixty-two US-American Mturks completed both a curiosity questionnaire and an experiment in the form of a cognitive puzzle game. High values in self-reported “Joyous Exploration” were not associated with larger numbers of experimental decisions. Contrary to findings where exploration served as a predictor for performance in human-computer interaction and as a mediator for “error learning” (Frese, 1994), participants who performed better relied less on exploratory decisions. Only about 10 % were able to overcome their routine behavior, confirming that induced routine can have a strong influence on behavior (Betsch, Haberstroh, Glöckner, Haar, & Fiedler, 2001). Those who overcame routine reported higher response times when being framed by unexpected feedback, but solved the experiment more efficiently. This is in line with research that fast, intuitive decisions increase framing effects (Guo et al., 2017) and exploration is time-costly (Athukorala, 2015). As described in research, this study further strengthens insights that performance in solving complex problems relies heavily on rule identification, rule knowledge and rule application (Wüstenberg, Greiff, & Funke, 2012).

Keywords:

complex problem solving, dynamic decision making, response time, decision bias, mental model, cognitive reflection, minimal complex system, overcoming routine, non-routine task, curiosity, emotion

JEL Classification: D81, D83, D91

Authors:

ULRICH STRUNZ, Universidad Católica San Antonio de Murcia, Germany, Email: ulrich.strunz@fom-net.de
CHRISTIAN CHLUPSA, FOM University of Economics and Management, Germany, Email: christian.chlupsa@fom.net

Citation:

ULRICH STRUNZ, CHRISTIAN CHLUPSA (2019). Overcoming Routine: A 21st Century Skill for a 21st Century Economy. International Journal of Economic Sciences, Vol. VIII(2), pp. 109-126., 10.20472/ES.2019.8.2.008

1 Introduction

The paper's research strategy is a valid application-test scenario developed to evaluate participants' ability to overcome routine, when thinking time is not incentivized and is considered as a resource by the participants. The goal of this paper is to show that performance in non-routine decision making boosts problem-solving efficiency, and to find common behavioral patterns of agents succeeding in overcoming routine effectively.

Economic working conditions are facing a dynamic and changing environment, commonly referred to as "VUCA" environment, short for "volatility, uncertainty, complexity, and ambiguity". Companies are surrounded by an "increasingly unstable and rapidly changing business world. This (…) will require HR and talent management professionals to change the focus and methods of leadership development." (Lawrence, 2013, p. 2).

Such a novel focus in modern workplaces is the rising importance of non-routine tasks, which are considered to be part of "21st Century Skills" and are related to performance in "complex problem solving" (Neubert, Mainert, Kretzschmar, & Greiff, 2015). Volatility, uncertainty, complexity and ambiguity share a common attribute, linking them to complex problem solving (CPS): they hide relevant information from the problem-solver on first sight. A problem is defined as being "complex" when not all relevant information is available at the outset and multiple interactions are required to successfully solve the problem (Eichmann, Goldhammer, Greiff, Pucite, & Naumann, 2019, p. 2).

Complexity in entrepreneurial decision-making is being dealt by Information Systems, such as Expert Systems, Enterprise Resource Planning and Supply Chain Management (Irani, Sharif, Kamal, & Love, 2014), to enhance forecasting quality. However, even the "best model for forecasting can potentially change over time" (Karel, Hebák, 2018, p. 66). Leaders and employees have to cope with CPS in a VUCA world, where circumstances vary frequently. Therefore, a critical asset to cope with CPS is "problem solving", which is defined as the act of finding "efficient solutions to novel tasks" for which no prior solutions are known, including challenges where "Humans have difficulties to execute counterintuitive movements" (Donnarumma, Maisto, & Pezzulo, 2016). It is suggested that problem solving, when modelled as "heuristic search", has to be described by implementing scientific research from various fields, being "perception, memory, intentionality, decision making, and judgement" (Ohlsson, 2012), linking human problem solving and heuristic decision making.

According to the famous work by Tversky and Kahneman (1974), many decisions are influenced by beliefs on the likelihood of uncertain events (Tversky & Kahneman, 1974). With the introduction of the "somatic marker" concept by Damasio, neuronal evidence of implicit processes influencing decision-making has been shown – a fact that is facing difficulty in being accepted in scientific fields, particularly in the field of "business administration" (Chlupsa, 2014, p. 443).

Such beliefs and implicit processes can lead to bias in decision-situations, where the decision-maker is lacking information to make a decision based on former knowledge (Fazio et al., 2015). Following an inner "status-quo" or "inertia" bias, the decision-maker might prefer consistency over positive feedback (Alós-Ferrer, Hügelschäfer, & Li, 2016). In other words, the decision-maker might fail to overcome routine, despite feedback, while others overcome their bias and proceed with non-routine decisions, to effectively react to novel circumstances.

Thinking time as a resource can be helpful to overcome these biases. Cognitive processes coping with complexity, e.g., answering survey questions of different lengths, are linked to response times (Yan & Tourangeau, 2008), which are a well-researched indicator for overcoming decision biases (Alós-Ferrer, Garagnani, & Hügelschäfer, 2016, p. 19). Response times have predictive power when decision-makers are facing strategic uncertainty (Kiss, Rodíguez-Lara, & Rosa-García, 2018, p. 19), e.g. decision-makers show longer response times when multiple options are seen as equally attractive (Krajbich et al., 2014).

Performance in CPS stems from thinking time, but also from the agents' ability to effectively "identify rules" governing a problem, gaining "rule knowledge" by understanding the problem's internal causal relations (true rule knowledge) and "applying knowledge" by controlling the problem and achieving goals (Wüstenberg, Greiff, & Funke, 2012).

Another important factor influencing CPS are emotions, such as motivation (Güss, Burger, & Dörner, 2017). Under stressful working conditions that allow little time to acquire knowledge about a complex problem's functionality, decision-makers might prefer simple and fast heuristics, such as imitation, which can for example explain "herding behavior" in stock markets (Özsu, 2015). Curiosity might play a role when decision-makers engage in time-costly knowledge investigation, despite stressful working conditions.

