UNIVERSITÀ DEGLI STUDI DI BERGAMO
Department of Economics

WORKING PAPERS

Target-the-Two: A lab-in-the-field experiment on routinization

Giuseppe Attanasi, Massimo Egidi, Elena Manzoni

December 2021 - WP N. 7 Year 2021

Target-the-Two: A lab-in-the-field experiment on routinization*†

Giuseppe Attanasi (GREDEG, CNRS, Université Côte d'Azur)‡
Massimo Egidi (School of Government, LUISS University, Rome)§
Elena Manzoni (Department of Economics, University of Bergamo)¶

December 15, 2021

Abstract

The paper investigates the determinants of routinization and creativity by means of a lab-in-the-field experiment run at the 20th edition of a mass gathering festival in Italy ("La Notte della Taranta"). In the experiment, subjects play repeatedly the puzzle version of Target-The-Two game (32 hands). We find that when we focus on expert subjects there is no difference in behavior between creative and routinized individuals. When we consider inexpert subjects, instead, routinized individuals perform better, due to the fact that they are faster. However, routinization, although increasing the likelihood to complete all the 32 hands of the game, it increases the number of moves needed to complete them, which ultimately decreases the likelihood to win the game.

Keywords: creativity, routinization, Target-The-Two game, lab-in-the-field experiment

JEL codes: C93, D91, O31

*We thank Gian Paolo Jesi, Plinio Limata, Maria Cristina Scarafile, Irene Staffieri, Elena Stancarelli, and Salvatore Vergine for research assistantship in the implementation of the lab-in-the-field experiment. We thank Daniela Di Cagno, Alexis Garapin, Patrick Llerena, and participants at the 2018 annual meeting of the SEET (Southern Europe Experimental Team) in Lecce, the 18th conference of the International Schumpeter Society at LUISS Rome, and at the seminar series of the Laboratoire d'économie appliquée de Grenoble (GAEL) for useful comments.

†*Funding*: The research leading to these results has received funding from the French Agence Nationale de la Recherche (ANR) under grant ANR-18-CE26-0018 (project GRICRIS). *Conflict of Interest*: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. *Ethical Conduct*: Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements; the experiment was conducted in compliance with ethical standards and with the procedures of CESARE Lab of LUISS Rome. *Data Availability Statement*: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

‡giuseppe.attanasi@unice.fr

§Corresponding author: megidi@luiss.it

¶elena.manzoni@unibg.it

1 Introduction

In the recent years, economics has started considering how creativity affects economic outcomes, how we can measure it, and what can be done to promote its development.

A relevant part of the literature dealing with creativity recognizes the role that individual characteristics, such as genetics, divergent thinking and cognitive development, play in fostering it (see, e.g., Simonton, 2000).

To study creativity, and its origins, however, we need to agree on a definition of creativity. From a psychological point of view, creativity has been seen as an innate feature of individuals, while from a social perspective, it is acknowledged that many external factors can influence it. Measuring creativity and identifying creative paths in the laboratory is one of the most recent challenges in experimental economics (see, e.g., Charness and Grieco, 2019). However, reliable measures of creative and innovative abilities in real-life situations have not been yet produced by the experimental literature in either psychology and economics (see the surveyed articles in Attanasi et al. 2021a).

Our research aims to fill this gap by considering a specific aspect of creativity, that is the ability of subjects to explore the space of possible solutions to a problem. In this respect, creativity is opposed to routinization, in the sense that routinized subjects maintain their routines and do not test alternative procedures that may prove to be more efficient than the adopted one (or not).

To do so, we have run an experimental study aimed at measuring creativity and routinization in the field and to detect correlations between attitudes to routinization on the one side, and subjects' cognition and other idiosyncratic features on the other side. Specifically, in this study we explore the causes and effects of creativity and routinization by means of a lab-in-the-field experiment, with a twofold aim. First, we aim to test the formation of routines outside the lab (lab-in-the-field), and the causes of routinization, by looking both at individual characteristics (gender, age and education) and individual behavior (such as alcohol consumption). Second, we investigate the effects of routinization and creativity on performance, in order to understand when it is the case that creative subjects outperform routinized ones, or vice-versa.

We run the experiment with a mobile lab positioned at the 20th edition of the traditional music Festival "La Notte della Taranta". To this aim, we have also developed a web platform —Graphgames—, which was a necessary condition to implement the field experiment with tablets and analyze creativity in real time. We proposed to subjects a 'puzzle' game experiment consisting in a repeated individual decision problem, namely the one-player version of the *Target-The-Two* game (Cohen and Bacdayan 1994). The Target-The-Two game admits two strategies that can be optimal, depending on the initial configuration. The experiment is composed by a maximum of 32 hands of the game, where in the first 16 hands (training phase) subjects face only problems that are all optimally solved by one of these two strategies. In the last 16 hands, instead, configurations are selected randomly, and therefore, on average, half of them can be solved with one strategy and the other half with the alternative one.

The two-phase structure of the game allows us to investigate both the learning process, i.e., the ability of understanding and adopting the optimal strategy in phase 1, where all hands have the same optimal solution, and the routinization process, i.e., the process for which agents keep using the same routine (strategy) even when they face tasks that possibly have different optimal solutions. Players who after the training phase routinized on one strategy stick to it once they have identified it. Creative players, instead, explore different strategies according to the specific configuration of the game, possibly selecting in each hand the strategy that would let them solve the game in a more efficient way, i.e. with a smaller number of moves.

We explore the possibility that creativity and routinization may have different effects on agents depending on their level of expertise. We measure expertise as the ability to learn and adopt the optimal strategy in the first phase, where subjects are trained with a series of problems that are all solved optimally with the same strategy. We find that in expert subjects overall performance does not differ between creative and routinized subjects, even though creative subjects seem to make a lower number of moves, and routinized subjects are faster. Things are different when focusing on inexpert subjects. In inexpert subjects routinization improves performance significantly. The channel of the improvement is twofold: on the one hand routinization reduces variance in the performance, on the other hand it reduces the average time by making subjects faster.

The contribution of our paper is threefold. First, we provide a measure of one particular aspect of creativity, thus contributing to the literature on experimental measures of creativity (see Attanasi et al. 2021a for a comprehensive survey of the literature).

Second, we contribute to the literature on the determinants and the effects of routinization, and on the link between expertise and routinization. From previous experiments (Luchins Luchins, 1959) we know that the more a player becomes expert, the more his reaction becomes fast and automatic. We explore this link further, by investigating the interaction between the level agents' expertise and the efficiency of routinization, which is a novel contribution in the literature. Moreover, we contribute to the specific literature on Target-The-Two and routinization, which is extensively discussed in Section 2, by implementing the puzzle version of Target-The-Two so as to isolate routinization that is not related to coordination issues, differently from existing work in the literature (with the exception of Egidi 2016).

Finally, we contribute to the literature on lab-in-the-field experiments, i.e., studies conducted in a naturalistic environment targeting the theoretically relevant population but using a standardized, validated lab paradigm. The "lab-in-the-field" methodology combines elements of both lab and field experiments in using standardized, validated paradigms from the lab in targeting relevant populations in naturalistic settings (Gneezy and Imas 2017). On the one side (lab), employing a standardized paradigm permits the experimenter to maintain tight control; on the other side (field), targeting the relevant population and setting increases the applicability of the results. Indeed, our study is the first lab-in-the-field study of the Target-The-Two game, with a sample of 480 participants. The gathering event where we carried it out is “La Notte della Taranta” Festival, held each year since 1998 in the province of Lecce (South of Italy) in late August (www.lanottedellataranta.it/en/). The event is among the most important European folk festivals: in the last 10 years, it was able to attract approximately 300,000 attendees on average per year. Our experiment was run during the rehersal of the final concert (50,000 attendees) and the final concert (100,000 attendees) of the 20th Edition (August 2017). Thus, the gathering event guaranteed high population size and heterogeneity (half of the attendees being non-local, the majority of tourists coming from Northern Italy, age from 14 to over 60, education from primary school to PhD, etc.), with the distribution of the main idiosyncratic features being representative of the Italian population (see Attanasi et al., 2017).

However, our lab-in-the-field experiment is different in spirit to what Harrison and List (2004) term as *artefactual field experiment*, which they define as “. . . the same as a conventional lab experiment but with a nonstandard subject pool.”[1] In line with Charness et al. (2013), we argue that the physical location of the lab is not what defines a method, and laboratory experiments that are run outside of the university are not best described as field experiments. In this regard, the gathering festival that we have chosen for our lab-in-the-field study had a second important feature: the small village where the event was held was transformed for two nights into a huge dance floor with musical contamination, prevalence of non-local attendants, and, more importantly, plenty of collateral events of different nature (social, artistic, entertainment) taking place in the squares around the main stage. Our experiment was framed as one of these events, as a series of tournaments taking place in a gazebo placed at about 300 meters from the main stage, the first tournament starting on the afternoon of the rehearsal of the final concert and the final tournament ending after the end of the final concert (more than 24 hours later). It was advertised in local newspapers and TV news some days before the concert, and across the concert through posters, flyers, and announcements on social networks. With this, although our subjects were aware of participating in an experiment, the naturalistic environment were it was implemented made it seem much more as a game of competition with cards, in the true spirit of the Target-The-Two game of Cohen and Bacdayan (1994).

The rest of the paper is organized as follows. Section 2 introduces the Target-The-Two game, which is at the heart of our experiment. Section 3 describes the experimental design and states the experimental hypotheses. Section 4 presents the results and Section 5 concludes.

[1]As an example of such artefactual field experiment, Harrison and List (2004) discuss the paper by Harrison et al. (2002) who use a standard lab experiment but instead of running it at a university run it in hotels in order to be able to attract a representative sample of the Danish population.

2 Target-The-Two

Target-The-Two (TTT henceforth) is a well-known card game in the field of behavioral and experimental economics. It was first introduced by Cohen and Bacdayan (1994) to study experimentally the behavioral routinization of pairs of card players who cooperate to achieve a common goal. Its properties in terms of detection of human cognition and decision processes have been extensively highlighted by Egidi and Narduzzo (1997).[2] TTT is our prototypical game for the following fundamental feature: it has a simple enough mathematical representation while being interesting and challenging for people (see Figure 1, which represents a deck of our experimental interface). It is appropriate for lab-in-the-field experiments, as on the one hand is easy to explain to subjects who are not familiar with experiments or with economic language, and on the other hand it is sufficiently rich to allow an interesting analysis of decision-making, routinization and creativity.

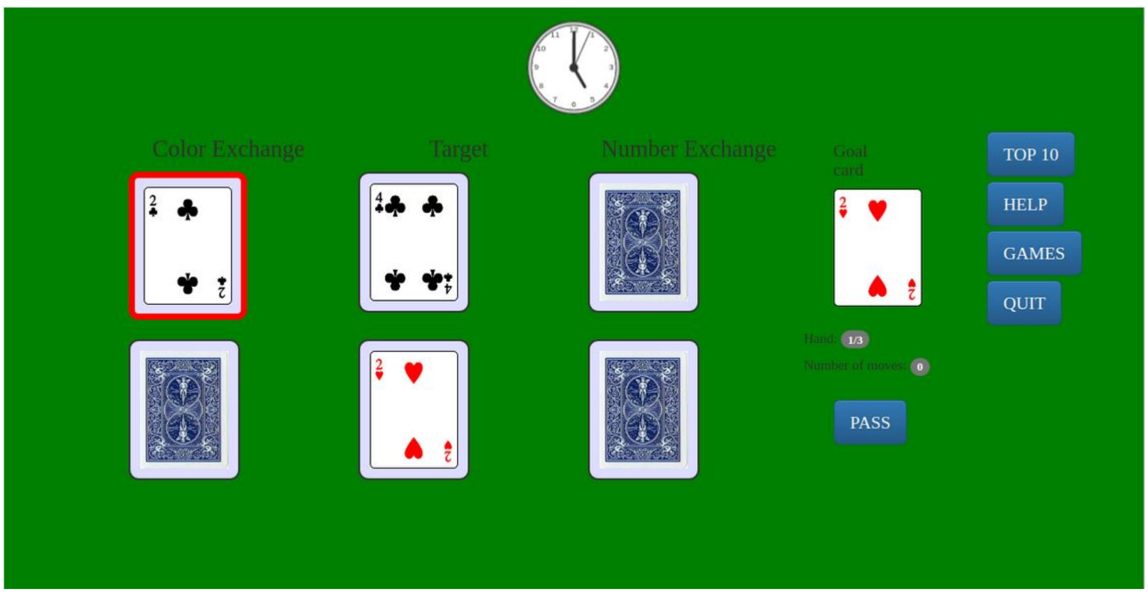

Figure 1: Target-The-Two Game: a deck.

TTT as a coordination game. TTT requires the use of six cards (numbers 2, 3 and 4) and two distinct seeds: Hearts (H) and Clubs (C), and it can be played as a cooperative two-player card game, or in a puzzle form (single player). The aim of the game is to put a goal card in the target position. For half of the treatments, and for the sake of this theoretical description of the game, we let the goal card be the 2 of Hearts (2H). In its original form, with two distinct players, the

[2] For a comprehensive analysis of the relevant literature on the TTT game, see Becker (2001, 2004).

players are called Color-Keeper and Number-Keeper. When cards are dealt, they are positioned in two rows of three cards each. In the top row we have the card in the hand of Color-Keeper, the target, and the card in the hand of Number-Keeper. In the bottom row we have three cards, two of which are covered (left and right), and one facing up (the central one, under the target position). Each player can see only the card in his hand, the card in target and the card facing up (under target). In each hand, the objective of the game is to put the goal card into the "Target" position. A move is any exchange of the player's card with one of the other board card (except for the other player's card, or "PASS": passing the turn). Players' moves to the target position are constrained as follows: Color-Keeper can exchange the card in his hand with the target card only if are of the same color, while the Number-Keeper can exchange the two cards only when they are of the same number.