Curiosity is part of "epistemic emotions". It is the recognition, pursuit and desire to explore events that are risky and ambiguous (Kashdan et al., 2018) and may also be defined as the desire to gain new information (Litman & Jimerson, 2004, p. 147). Curiosity increases with the knowledge one possesses in a certain area (Loewenstein, 1994, p. 94) and enables creative performance, even without feedback or external stimuli such as informational resources (Hagtvedt, Dossinger, Harrison, & Huang, 2019, p. 10). It motivates exploratory behavior to fill information gaps (Litman, Hutchins, & Russon, 2005, p. 579), and according to the learning progress hypothesis, curiosity is causally linked to learning (Oudeyer, Gottlieb, & Lopes, 2016). Curiosity is of crucial importance to global managers' decision-making processes to discover and understand new knowledge (Harvey, Novicevic, Leonard, & Payne, 2007, p. 58). Curiosity has high predictive validity for task performance, including the Big Five personality traits and tacit knowledge (Mussel, 2013, p. 467).

Despite mentioned insights little is known about non-routine task performance. This paper introduces an experimental setup, which includes testing for self-reported "Joyous exploration". It tests the capacity to solve non-routine problems under "working conditions", where thinking time is not incentivized, but is regarded as a valuable resource for the participants. Results are to shed light on what percentage of agents show exploratory behavior and successfully overcome a routine task, and whether or not exploration positively influenced performance in this regard.

The study begins with a short introduction on the theoretical background, followed by a description of the participants of the study, its research design and a detailed instruction of the experiment's procedure. The next chapter will list the study's results about completion time, amount of actions, response time and "Joyous Exploration". Findings and limitations of the present study are then discussed, and brought into economic day-to-day reality in the final chapter "Conclusions and Future Prospects".

2 Theoretical Background

Puzzles are used in cognitive sciences to research human problem solving. E.g. "Tower of Hanoi" is an important algorithmic puzzle in research areas such as memory, intelligence and executive function (Levitin, 2017, p. 7). The "Tower of Hanoi" problem is regarded as a "simple problem solving task", which does not reflect real-life situations (Spering, Wagener, & Funke, 2005, p. 1253). A problem is defined as being "complex" when not all relevant information is available at the outset and when multiple interactions are required to successfully solve the problem (Eichmann, Goldhammer, Greiff, Pucite, & Naumann, 2019, p. 2).

"Complex Problem Solving" is "typically measured via dynamic systems that contain several interrelated variables that participants need to alter (Dörner & Funke, 2017, p. 2), and is commonly researched using computer-simulations (Güß, 2017).

There has been a heated controversy on the question how to clearly define and set-up "Complex Problem Solving" (CPS) experiments (Greiff, Stadler, Sonnleitner, Wolff, & Martin, 2015; Funke, Fischer, & Holt, 2017; Greiff, Stadler, Sonnleitner, Wolff, & Martin, 2017). There exists agreement that during CPS participants have to overcome barriers stemming from a changing, dynamic problem environment, uncertainty of relevant information and how system states are interconnected.

Environmental changes, such as a change in rules, influence performance when those changes influence individual strategies (Cañas, Quesada, Antolí, & Fajardo, 2003). Companies are required to react to changing markets and to face novel tasks, which not only bring along complexity but also unforeseeable uncertainty (Sommer & Loch, 2004). Understanding complex systems requires the learning of counterintuitive concepts, including the realization that mental models are always incorrect, as they are by definition a simplification of reality (Sterman, 2002).

In contrast to the concept of "rational decision-making", human decision-making is also influenced by the way available choice options are presented, described by Kahneman and Tversky as the "framing effect" (Kahneman & Tversky, 1979; Tversky & Kahneman, 1981). The "framing effect" has been connected to amygdala activity, offering data to support the idea that "more rational" individuals have a better representation of their own emotional biases, which enables them to adapt their behavior when such biases lead to suboptimal decisions (De Martino, Kumaran, Seymour, & Dolan, 2006). Fast and intuitive responses increase framing effects in risky decision-making (Guo, Trueblood, & Diederich, 2017), and according to neuroimaging studies, impulsive/emotional and reflective/cognitive processes are two separate systems that guide human decision-making, violating traditional economic theory (Stallen & Sanfey, 2011, p. 7).

While there exists rich data on the "framing effect", modern approaches to analyze framing have resulted in the understanding that external framing does not influence choice behavior alone and that the original approaches by Kahneman and Tversky were restricted choice problems. Choice behavior is influenced by decision strategies, and how internal representations of framed problems are transformed into choice behavior has to be experimentally tested to ultimately understand decision framing (Maule & Villejoubert, 2007).

Based on mentioned research in the introduction and on the theoretical background, the hypotheses are as follows:

[1] Participants who gain true rule knowledge during the experiment and participants who do not will differ in their exploratory behavior.

[2] Participants who gain true rule knowledge during the experiment will finish the puzzle faster than those who do not.

[3] Participants who gain true rule knowledge during the experiment will finish the puzzle using fewer inputs than those who do not.

[4] Participants who gain true rule knowledge during the experiment show longer response times for their inputs.

[5] Participants who gain true rule knowledge during the experiment show higher scores on the questionnaire for “Joyous Exploration”.

3 Method

Kashdan et al.’s 2018 multidimensional model to measure curiosity via a 25-item questionnaire is applied (Kashdan et al., 2018). It measures self-reported curiosity via five dimensions: “Joyous Exploration”, “Deprivation Sensitivity”, “Stress Tolerance”, “Social Curiosity” and “Thrill Seeking”. Five questions for each category comprise the 25-question survey.