If we exclude some very elementary card distribution (which we do not use in the experiment), the TTT game admits two different strategies to achieve the goal, each of which is optimal for a different subset of the initial card distributions (Egidi 1997). Using the deck of Figure 1 as example, the first strategy requires first Color-Keeper to search and put in target 2C, and then Number-Keeper to search and put in target 2H; the sequence of cards on target is therefore 4C-2C-2H (hereafter 422). The second strategy, 442, requires first Number-Keeper to search and put in the target 4H, then Color-Keeper to search and put in the target 2H; the sequence of cards on target is therefore 4C-4H-2H (hereafter 442). The two strategies are reciprocally incompatible, because either Color-Keeper puts the 2H in target (442) or Number-Keeper who puts 2H in target (422). Therefore, in the two-player version of the game, coordination is strictly necessary. Coordination occurs because each strategy defines two compatible sub-goals for the pairs, i.e., two different key-cards which trigger their actions in a coordinated way (Figure 2).

	Color-Keeper's sub-goal	Number-Keeper's sub-goal
Strategy 1 (422)	2C	2H
Strategy 2 (442)	2H	4C

Figure 2: Key cards (i.e., sub-goals) for coordinated strategies

In the particular case of Figure 1 it is clear that the first strategy (422) is more efficient than the second one (442), but of course there are different card distributions which reverse the situation by making the 442 strategy more efficient than 422. Both strategies are applicable when the card in the target at the beginning are 3C or 4C (Figure 2). In any other case the hand can be solved by one player only: if 2C is in the target, only Number-Keeper moves (his sub-goal of 422); if 3H or 4H are in the target, only Color-Keeper moves (his sub-goal of 442).

Experiments show that, if the game is played repeatedly, pairs of players jointly learn at least one strategy and become familiar with it (Egidi 1996, Egidi and Narduzzo 1997, Egidi 2001). When they apply the strategy they have learnt, the pairs are clearly acting in a strongly coordinated way; this means that each member of a pair has identified the key-cards of the strategy and make the action required to realize his sub-goal within the joint target (see Figure 2). At this point of the tournament the execution of a strategy becomes largely automatic because the key cards trigger the actions of the pair in a coordinated way.

TTT game and routines. Cohen and Bacdayan (1994) define organizational routines as "patterned sequences of learned behavior involving multiple actors who are linked by relations of communication and/or authority" (1994, p. 555). They consider the occurrence of repeated sequences of action to be the most salient feature of routinized behaviors. As a consequence, in order to verify if an individual's actions fit to a routine, we need to check whether the sequences of actions are repetitive: in our case this means that players are supposed to react with the same sequence of actions if the same distribution of cards appears on the board again. To check routinization according to this definition, we would need to set up an experiment where the same card distributions are presented two or more times: on top of making the design of the experiment heavier, this structure of the experiment could skew the test results by inducing players to biased behaviors, in a situation where there are two alternative strategies.

We can instead verify the application of a strategy in a more general and accurate way simply by comparing the sequences produced by the pairs with the sequences produced according to the division in sub-goals sketched in Figure 2. This means that we do not consider only the sequence of actions but the sequence of rules that a pair must use to produce the actions. In general we can completely define player's behavior through condition-action rules, and in consequence the routines will be detected by checking if players' action patterns are fitting with the action sequences generated by a system of condition-action rules.

Moreover, this approach allows us to measure the degree of routinization of a pair of players in a precise way. According to March and Simon (1958), routinization exists when players do not modify their behavior even when the decision context changes. According with this view (and with the findings of Luchins and Luchins 1959), the pairs who, after the discovery of one strategy, continue to use it for the rest of the tournament even when a better strategy exist, can be defined "routinized".

Therefore, to say that a pair is routinized means that players are locked in one strategy only, and that they cannot discover and use the alternative one when it is more efficient. As a consequence, in this context we consider creativity as measured by the ability to explore widely the space of the game configurations: players who remain locked in one strategy are less creative than players who are able to get out from the lock-in and discover a new strategy.[3]

Routines and automaticity. An interesting question emerges from the features of TTT players' behavior: can the process of "routinization" be attributed only to the players' cooperative interaction or it depends also on the individual cognitive process during the search for a strategy? Cohen and Bacdayan (1994) assume that routinized behaviors are stored as procedural memory, a property which directly relates to the opaque nature of the knowledge embodied in routinized behaviors and their partially inarticulate nature. Their view suggests that the automaticity with which players repeat the same sequences of actions can be explained in terms of automaticity in their mental processes. Studies on the mechanization of thinking - the so-called "Einstellung effect" - have a long tradition in psychology (Luchins 1942, Luchins and Luchins 1950). The literature has suggested that routinized behaviors are based on "routinized thinking", i.e., on the automatic use of "chunks" which enable individuals to save on mental effort (Newell and Simon 1972, Laird et al. 1987, Newell 1990). The experiments by Luchins and Luchins (1959) lead to the conclusion that the routinized behaviors can be explained in terms of bounded rationality or more precisely in terms of the dual model account of reasoning:[4] according to this model, when players have discovered one strategy, the key elements of this familiar strategy will come automatically to their mind (i.e., become more accessible) for the following hands of the game; accessibility then governs players' attention making the application of the familiar strategy easier then the search for a new one. According to this hypothesis the routinized behavior of a pair in the TTT game could be explained as originated also by the cognitive features of the individual learning process.

TTT as puzzle for testing creative abilities. In order to eliminate the strategic interactions among the pairs of players and to verify the individual features of the lock-in processes, we created a puzzle version of the TTT game in which only one player makes both roles in turn. Given the positive results of a preliminary experiment with a group of university students (Egidi 2016), we have decided to verify the emergence of the routinization process in a context of a large number of individuals selected among the attendees of an important European folk festival, using the TTT game puzzle. The main feature of the puzzle game is that it eliminates strategic interactions between players, maintaining all other characteristics of the original TTT game; having an individual rather than a strategic task constrains players to use cognitive abilities for the task solution only, thereby

[3] A more general definition of creativity is related to individual skills in term of mental manipulation, i.e., ability of representing in different way the space of the game configurations and the rules to achieve their goal.