An online behavior experiment in the form of a simple puzzle game was developed to analyze correlations between curiosity and non-routine problem solving. The different behavior approaches of the participants, e.g., whether participants repeated ineffective inputs or tried new approaches, were directly observed.

4 Participants

Two hundred sixty-two Amazon Mechanical Turk (AMT) workers were recruited via Amazon’s Mechanical Turk, a website containing an integrated participant compensation system for online freelancers, who are commonly recruited for behavioral experiments due to AMT’s workers pool size, low costs and being able to produce high-quality data fast (Buhrmester, Kwang, & Gosling, 2011). Amazon Mechanical Turk workers exhibit same biases and heuristics as traditional experiment recruits (Paolacci, Chandler, & Ipeirotis, 2010, p. 417).

The main motivator for AMT Workers (Mturks) is compensation (Lovett, Bajaba, Lovett, & Simmering, 2018), making Mturks suitable candidates to simulate working conditions, where every second counts. We provided 15 minutes completion time, and offered 2.05 USD for this task, 95 USD cents above US minimum wage (7.25 USD per hour). The experiment can be regarded as complex-problem solving under risk, as time is considered a valuable resource for Mturks.

AMT is a suitable recruitment source for any study, when the research object is not expected to vary over cultures and is not specifically relevant to US or Indian samples (Cheung, Sliter, & Sinclair, 2016). While the latter is the case for complex problem-solving, cultural influence on decision-making under uncertainty is mixed. Medium uncertainty avoidance values for US American and low uncertainty avoidance values for Indian citizens have been found in experimental setups (Güss, 2011). No significant differences in planning that is skillful adaption of decision-making to successfully allocate resources over a long-term period, were measured between US American and Indian citizens (Güss, Tuason, & Gerhard, 2010). When it comes to Complex Problem Solving (CPS) cultural differences will be more visible, since it requires causal cognition which is influenced by the cultural learning environment (Funke, 2014). For this reason, only Mturks from the United States of America very accepted. HIT approval rate was required to be equal or above 99 %.

5 Research Design

Participants had to provide login code, age and sex to start the questionnaire. After the final question the experiment started.

During the “Flag Run” experiment all clicks on provided graphical user interface (GUI) items are saved and assigned to an individual participant with a timestamp.

The experiment is designed as a puzzle game, consisting of ten levels. The goal is indicated with a green space and flag, as the interpretation of decision options requires clear goals to be apparent (Jeschke, 2017, p. 20). The playing piece always moves to the left and is therefore only controlled by the number buttons, not by the directional buttons.

This relevant information to control the game is hidden, making it a complex problem-solving task. Knowledge about the true rules has to be acquired by exploration. No change in rules will be implemented; only structural changes will be implemented. These structural representations are called levels. We expect participants to build up and act upon a strategy to solve the puzzle that suits the solving of early levels, which deviates from the true rules governing the game. In later stages, participants will be confronted with uncertainty, as their former and biased strategy will fail to be effective in controlling the game. The game was designed that it was likely for participants to develop a routine during levels 1 to 6, which they had to overcome in order to effectively control the „non-routine” levels 7 to 10.

Participants starting the experiment were provided with a short explanation text on what to do, which read as follows:

“Move the playing piece onto the green goal field. You cannot do this by drag-and-drop.”.

Participants do not see the playing piece actually moving. Only start and end position of the playing piece are visualized.

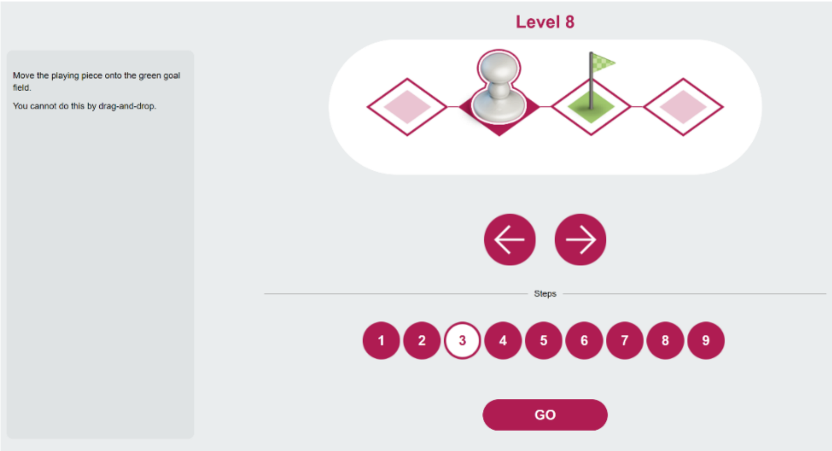
Each level can be solved by one action, consisting of multiple inputs. An action consisting of four inputs is clicking on the playing piece (first action), one of the directional buttons (second action), a valid number button (third action), and “GO” (fourth action).

There are 10 levels. The first 6 stages are designed in such a way that the left and right distances from the playing piece to the goal space are equal. When the playing piece reaches the left end of the playing field, it appears on the right side. The playing piece can reach the goal space even when participants have an incorrect mental model of the rules, e.g., trying to make the playing piece move to the right.

Clicking on the playing piece was mandatory during the first three levels to make the directional buttons appear. Choosing the playing piece will make a white outline appear. It is not possible to “de-select” the playing piece.

Figure 1: screenshot of level 8 of "Flag Run" experiment

authors' own creation



A directional button had then to be selected, making the number buttons appear and filling the chosen direction button with white color. Even though the directional buttons lacked any functionality, clicking on an already selected directional button would remove the white color, giving the impression of being "de-selected". When a directional button was de-selected, the number buttons would still remain. Choosing a number button fills it with white color. With N game fields, a chosen number button with number n had to be lower than or equal to N-1 to make the "GO" button appear. A number button could not be "de-selected". Switching from a valid to an invalid number button makes the GO button disappear. Clicking on the GO button fulfills an action. It is never mandatory to have a directional button selected, i.e. being filled with white color. During the first three levels, a minimum of 4 inputs is required to perform an action: clicking on the playing piece, a directional button, a valid number button and GO.

During level 4 to 6, a click on the playing piece will make both directional and number buttons appear. Only three inputs are required to perform an action: clicking on the playing piece, a valid number button and GO.

Levels 7 to 10 have the playing piece automatically pre-selected, so that direction and number buttons are already displayed. Only two inputs are required to perform an action: click on a valid number button and GO.

6 Procedure

During the first 6 levels, any departure from the minimal amount of inputs and actions required to solve the game is counted. Throughout levels 7-10, each action is assigned an "action state". There are 5 distinguished action states: "framed logic", "random", "experimental", "realize", and "king". These action states are explained in more detail below, and are further explained in Table 1.

Table 1: move states and categories

authors' own creation

Playing stone left of goal field						**Playing stone right of goal field**					
Direction: left		Direction: none		Direction: right		Direction: left		Direction: none		Direction: right	
R	W	R	W	R	W	R	W	R	W	R	W
10	**1**	**11**	**7**	**6**	**3**	**9**	**2**	**12**	**8**	**5**	**4**

If distance and direction, assuming direction could actually be manipulated, lead to goal field = R.

If distance and direction, assuming direction could actually be manipulated, would not lead to goal field = W

If Direction: none, = R when distance leads to goal field, = W when not.

1. <u>Actions which do not result in completion of level:</u>
   6 = framed logic
   1, 2, 3, 4, 5, 7, 8 = experimental
2. <u>Actions leading to completion of level:</u>
   1, 2, 4, 5, 6, 7, 8 = random
   9 = random
   3, 10, 11, 12 = random, when not all following levels are completed in one action
   3, 10, 11, 12 = realize, when all following levels are completed in one action
   3, 10, 11, 12 = king, when previous level was solved in one action and/or all following levels are solved with a 3, 10, 11, or 12 in one action

With the playing piece being positioned to the left of the goal space, the player commits a "framed logic" input by using the "right" directional button and choosing the number of steps that would place the playing piece onto the goal if the "right" directional button would actually move the playing piece to the right, which it does not.

Any other combination of inputs leading to an action that does not place the playing piece onto the goal is considered an "experimental" action state.

A level is solved when the playing piece lands on the goal space. When a player commits an action that solves a level, it is assigned the action state of "random", "realize" or "king".

All actions putting the playing piece onto the goal field are considered "random" actions when they do not stem from "true rule knowledge" (TRK). True rule knowledge implies the realization that the playing piece jumps edges, that the default direction is set to "left" and that the directional buttons do not work. Participants who have not shown by their behavior of having at least partly obtained TRK are of the "rule knowledge" (RK) group.

The "realize" action is performed when the last step shows "true rule knowledge", and all remaining levels are also solved in this manner. When the following levels include "framed logic" or "experimental" actions, however, it is still considered a "random" action state.

When participants solve a level in one step showing "true rule knowledge", then this action is called "king".

Participants with one or more “king” actions are regarded as being part of the group that understands the “true rule knowledge” of the game, while all other participants are regarded as being part of the group that simply holds “rule knowledge”.

7 Results

The statistical analysis was performed with IBM SPSS Statistics 25.0. The metric-dependent variables are not normally distributed. The distinguishing factors between the two groups “true rule knowledge” (TRK) and “rule knowledge” (RK) are researched with the Mann-Whitney U-test. All ordinal dependent data are also tested with the Mann-Whitney U-test for independent variables.

Out of 262 participants, only 28 are part of the TRK group, while 234 participants are part of the RK group. These two groups are compared. No significant difference between women and men is detected ($U(28,234) = 2805$, $z = -1.519$, $p = 0.129$). Age shows also no significant differences ($U(28,234) = 2616$, $z = -1.744$, $p = 0.081$). The average age of the TRK group is 32.04 years ($SD = 7.918$), and the average age of the RK group is 35.15 years ($SD = 9.967$).

7.1 Exploratory Behavior

The difference in average number of inputs deviating from the minimal requirement to solve the level in one attempt from levels 1 to 6 between TRK group and RK group was not significant ($U(28,234) = 2820.5$, $z = -1.209$, $p = 0.227$). TRK requires 10.79 ($SD = 13.636$) and RK 16.19 ($SD = 30.227$) more inputs on average to solve level 1 to 6 than is actually necessary.

During levels 1 to 6 both groups require more inputs than are necessary, because of exploration or confusion about the functionality. The deviation of inputs from level 1 to 6 is not a significant predictor of obtaining TRK.

The difference in the amount of attempts required to solve level 1 to 6 between TRK and RK is not significant ($U(28,234) = 2957$. $z = -1.047$, $p = 0.295$). TRK required an average of 6.79 ($SD = 2.331$) and RK 7.09 ($SD = 2.422$) attempts to solve the first 6 levels. TRK and RK show comparable values in both amount of inputs and attempts required during the first 6 levels.

Whether or not a participant obtained TRK was measured from level 7 to 10. As soon as a participant grasps the true rules of the experiment, the participant does not have to experiment anymore and will most likely stick to “realize” actions. To find differences in behavior, the relationship between “framed logic” or “experimental” actions and the total amount of actions in levels 7 to 10 is analyzed.

“Framed logic” actions per total amount of attempts from level 7 to 10 does not differ significantly between the TRK and RK groups ($U(28,234) = 2565.5$, $z = -1.876$, $p = 0.061$), meaning that TRK participants perform as many “framed logic” actions in relation to the total amount of actions as do RK participants.

The amount of “framed actions” in relation to the total amount of actions in level 7 to 10, with “realize” and “king” actions subtracted from that total amount however does differ significantly ($U(28,234) = 1295$, $z = -5.229$, $p = 0.000$). This is because TRK participants perform significantly less “experimental” ($U(28,234) = 1393.5$, $z = -4.986$, $p = 0.000$) or “random” ($U(28,234) = 35$, $z = -10.020$, $p = 0.000$) actions from level 7 to 10 than do RK participants.