[4] The dualism between the automatic process of recall and the effortful process of symbol manipulation has been deeply explored (see Shiffrin and Schneider 1977, and Kahneman's Nobel lecture). The dual view is based on the evidence that a large part of neural activity is related to automatic processes, which are faster than conscious deliberations and which occur with little or no awareness of effort. As within Simon's analysis, thinking is supposed to be composed of two different cognitive processes: on the one hand a controlled, deliberate, sequential, and effortful process of mental manipulation of items; on the other hand, a non-deliberate, automatic, effortless, and fast process of eliciting mental items from long-term memory.

providing a cleaner measure of individual creativity/routinization. In the puzzle form that we implement in our study, the TTT game has the same rules as in the original version, but with only one player who plays both roles in turn (Color-Keeper and Number-Keeper). Therefore, the constraints on card exchanges are the same, i.e., the card in position “Color” can be exchanged with the card in Target only when the two cards have the same color, while the card in position “Number” can be exchanged with the Target only when the cards number is preserved. The player moves alternatively from the position “Number” or the position “Color” according to a red frame that illuminates the position from which he has to move (in Figure 1, the card in the top left corner). Importantly, the TTT puzzle has exactly the same strategies and properties than the original game.

3 Experimental design and hypotheses

This section describes the experimental procedures (Section 3.1) and design (Section 3.2), and then introduces and discusses the experimental hypotheses (Section 3.3)

3.1 Procedures

The field experiment was run during the 20th edition of the traditional music Festival “La Notte della Taranta”, that takes place each year in August in the most southern part of the Apulia region (South of Italy). The event is among the most important European folk festivals, and attracted approximately 300,000 attendees per year during editions 2012-2017. It consists of 15 itinerant minor concerts (approximately 85,000 attendees per year, with a median of 7,000 attendees per concert) and a final concert (approximately 200,000 attendees per year). For both types of concerts entry is free. The data employed in our analysis was collected during the rehearsal of the final concert (Friday, August 24 2017: 50,000 attendees) and during the final concert itself (Saturday, August 24 2017: 200,000 attendees), where we observe a higher mass-gathering effect and a higher level of individual alcohol consumption. The final concert consists of a one-night huge dance floor, and participants to the concert come from every region of Italy (see Attanasi et al. 2013, 2019).

The sessions were run in a mobile laboratory (a gazebo), installed in the middle of the concert area, with the help of 15 experimental assistants who were continuatively in the field from Friday, August 23 2017 at 5 pm until 5 am of Sunday, August 25 2017. Each session had 20 participants, who played the experimental game on tablets. Each session had a 25 minutes time limit, and the best performer of each session was rewarded with 50€, so that the average payment was 2.5€. Overall, 480 subjects participated in the experiment (61% males, average age 24).

Since alcoholic drinks were extensively available during the festival, each player’s alcohol level was monitored before playing the experimental game. Specifically, we measured Blood Alcohol Concentration (BAC) through electronic breathalyzers (Testmed Safety digital professional alcoholometers).[5]

3.2 Experimental design

In the experiment, subjects played repeatedly the TTT game in its puzzle (solo) version. We selected a set of starting configurations of the board, all of which were solvable with few moves if the same strategy (422 or 442) was played, while the alternative strategy required a higher number of moves.

Structure of the experiment. Before the beginning of the experiment, subjects were asked to fill in a questionnaire about their idiosyncratic features and take an alcohol test, meant to measure their inebriation level. Then, they were assigned a tablet and the experiment began. Subject faced a maximum of 32 hands of TTT in its solo version, and they had a 25 minutes to complete the 32 hands, trying to minimize the number of moves needed to complete the task. The first 16 hands composed the *training phase*. The training phase was designed to teach the player the best strategy to solve the provided hands. In the training phase of the experiment, subjects are exposed to configurations that are more easily solved by one of the two strategies (422 or 442) only, so to familiarize with it. In the *second phase* of the experiment, instead, subjects face a set of configurations where the optimal strategy is not constant: in the last 16 hands, configurations are selected randomly, and therefore, on average, half of the hands can be more easily solved (i.e., through less moves) with the 422 strategy and the other half with the 442 strategy.

Treatments. The experiment has a 2 x 2 x 2 between-subject design, where the treatment variables are (i) the right contextual strategy to solve the game in the training phase (422 vs. 442); (ii) whether the opponent role's cards are uncovered vs. covered (iii); whether the goal card is fixed vs. goal card changes at each hand. We run 3 sessions per treatment, with a total of 60 observations for each treatment.[6]

422 vs. 442 strategy. The first treatment variable is the type of strategy (422 or 442) which is optimal in the training phase (first 16 hands) and on which subjects are therefore trained (in the example of Figure 1, the optimal strategy is 422).

Uncovered vs. covered cards. The second treatment variable is whether the opponent role's cards are uncovered or covered (in the example of Figure 1, the opponent role's cards are covered). Note that in the covered treatment it is more difficult to identify the optimal strategy.

[5]The same type of electronic breathalyzers with the same technique were used to measure BAC for other field studies run during the same event (see Attanasi et al. 2017, 2021c).

[6]Due to technical failure (failed data recording during some experimental sessions), we turned out to have a different number of subjects per treatment, and we had to re-run one session entirely (with different subjects). The average number of subjects with no technical problem per session is 57 (total number = 456), with their minimum number per session being 52.

Fixed vs. changing goal. The third treatment variable is whether the goal card is fixed across different hands, or whether it changes with each hand (in the example of Figure 1, the goal card is the two of hearts).

Payment. In each session, composed by 20 subjects, the winner received 50€. The winner was selected according to the lexicographic order: (i) higher number of solved hands (ii) smaller number of moves (iii) lower amount of time.

3.3 Experimental hypotheses

Let us first state two preliminary hypotheses that do not concern the routinization pattern, but only the effects that idiosyncratic characteristics of the players or of the game may or may not have on their performance. Recall that the winner in each experimental session of 20 participants was the player with the highest number of completed hands, where ties were broken by choosing the player with the lower (total) number of moves, and (if needed) the lower completion time. Therefore, we measure performance in three ways, which constitute the three levels of the above mentioned lexicographic order: number of hands completed, moves/hands ratio and time/hands ratio. The first is a direct measure of performance, while the other two indicators are higher the lower the performance of the player. In fact, our results show a -0.79 and a -0.93 Spearman's rank correlation between number of hands completed and, respectively, moves/hands and time/hands ratio, both significant at the 0.1% level. From now on, with the term "*performance*" in the TTT game we mean all the three indicators: higher number of hands and lower moves/hands and time/hands ratios.

The first preliminary hypothesis states that we expect a worse performance in those treatments that are more complicated to play, i.e., those treatments with a higher cognitive load. In our experiment, we expect treatments with changing goal to be more complicated than treatments with fixed goal. The comparison between covered and uncovered treatments is less trivial: on the one hand the subject has to process less information in covered than in uncovered treatments; on the other hand, it is more complicated to figure out the optimal strategy. Therefore we expect the performance of covered treatments to be better in terms of average time, and possibly worse in terms of average moves. The total effect on the number of hands depends on which of the two aforementioned effects dominates. As for the strategy subjects are trained to, namely 422 vs. 442, given the symmetry of the two strategies, we expect no significant treatment difference. With this, we state **H0.A** as follows:

H0.A: (i) Performance is not affected by the strategy subjects are trained to. (ii) Treatments with changing goal have worse performance than treatments with fixed goal. (iii) Covered treatments bring higher moves/hands ratio and lower time/hands ratio than uncovered ones, with opposite effects on the number of hands.

The second preliminary hypotheses considers the relation between players' individual characteristics and their performance. We have information on age, gender, education and inebriation level. As the task is implemented on young-friendly tablets, the higher familiarity with technology of younger (and possibly male) subjects should have a positive impact on their performance (see, e.g., the Attanasi et al. 2021b, and references therein, as for higher videogame addiction by younger and male subjects). Thus, both age and gender should have a negative impact on subjects' performance. For gender, there is a reinforcement effect in the same direction: the experiment is designed as a tournament, and it is a common finding in the literature that females underperform in competitive environments (see, for example, Gneezy et al. 2003, and Gneezy and Rustichini 2004). All in all, we expect the performance of female participants to be lower than the males' one. Furthermore, as the task requires logic reasoning, we expect education to have a positive effect on performance. For the same reason, we expect inebriation to have a negative effect on performance, as inebriation reduces deductive thinking (Gustafson and Nordlander 1994). With this, we state **H0.B** as follows:

H0.B: Performance is: (i) negatively affected by participants' age; (ii) positively affected by participants' education; (iii) significantly worse for female than for male participants; (iv) negatively affected by inebriation.

Now, before we discuss the experimental hypotheses on learning in the training phase and routinization in the second phase of the TTT game, we introduce our classification of the subjects in these two dimensions. First of all, we note that we cannot define an agent routinized on the basis of the behavior in the first 16 hands, as using the same strategy in the first 16 hands is the optimal behavior. Rather, the training phase allows us to investigate the learning process of the subjects, and their (acquired) expertise in the game. Therefore, we call *expert* those subjects who implemented the right contextual strategy during at least 75% of the hands of the training phase, and *inexpert* those subjects who did not manage to learn the optimal contextual strategy in the training phase, implementing it in less than 75% of the hands of this phase. To investigate routinization, instead, we have to focus on the second phase of the game, in which the optimal strategy varies across hands, so that we can distinguish subjects who routinize and always implement the same strategy from creative subjects who (try to) adapt their strategy to the specific problem they have to solve. However, we note that not all subjects manage to reach the second phase. We call *early drop outs* those subjects who completed a number of hands strictly lower than 17, i.e., those subjects who did not complete hands in the second phase. Among those subjects who reach the second phase, we call (practically) *routinized* the subjects who in the second phase implemented the contextual strategy they have learned in the training phase for at least 75% of the hands completed in this phase. Finally, we call (practically) *creative* the subjects who in the second phase implemented the contextual strategy learned in the training phase in less than 75% of the hands completed in this phase.

We are now ready to start discussing our main experimental hypotheses. First of all, we expect both learning and routinization to be task-dependent (Ohly et al. 2006). Specifically, we expect learning to be lower in more difficult tasks. We instead expect routinization to increase in more difficult tasks for a twofold reason: first, it is more complicated for a subject to understand that a different strategy may be optimal in the new context; second, difficult tasks may induce more time pressure on the subject, thereby increasing routinization, consistently with the findings of Ohly et al. (2006). We can therefore state **H1** as follows:

H1: (i) Learning in the training phase and routinization in the second phase are both independent of the strategy subjects are trained in. (ii) Learning is higher in treatments with fixed goal, and routinization is higher in treatments with changing goal. (iii) Learning is higher in uncovered treatments, and routinization is higher in covered treatments.

While we expect learning and routinization to be task-dependent, we expect them to be independent of personal traits such as age, gender and education. As a matter of fact, as shown in Chae and Choi (2019), correlation between routinization and age, gender or education is small and not significantly different from zero. There is instead evidence that alcohol consumption affects creativity and routinization. Norlander (1999) contains a review of the literature on the links between creativity and inebriation, and shows how alcohol may increase or decrease creativity depending on the specific type of creativity and on the phase of the creative process that we consider. We think that, in the literature discussed there, the most relevant result given our decision task is the one in Gustafson and Norlander (1994), who suggest that inebriation impair deductive reasoning. For this reason, we think that inebriated subjects should be less able than sober subjects to get expertise in the training phase and to routinize in the final phase. Therefore, we can state **H2** as follows:

H2: (i) Learning in the training phase and routinization in the second phase are both: (i) independent of personal traits: age, gender and education; (ii) negatively dependent on the current state of the subject: inebriation.

The adoption of one strategy for all hands decreases the decision time of agents, by reducing the set of reasonable moves. As a consequence, it surely increases performance in the training phase, where expert subjects choose the correct contextual strategy faster. The effect of routinization on the second phase is more controversial. On the one hand, routinized agents are faster: they know where they have to go, since they are convinced that there is only one strategy to implement (the one they have been trained to); on the other hand, this is not always the right contextual one in the second phase, and therefore they might have a higher moves/hands ratio than creative subjects, who instead try to get the right strategy for each of the 16 hands of the second phase. Therefore, we formulate **H3** as follows:

H3: (i) Learning in the training phase is efficient: expert players perform better than inexpert ones. (ii) Routinization in the second phase has opposite effects on performance: routinized players are faster (lower time/hands ratio) but they have a higher moves/hands ratio then creative ones. If the first effect predominates, then routinization is efficient.

4 Results

We begin our analysis of the experimental data by testing the two preliminary hypotheses that consider the effects of game characteristics, individual characteristics and individual behavior on performance.

We start from hypothesis **H0.A**, which considers the effects of the treatment variables on performance. Recall, from Section 3.3, that the three measures of performance in our experiment are: number of hands, moves/hands ratio, and time/hands ratio. A subject's better performance in the TTT game is characterized by higher number of hands and lower moves/hands ratio and time/hands ratio.

First, we note that training a subject on 422 or on 442 makes no significant difference in terms of performance. Indeed, the average number of hands (24.26 vs. 23.73, p-value = 0.845), the average moves/hands ratio (15.14 vs. 12.09, p-value = 0.568) and the average time/hands ratio (125.88 vs. 98.34 seconds, p-value = 0.755) are not significantly different between treatments according to the strategies subjects have been trained to (Mann-Whitney test of equality of medians between two independent sampled populations). Hence, **H0.A.i is verified**.