The relationship between "experimental" actions and the total amount of actions (including "realize" and "king" actions) from levels 7 to 10 between the TRK and RK groups is significant (U(28, 234) = 1572, z = -4.499, p = 0.000). Same is true with "realize" and "king" actions excluded from the total amount of actions from levels 7 to 10 (U(28,234) = 2066, z = -3.194, p = 0.001).

7.2 Completion Time

No significant difference exists in completion time of levels 1 to 6 between TRK and RK participants (U(28, 234) = 2747.5, z = -1.395, p = 0.163).

To solve level 7 to 10, TRK participants require 38 secs to 168 secs and, on average, 82.679 secs (SD = 37.82 secs). It takes RK participants on average 104.49 secs (SD = 66.93 secs), ranging from 38 secs to 507 secs.

The difference in completion time of levels 7 to 10 between the two groups in total is significant (U(28, 234) = 2514, z = -2.011, p = 0.044).

7.3 Amount of Actions

From levels 7 to 10, TRK participants require 5 to 31 actions to solve these levels, or 12.54 actions on average (SD = 6.119). RK participants require 5 to 138 actions to solve levels 7 to 10, or 19.09 (SD = 17.405) actions on average. This difference between the groups is significant (U(28,234) = 1841, z = -3.794, p = 0.000).

7.4 Response Time

The total average response time per action from level 7 to 10 between the two groups is significant (U(28,234) =1619, 5, z=-4.373, p=0,000). Response times begin to deviate and its significance grows as levels' difficulty and complexity grow.

It takes TRK participants 6.60 secs (SD = 3.295 secs) and RK participants 6.855 secs (SD = 3.608 secs) on average per action during level 7. Differences in response time during level 7 are not significant between the two groups (U(28, 234) = 2879.5, z = -1.047, p = 0.295).

Response times begin to clearly deviate at level 8. It takes TRK participants 7.37 secs (SD = 2.639 secs) and RK 6.53secs (SD = 2.718 secs) per action on average in level 8, barely showing no significant difference (U(28,234) = 2539, z = -1.949, p = 0.051).

The response times per action during level 9 differ further, and are significant (U(28,234) = 2295.5, z = -2.591, p = 0.010) between TRK 7.25 secs (SD = 2.365 secs) and RK 6.43 secs (SD = 5.629 secs).

Level 10 shows the largest difference in response times on average. Difference between TRK participants 12.07 secs (SD = 4.56 secs) and RK participants 5.96 secs (SD = 3.97 secs) is significant (U(28,234) = 368.5, z = -7.675, p = 0.00). All participants of the "true rule knowledge" group solve level 10 in one action by definition and show a "realize" action beforehand. It can be assumed that the large response times of level 10 stem from participants counting playing spaces to make a correct move. With 9 playing spaces and the playing piece starting at the leftmost playing space, participants have to correctly count 8 playing spaces to solve this level in one action.

7.5 Joyous Exploration

Participants also self-report their curiosity levels by a 25-item questionnaire in five dimensions: JOY, NEED, STRESS, SOCIAL, and THRILL. JOY best describes the exploration of novelties just for the sake of knowing more. No significant difference in any of the self-reported items was found between TRK and RK participants, including JOY ($U(28, 234) = 3175.5$, $z = -0.266$, $p = 0.790$). TRK self-reported JOY ranging from 16 to 35, and RK from 5 to 35, leading to the conclusion that very low self-reported JOY values were not part of the TRK group. Correlation between self-reported JOY values and "experimental actions" in relation to the total amount of actions (including "realize" and "king" actions) from level 7 to 10 is not significant ($p = 0.340$). Correlation between total average response time from levels 7 to 10 and self-reported JOY is not significant ($p = 0.286$).

8 Discussion

TRK does not perform better or worse during the "routine" levels 1-6. This confirms findings that as long as circumstances, such as framing, do not influence decision-making strategies, performance is not necessarily altered (Cañas, Quesada, Antolí, & Fajardo, 2003). Starting with level 7, routine action was no longer effective. TRKs performed as many routine actions in relation to the total amount of actions during levels 7-10 as RK participants. However, TRK performed less actions in total and performed fewer experimental and random actions in levels 7-10 than RK in relation to the total amount of actions. TRK had a significantly higher proportion of "framed logic" moves, when "realize" and "king" actions are subtracted from their total amount of actions. This is in line with experiments, where participants would stick with their old routine decisions despite information clearly indicated that routine action became obsolete and that failing to overcome such routine depended on the level of strength in routine (Betsch, Haberstroh, Glöckner, Haar, & Fiedler, 2001). While all participants in the "Flag Run" experiment were confronted with the equal amount of levels where such routine was deliberately nudged, it can be assumed that individual routine levels among the participants varied, having an influence on the individual behavior and effectiveness of finding an efficient strategy to solve level 7 to 10.

This result also means that exploratory behavior, including random or experimental actions, did not enhance the chances of obtaining TRK. Reflection on feedback and understanding of information provided by exploratory behavior comprised the true critical success factor in obtaining true rule knowledge. TRK and RK performed equally well during level 1 and 6, showing comparable behavior when measured in terms of total actions, input deviations and completion time. Behavior changes between both groups started with level 7 and differences became more significant with growing level complexity and difficulty. Hypothesis [1] is falsified for behavior in levels 1-6 and confirmed for levels 7-10. This strengthens findings that rule identification and correct mental formulation of the causal model governing the complex problem is essential for being successful in applying knowledge (Wüstenberg, Greiff, & Funke, 2012). This means that participants who performed more experimental actions failed to identify the true rules since level 7, and failed to construct rule knowledge, which results in more "useless" actions than participants who outperformed in rule identification and were then able to "stick" with an efficient strategy based on a mental model, which at least partly represented the true rules of the problem.