We now move to compare the treatments with fixed goal and the treatments with changing goal. As discussed in Section 3.3, we expect subjects in treatments with changing goal to have on average worse performances. Specifically, we find that treatments with changing goal have lower average number of hands (23.10 vs 24.87, p-value = 0.013), higher moves/hands ratio (14.81 vs. 12.49, p-value = 0.002), and higher time/hands ratio (114.69 vs. 109.86 seconds, p-value = 0.002) than treatments with fixed goal. Hence, **H0.A.ii is verified**.

Finally, we consider the effect of covering the opponent role's cards. Consistently with H0.A.iii, we find a higher moves/hands ratio in covered treatments than in uncovered ones (14.16 vs. 13.04), but the difference is not significant (p-value = 0.774). Again consistently with H0.A.iii, we detect a significantly lower average time/hands ratio in covered than uncovered treatments (97.94 vs. 127.83, p-value < 0.001). Recall that H0.A.iii is silent as for which of the two opposite effects – higher moves/hands ratio vs. lower time/hands ratio – would prevail. Having detected only the second effect to be significant, which goes in the direction of a better performance by covered treatments, it is not surprising that average number of hands is significantly higher in covered than uncovered treatments (25.60 vs. 22.26, p-value < 0.001). This leads us to conclude that in the (less difficult) uncovered treatments subjects complete a lower number of hands because they think for a longer time, and hence they are slower. With this, we conclude that **H0.A.iii is mostly verified**.

All in all, we can conclude that **H0.A is verified**, with the exception of a non-significant difference in the moves/hands ratio between covered and uncovered treatments, because the expected positive difference is not a statistically significant one.

We now proceed to test **H0.B**. Consistently with H0.B.i, we find that age has a slight negative effect on subject's performance. In fact, the Spearman's rank correlation index is significantly negative between age and number of hands (-0.09, p-value = 0.055), non-significant between age and moves/hands ratio (-0.05, p-value = 0.304) and significantly positive between age and times/hands ratio (0.10, p-value = 0.047). With this, we conclude that **H0.B.i is mostly verified**.

As for the effect of education on performance, **H0.B.ii is only partially verified**. In fact, we report a significantly negative effect on moves/hands ratio (-0.11, p-value = 0.023) and essentially no effect on number of hands and time/hands ratio (respectively, 0.04, p-value = 0.392; -0.02, p-value = 0.620).

As for gender, **H0.B.iii is strongly verified**, instead. Indeed, females complete a lower number of hands (22.85 vs. 24.69) with higher moves/hands ratio (13.78 vs. 13.72) in a higher time/hands ratio (118.31 vs. 111.65 seconds), all these differences being highly significant (p-value equal respectively to 0.019, 0.004, 0.003).

Finally, as for inebriation, only 32.46% of our subjects were found with a positive BAC, which explains the low average BAC = 0.09 g/l across the whole sample. Therefore, the overwhelming majority of the experimental participants were sober, while inebriated subjects reported an average BAC = 0.33 g/l, which is below the legal amount for driving established by the Italian legislation (0.50 g/l). We detect no correlation at all between subjects' BAC and their performance in the TTT game (Spearman's rho = 0.02, p-value = 0.724 for number of hands; Spearman's rho = 0.01, p-value = 0.810 for moves/hands ratio; Spearman's rho = 0.00, p-value = 0.990 for time/hands ratio). We find a slight negative correlation if we focus on inebriated subjects (i.e., with BAC>0), but the correlation index is usually not significant: Spearman's rho = -0.14, p-value = 0.142 for number of hands; Spearman's rho = 0.21, p-value = 0.025 for moves/hands ratio; Spearman's rho = 0.12, p-value = 0.223 for time/hands ratio. This result is confirmed by the fact that no significant difference in performance is found between sober (i.e., with BAC=0) and inebriated subjects, regardless of the measure of performance (lowest p-value = 0.553 for average number of hands). With this, we can state that **H0.B.iv is not verified**. We consider this result as relevant for the following analysis, since it allows us to consider the whole sample of participants – rather than the subsample of sober ones – when testing our experimental hypotheses on learning and routinization.

We now turn to the analysis of learning in the training phase and routinization in the second phase of the TTT game, by testing our main hypotheses H1-H3. First, let us recall that learning in the training phase is captured by a dummy variable that classifies a subject as *expert* (value 1) if he chooses the right contextual strategy (according to treatments 422 or 442) in at least 75% of the hands he plays out of the first 16 (i.e, in the training phase), and *inexpert* (value 0) otherwise. We also recall that routinization and creativity in the second phase are captured by a dummy variable that classifies a subject as *routinized* (value 1) (resp., *creative*, value 0) if he chooses, in at least (resp., less than) 75% of the hands he plays in the second phase of the experiment, the contextual strategy he has been trained to use in the training phase (i.e., according to treatments 422 or 442).[7]

With this, we can test **H1** by relying on the two dummy variables *expert* and *routinized.* The detected share of *expert* subjects is low (10.09%), and significantly higher in 422 than in the 442 treatments (14.35% vs. 5.75%, p-value = 0.002, χ^2 test). We interpret this finding as a signal that strategy 442 is more difficult to learn, so that subjects needed a higher number of hands to converge on it. However, and possibly for the same reason, once they have learned strategy 442, they routinize on it more strongly. In fact, we note that the fraction of subjects completing more than 16 hands of the game (i.e., overcoming the training phase) is not significantly different between the 422 and the 442 treatments (75.22% vs. 70.35%, p-value = 0.243, χ^2 test). However, among those who reached at least hand 17, we find a significantly lower share of *routinized* subjects in 422 than in the 442 treatments (19.08% vs. 34.59%, p-value = 0.001, χ^2 test).

With this, we can state that **H1.i is not verified**. This is represented in Figure 3, where we report the share of subjects not overcoming the training phase (namely, *early drop out* subjects), the share of experts and the share of routinized subjects. Note that routinized subjects belong to the complementary subsample of subjects who did overcome the training phase. H1.i is not verified since, despite a similar share of early drop outs in 422 and 442 treatments, the share of expert (resp., routinized) subjects is significantly higher in treatment 422 (resp., 442).

We now turn to the second treatment variable, by comparing treatments with fixed goal to treatments where the goal card is changing (H1.ii). The hypothesis that the share of expert subjects is higher in treatments with fixed goal than in treatments with changing goal is verified (12.50 vs. 7.59, p-value = 0.082). The hypothesis that the share of routinized subjects is higher in treatments with changing goal than in treatments with fixed goal is instead not verified. As a matter of fact, the share of routinized subjects is significantly higher in fixed than in changing treatments (32.39

[7] Note that given our definition of routinized and creative subjects, we can only classify subjects as routinized or creative if they play at least 17 hands.