To an Amazon Mechanical Turk worker time is a resource, and thinking time per action can be considered a risk. The costs of an action are equal to the time taken and number of inputs used to perform this step. Even though reflection time brings about higher costs, it pays off.

TRK participants did not outperform RK participants during level 1 to 6 in regards to completion time, but completed the total experiment faster, since TRK participants required less moves to complete levels 7 to 10, where TRK was an advantage regarding efficiency. These results are in line with recent research that exploratory behavior results in larger completion time compared to less costly decision heuristics (Athukorala, 2015).

This confirms hypothesis [2] and shows that investing thinking time in non-routine problem solving pays off as task complexity and difficulty grows.

TRK require more time per action but require fewer actions in total. This confirms hypothesis [3]. TRK participants are more both more time- and action-efficient when facing the non-routine problem experiment. As a more precise mental model of the true rules governing the game enables the participants to solve e.g. level 10 in one step.

As intuitive responses in risky decision-making increase framing effects (Guo et al., 2017), participants who failed to understand the true rules governing the complex problem task due to framing effects were expected to show shorter response times. This expectation has been shown statistically, confirming hypothesis [4]. In other words, short reflection times ultimately decreased the probability of participants understanding the true rules of the puzzle. Participants who understand at least parts of the true rules governing the game report higher response times and reflect longer with increasing level complexity and difficulty. These findings are in line with the extant research.

Hypothesis [5] cannot be confirmed. While participants with low scores in "Joyous Exploration" were not found to be in the TRK group, no statistically significant differences were found. Participants with high self-reported JOY values do not show more experimental actions, are not correlated to response times and are not more likely to obtain true rule knowledge, as incorrectly assumed by hypothesis [5]. Exploratory behavior and curiosity alone are not enough to see through complexity, putting more weight on the importance of learning under uncertainty. These results also show that self-reported curiosity should be rated with care. While it was shown that exploration predicts performance in human-computer interaction, and also serves as a mediator to "error training" (Frese, 1994), more research has to be conducted to understand the role of exploratory behavior on problem solving performance when unexpected feedback can be wrongly perceived as "errors".

9 Limitations of the Present Study

The experimental results do not reflect important influencers on economic decision-making such as social influencers that can be tested in multiplayer experiments (van den Bos, Talwar, & McClure, 2013).

While there does not exist consensus on how to clearly define deception and when deception occurs, "Deception may be defined as concealing or camouflaging the real purpose of an experiment (i.e. the data in which the scientist is interested) to avoid conscious reactivity of participants that would make these data worthless" (Krawczyk, 2019, p. 111).

The purpose of the experiment was to measure patterns in behavior that can be found with participants, who successfully overcome a false causal mental model. Conscious reactivity, being the participants' behavior to overcome a wrong understanding of the game's rules e.g. the directional buttons influencing the direction, was not avoided by the setup, but was the data in which this paper is interested in. In fact, such conscious reactivity was anticipated and encouraged by level design, starting with level 7.

No “deceiving” information was provided: The explanation text did not include the information that GUI elements would appear, when participants clicked on the playing piece. This is because in later levels, this would not be true and could also manipulate participants into thinking that clicking on the playing piece was required during levels 7 to 10, where the playing piece is already pre-selected by the program. The explanation text did not include any information about the direction and number buttons. Players would build their own understanding of how the game works, while some outperformed others in successfully updating their understanding when being confronted with very subtle changes of the “environment”. No deception was used in this study.

10 Conclusions and Future Prospects

Thinking time during non-routine problem solving pays off. Participants who reflected longer were more likely to figure out hidden information by learning from feedback despite uncertainty. The costs associated with reflection when facing non-routine problems are well spent as complexity and difficulty rises. Those who reflected more required fewer steps and less time to solve complex levels. This is no easy task however. Only about 10 % of the participants were able to overcome their routine, obtain true rule knowledge and outperform in terms of efficiency. This insight could be of value for enterprises that face high levels of uncertainty, which comes along with economic decision making.

The influence of exploratory behavior on obtaining hidden information during non-routine problem solving was sobering. The main indicator of participants who performed better in realizing how the experiment actually works was thinking time. Participants with high “Joyous Exploration” were not more likely to obtain true rule knowledge nor did participants who performed more “experimental” actions.

Implicit motives that might have had an effect on all participants assigning functionality to the directional buttons should be analyzed in future experimental setups, as not a single participant determined that the directional buttons were actually useless. This could be due to strong implicit associations with directional buttons. The concept of influencing the direction of objects via buttons is certainly strongly pattern-embedded in our day-to-day reality. The human brain is in constant search of known patterns, acting as an “association machine” (Chlupsa, 2017, p. 262). Routines might be heavily influenced by strong associations, which are then costly to overcome, building up barriers to successfully overcome routine behavior. This might promote „status quo“, „inertia“ or “confirmation” bias, as was shown by many participants, who repeatedly performed the same routine action over and over again, when being faced with the non-routine problem for the first time in level 7, despite feedback clearly indicated lack of progress by doing so. Learning from unexpected feedback seems to be a critical success factor in overcoming routine.

While many researches and economists press the importance of skills that enhance adaption to changing conditions, it has to be understood that overcoming routine and its linked set of behavioral biases is not easily performed, and can probably only be done by a small fraction of leaders and employees, when there is not much time to reflect on the problem at hand. In an interconnected, globalized world, where information will reach enterprises at a blink of an eye, quick and non-routine decision-making is essential. Further research is required to better understand possible influencers on non-routine decision making, such as cognitive reflection, learning from unexpected feedback, social interaction with uncertainty, learning agility and mental strategy making.