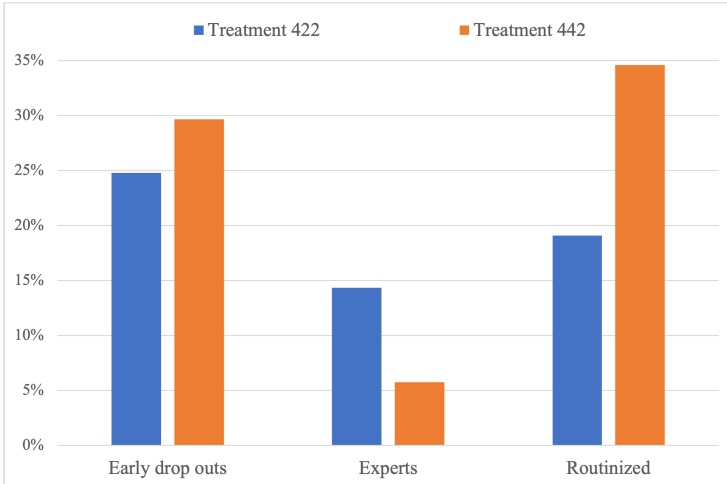

Figure 3: Share of early drop out, expert and routinized subjects in 422 vs. 442 treatments

vs. 19.87, p-value = 0.010). As such, **H1.ii is verified for learning but contradicted for routinization**, as it is shown in Figure 4, which also reports the share of early drop outs in the two treatments. However, we recall that, given the nature of the TTT game, we can only observe subjects' routinization if they managed to learn the routine in the training phase. We believe that the explanation of the comparison between fixed and changing treatments has to be found in the fact that the changing treatment is so complicated that only few subjects managed to learn the correct contextual strategy in the training phase, as it is shown by the negligible share of expert subjects (less than 8%) and the high share of early drop outs (more than 30%) in the treatments with changing goal.

Finally, let us investigate the effects of the last treatment variable on learning and routinization, by comparing treatments where the opponent role's cards are covered to treatments where they are uncovered (H1.iii). Figure 5 shows the same (low) share of expert subjects between uncovered and covered treatments (10.09 vs. 10.08, p-value = 1.000), and a non-significant difference in the share of routinized subjects (28.87 vs. 24.74, p-value = 0.398). Hence, we conclude that **H1.iii is not verified.** This confirms our findings on performance (test of H0.A.iii), which suggested that we cannot conclude that covered treatments are cognitively more complicated than uncovered ones, as, on the one hand, they make it harder to figure out the optimal strategy but on the other hand they require the subject to process less information. And in fact, as Figure 5 shows, the share of early drop outs is significantly higher in uncovered than in covered treatments (34.86% vs. 20.17%, p-value < 0.001).

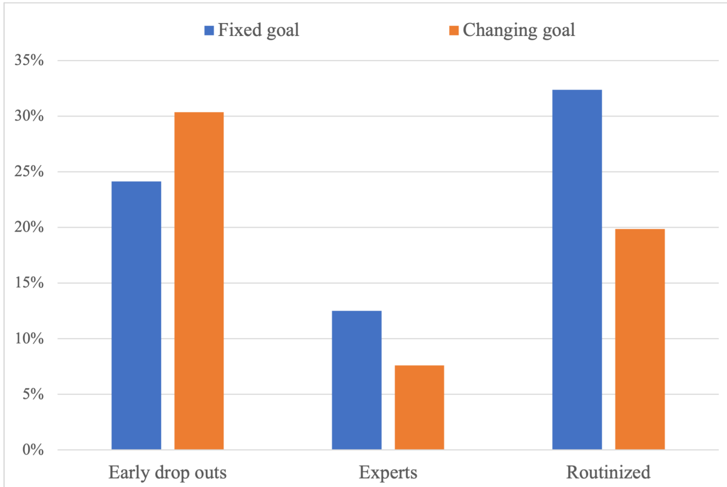

Figure 4: Share of early drop out, expert and routinized subjects in fixed vs. changing goal

We now proceed to test **H2**, i.e., the potential correlations between learning and routinization on the one side and subjects' idiosyncratic features on the other side. We detected a slight negative effect of age on routinization (Spearman's rho $= -0.10$, p-value $= 0.068$), while no effect was found for learning (Spearman's rho $= -0.02$, p-value $= 0.737$). Therefore, it seems that an older age leads to more routinization in the second phase: once learned the right contextual strategy, older subjects stuck to it more, without trying anything different in the hands where this was needed to reduce the number of moves. As expected, education had no effect on either learning or routinization (respectively, Spearman's rho $= -0.01$, p-value $= 0.961$; -0.01, p-value $= 0.886$). Also gender had no effect on routinization in the second phase (χ^2 test, p-value $= 0.376$), while males shew significantly more learning than females in the training phase (11.99% vs. 6.43%, p-value $= 0.057$, χ^2 test), which is consistent with their better performance (test of H0.B.iii). With this, we can conclude that **H2.i is partially verified**.

As for the effect of inebriation, the correlation between BAC and both learning and routinization is negative (respectively, Spearman's rho $= -0.08$ and -0.10) and significant only for routinization (respectively, p-value $= 0.103$ and 0.077). Focusing only on inebriated subjects, we find again a negative correlation of BAC with both learning and routinization (respectively, -0.16 and -0.09), significant only for the former (respectively, p-value $= 0.092$ and 0.406). A Mann-Whitney test confirms that inebriated subjects routinized significantly less than sober ones in the second phase of the game (19.77% vs. 29.09%, p-value $= 0.097$). They also made less learning in the training phase, but not significantly so (6.36% vs. 11.04%, p-value $= 0.158$). With this, we conclude that **H2.ii is verified**.

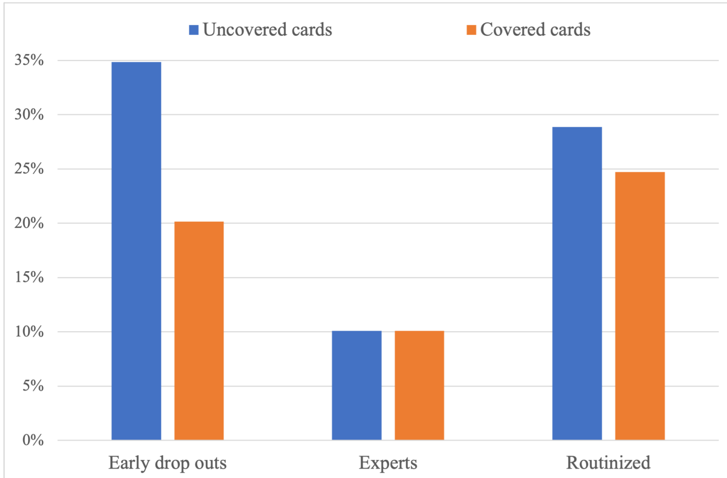

Figure 5: Share of early drop out, expert and routinized subjects in uncovered vs. covered cards

We conclude the analysis with the test of **H3**, which deals with the performance of expert vs. inexpert subjects in the training phase (H3.i) and of routinized vs. creative subjects in the second phase (H3.ii). **H3.i is strongly verified**: expert subjects completed a significantly greater number of hands than inexpert ones (31.37 vs. 23.18, p-value < 0.001), and presented a significantly lower moves/hands ratio (6.04 vs. 14.48, p-value < 0.001) and time/hands ratio (32.12 vs. 121.22 seconds, p-value < 0.001).