11 References

Alós-Ferrer, C., Garagnani, M., & Hügelschäfer, S. (2016). Cognitive reflection, decision biases, and response times. Frontiers in Psychology, 7, 1402. doi: 10.3389/fpsyg.2016.01402

Alós-Ferrer, C., Hügelschäfer, S., & Li, J. (2016, February 16). Inertia and Decision Making. frontiers in Psychology, pp. 1-9. doi: 10.3389/fpsyg.2016.00169

Athukorala, K., Głowacka, D., Jacucci, G., Oulasvirta, A., & Vreeken, J. (14. July 2015). Is Exploratory Search Different? A Comparison ofInformation Search Behavior for Exploratory andLookup Tasks. Journal of the Association for Information Scienceand Technology, S. 1-17. doi:10.1002/asi.23617

Betsch, T., Haberstroh, S., Glöckner, A., Haar, T., & Fiedler, K. (2001, January). The Effects of Routine Strength on Adaptation and Information Search in Recurrent Decision Making. Organizational Behavior and Human Decision Processes, pp. 23-53. doi:10.1006/obhd.2000.2916

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011, February 3). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?. Perspectives on Psychological Science, pp. 3-5. https://doi.org/10.1177/1745691610393980

Cañas, J., Quesada, J., Antolí, A., & Fajardo, I. (2003). Cognitive flexibility and adaptability to environmental changes in dynamic complex problem-solving tasks. Ergonomics, 46(5), 482–501. doi: 10.1080/0014013031000061640

Cheung, J. H., Sliter, M. T., & Sinclair, R. R. (2016, June). Amazon Mechanical Turk in Organizational Psychology: An Evaluation and Practical Recommendations. Journal of Business and Psychology, pp. 347-361. https://doi.org/10.1007/s10869-016-9458-5

Chlupsa, C. (2014). The impact of implicit motives on the business to business decision-making process (Doctoral dissertaton), University of Plymouth, Plymouth.

Chlupsa, C. (2017). Der Einfluss unbewusster Motive auf den Entscheidungsprozess - Wie implizite Codes Managemententscheidungen steuern. Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-07230-8

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. Science, 313(5787), 684–687. doi: 10.1126/science.1128356

Donnarumma, F., Maisto, D., & Pezzulo, G. (2016). Problem solving as probabilistic inference with subgoaling: Explaining human successes and pitfalls in the Tower of Hanoi. PLoS Computational Biology, 12(4), e1004864. doi: 10.1371/journal.pcbi.1004864

Dörner, D., & Funke, J. (2017, July 11). Complex Problem Solving: What It Is and What It Is Not. Front. Psychol., pp. 1-11. doi: 10.3389/fpsyg.2017.01153

Eichmann, B., Goldhammer, F., Greiff, S., Pucite, L., & Naumann, J. (2019). The role of planning in complex problem solving. Computers & Education, 128, 1–12. doi: 10.1016/j.compedu.2018.08.004

Fazio, R. H., Pietri, E. S., Rocklage, M. D., & Shook, N. J. (2015). Positive versus negative valence: Asymmetries in attitude formation and generalization as fundamental individual differences. In J. M. Olson & M. P. Zanna (Eds.), Advances in Experimental Social Psychology (pp. 97–146). Burlington: Elsevier. https://doi.org/10.1016/bs.aesp.2014.09.002

Frese, M. (1994, October). Error training: Replication and the function of exploratory behavior. International Journal of Human-Computer Interaction, pp. 365-372. doi: 10.1080/10447319409526101

Funke, J. (2014, July 16). Analysis of minimal complex systems and complex problem solving require different forms of causal cognition. Front. Psychol. https://doi.org/10.3389/fpsyg.2014.00739

Funke, J., Fischer, A., & Holt, D. (2017, January 5). When Less Is Less: Solving Multiple Simple Problems Is Not Complex Problem Solving—A comment on Greiff et al. (2015). Intelligence, pp. 1-11. doi: 10.3390/jintelligence5010005

Greiff, S., Stadler, M., Sonnleitner, P., Wolff, C., & Martin, R. (2015, March 14). Sometimes less is more: Comparing the validity of complex problem solving measures. Intelligence, pp. 100-113. doi:10.1016/j.intell.2015.02.007

Greiff, S., Stadler, M., Sonnleitner, P., Wolff, C., & Martin, R. (2017, January 5). Sometimes More is Too Much: A Rejoinder to the Commentaries on Greiff et al. (2015). Intelligence, pp. 1-10. doi: 10.3390/jintelligence5010006

Guo, L., Trueblood, J. S., & Diederich, A. (2017). Thinking fast increases framing effects in risky decision making. Psychological Science, 28(4), 530–543. doi: 10.1177/0956797616689092

Güss, C. D. (2011). Fire and Ice: Testing a Model on Culture and Complex Problem Solving. Journal of Cross-Cultural Psychology, pp. 1279-1298. https://doi.org/10.1177/0022022110383320

Güss, C. D. (13. May 2017). Comparing Business Experts and Novices in Complex Problem Solving. Intelligence. doi: 10.3390/jintelligence5020020

Güss, C. D., Burger, M. L., & Dörner, D. (23. May 2017). The Role of Motivation in Complex Problem Solving. Front. Psychol., S. 1-5. doi: 10.3389/fpsyg.2017.00851

Güss, C. D., Tuason, T. M., & Gerhard, C. (2010, March). Cross-National Comparisons of Complex Problem-Solving Strategies in Two Microworlds. Cognitive Science. https://doi.org/10.1111/j.1551-6709.2009.01087.x

Hagtvedt, L. P., Dossinger, K., Harrison, S. H., & Huang, L. (2019). Curiosity made the cat more creative: Specific curiosity as a driver of creativity. Organizational Behavior and Human Decision Processes, 150, 1–13. doi: 10.1016/j.obhdp.2018.10.007

Harvey, M., Novicevic, M., Leonard, N., & Payne, D. (2007). The role of curiosity in global managers' decision-making. Journal of Leadership & Organizational Studies, 13(3), 43–58. doi: 10.1177/10717919070130030401

Irani, Z., Sharif, A., Kamal, M. M., & Love, P. E. (2014). Visualising a knowledge mapping of information systems investment evaluation. Expert Systems with Applications, pp. 105-125. doi: 10.1016/j.eswa.2013.07.015

Jeschke, B. G. (2017). Entscheidungsorientiertes Management. Einführung in eine konzeptionell fundierte, pragmatische Entscheidungsfindung. Berlin/Boston: De Gruyter. https://doi.org/10.1515/9783110518207

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263–292. doi: 10.2307/1914185

Karel, T., & Hebák, P. (2018). Forecasting Czech GDP using Bayesian dynamic model averaging. International Journal of Economic Sciences, Vol. VII(1), pp. 65-81., doi: 10.20472/ES.2018.7.1.004

Kashdan, T. B., Stiksma, M. C., Disabato, D. J., McKnight, P. E., Bekier, J., Kaji, J., & Lazarus, R. (2018). The five-dimensional curiosity scale: Capturing the bandwidth of curiosity and identifying four unique subgroups of curious people. Journal of Research in Personality, 73, 130–149. doi: 10.1016/j.jrp.2017.11.011

Kiss, H. J., Rodriguez-Lara, I., & Rosa-Garcia, A. (2018). Does response time predict withdrawal decisions? Lessons from a bank-run experiment. SSRN Electronic Journal. doi: 10.2139/ssrn.3292159

Krawczyk, M. (2019). What should be regarded as deception in experimental economics? Evidence from a survey of researchers and subjects. Journal of Behavioral and Experimental Economics, pp. 110-118. doi: 10.1016/j.socec.2019.01.008

Lawrence, K. (2013). Developing Leaders in a VUCA Environment. UNC Executive Development 2013, pp. 1-15.

Levitin, A. (2017). Algorithmic puzzles: History, taxonomies, and applications in human problem solving. The Journal of Problem Solving, 10(1). doi: 10.7771/1932-6246.1194

Litman, J., Hutchins, T., & Russon, R. (2005). Epistemic curiosity, feeling-of-knowing, and exploratory behaviour. Cognition & Emotion, 19(4), 559–582. doi: 10.1080/02699930441000427

Litman, J. A., & Jimerson, T. L. (2004). The measurement of curiosity as a feeling of deprivation. Journal of Personality Assessment, 82(2), 147–157. doi: 10.1207/s15327752jpa8202_3

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. Psychological Bulletin, 116(1), 75–98. doi: 10.1037//0033-2909.116.1.75

Lovett, M., Bajaba, S., Lovett, M., & Simmering, M. J. (2017, December 05). Data Quality from Crowdsourced Surveys: A Mixed Method Inquiry into Perceptions of Amazon's Mechanical Turk Masters. Applied Psychology. https://doi.org/10.1111/apps.12124

Maule, J., & Villejoubert, G. (2007). What lies beneath: Reframing framing effects. Thinking & Reasoning, 13(1), 25–44. doi: 10.1080/13546780600872585

Mussel, P. (2013). Introducing the construct curiosity for predicting job performance. Journal of Organizational Behavior, 34(4), 453–472. doi: 10.1002/job.1809

Neubert, J., Mainert, J., Kretzschmar, A., & Greiff, S. (2015, June). The Assessment of 21st Century Skills in Industrial and Organizational Psychology: Complex and Collaborative Problem Solving. Industrial and Organizational Psychology, pp. 238-268. doi:10.1017/iop.2015.14

Ohlsson, S. (2012). The Problems with Problem Solving: Reflections on the Rise, Current Status, and Possible Future of a Cognitive Research Paradigm. The Journal of Problem Solving, pp. 101-128. doi: 10.7771/1932-6246.1144

Oudeyer, P. Y., Gottlieb, J., & Lopes, M. (2016). Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. Progress in Brain Research, 229, 257–284. doi: 10.1016/bs.pbr.2016.05.005

Özsu, H. H. (2015). Empirical Analysis of Herd Behavior in Borsa Istanbul. International Journal of Economic Sciences, Vol. IV(4), pp. 27-52., doi: 10.20472/ES.2015.4.4.003

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010, August). Running experiments on Amazon Mechanical Turk. Judgment and Decision Making, pp. 411-419.

Sommer, S. C., & Loch, C. H. (2004). Selectionism and learning in projects with complexity and unforeseeable uncertainty. Management Science, 50(10), 1334–1347. doi: 10.1287/mnsc.1040.0274

Spering, M., Wagener, D., & Funke, J. (2005). The role of emotions in complex problem solving. Cognition & Emotion, 19(8), 1252–1261. doi: 10.1080/02699930500304886

Cochran, J. J., Cox, L. A., Keskinocak, P. , Kharoufeh, J. P., Smith, J. C., Stallen, M. and Sanfey, A. G. (2011). Neuroeconomics and Game Theory. In J. J. Cochran, L. A. Cox, P. Keskinocak, J. P. Kharoufeh, & J. C. Smith (Eds.), Wiley Encyclopedia of Operations Research and Management Science. doi: 10.1002/9780470400531.eorms0566

Sterman, J. D. (2002). All models are wrong: Reflections on becoming a systems scientist. System Dynamics Review, 18(4), 501–531. doi: 10.1002/sdr.261

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185(4157), 1124–1131. doi: 10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science, 211(4481), 453–458. doi: 10.1126/science.7455683

van den Bos, W., Talwar, A., & McClure, S. M. (2013). Neural correlates of reinforcement learning and social preferences in competitive bidding. Journal of Neuroscience, 33(5), 2137–2146. doi: 10.1523/jneurosci.3095-12.2013

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving — More than reasoning? Intelligence, 40(1), 1–14. doi: 10.1016/j.intell.2011.11.003

Yan, T., & Tourangeau, R. (2008). Fast times and easy questions: The effects of age, experience and question complexity on web survey response times. Applied Cognitive Psychology, 22(1), 51–68. doi: 10.1002/acp.1331