The fact that expert subjects perform better than inexpert ones in terms of (lower) moves/hands ratio informs us that inexpert subjects are kind of lost in their effort to find the right strategy in the training phase, this strategy being the same regardless of the first 16 hands of the game. For this reason, when testing H3.ii as for the effect of routinization on performance, we disentangle expert and inexpert subjects. With this, we rely on four mutually-exclusive categories of players: inexpert creative (67.47% of the sample), inexpert routinized (26.72% of the sample), expert creative (8.62% of the sample), and expert routinized (11.21% of the sample). These four categories disentangle the subsample of subjects overcoming the 16-hand training phase, i.e., completing at least 17 hands of the game (72.81% of the sample). The residual category of early drop outs in Figures 3-5 (27.19% of the sample) is not considered in the test of H3.ii. Figure 6a reports two out of the three measures of performance of the four categories of subjects overcoming the 16-hand training phase, namely moves/hands ratio and time/hands ratio. Figure 6b reports the same two measures for the same four categories of subjects, conditionally on having completed the whole 32 hands of the TTT game. We highlight that the subsample of subjects in Figure 6b is modal among those subjects overcoming the training phase. In fact, 62.95% of inexpert creative, 83.87% of inexpert routinized, 90% of expert creative and 100% of expert routinized completed the 32 hands of the TTT game.

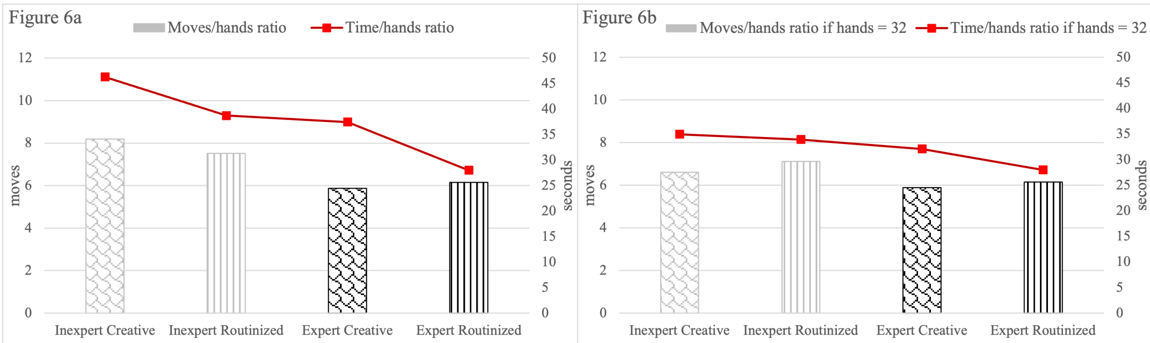

Figure 6: Performance of creative vs. routinized subjects, disentangled by expertise

From Figure 6a it is easy to notice that, among inexpert subjects, routinized subjects perform significantly better than creative ones in terms of lower time/hands ratio (38.72 vs. 46.29, p-value = 0.001), which confirms H3.ii, but not significantly worse in terms of higher moves/hands ratio (7.52 vs. 8.19, p-value = 0.291), which is inconsistent with H3.ii. Given that the first effect is predominant, inexpert routinized subjects complete a significantly higher number of hands than inexpert creative ones (30.77 vs. 28.95, p-value = 0.002). If we focus on subjects completing the whole TTT game (Figure 6b), H3.ii is verified for moves/hands ratio but not for time/hands ratio: inexpert routinized and inexpert creative subjects show a similar time/hands ratio (33.94 vs. 34.97, p-value = 0.001) and the former performs significantly worse in terms of higher moves/hands ratio (7.11 vs. 6.61, p-value = 0.096). With this, we conclude that **H3.ii is partially verified for inexpert subjects**.

Moving to expert subjects, Figure 6a shows that routinized subjects perform significantly better than creative ones in terms of lower time/hands ratio (28.01 vs. 37.46, p-value = 0.015), and worse in terms of higher moves/hands ratio, although not significantly so (6.16 vs. 5.88, p-value = 0.258). Both results are consistent with H3.ii, and are confirmed if focusing on expert subjects completing the whole TTT game (Figure 6b): the positive gap for time/hands ratio and negative gap for moves/hands ratio are respectively similar to those detected in the general case of Figure 6a, with only the first gap being significant (respectively, p-value equal to 0.045 and 0.272). With this, we conclude that **H3.ii is verified for expert subjects**.

All in all, our test of H3 seems to show that learning the suggested contextual strategy in the training phase generates a significant gap in the performance in the second phase (H3.i). However, for expert subjects, routinizing in this (unique) strategy has no impact on the number of completed hands, and decreases performance as for moves/hands ratio, although not significantly so. Furthermore, for inexpert subjects, although significantly increasing the number of completed hands, routinization has no impact on the moves/hands ratio and, among those subjects completing the whole game, it even has a significant negative impact on the moves/hands ratio. In one sentence, given that 71.39% of our subjects overcoming the training phase were able to complete the whole game, routinization decreased overall the likelihood to win the game (H3.ii). Indeed, in each of the 24 sessions of our TTT game, what counted for the final ranking was the moves/hands ratio, since usually more than half of the subjects in a session completed the 32 hands (51.97% on average across all sessions).[8]

5 Conclusions

Understanding the process by which individuals routinize, and whether routinization is efficient in completing difficult tasks is relevant for most organizations.

In the paper, we tackle this issue through a lab-in-the-field experiment in which we first train a subject to solve a decision problem, the TTT game in its puzzle form, with one optimal strategy, and then we propose to the same subject different versions of the puzzle where the strategy that was used in the training phase may no longer be optimal. Our analysis focuses on this second phase of the game, and on the effects of a creative approach vs. routinization.

Our experiment is built to test the hypothesis that by exposing players to a set of preliminary hands characterized by starting configurations all easily solved by the same strategy, they would be "induced" to discover this solution more easily than the alternative one and to memorize it more deeply. This mechanism is strengthened by exogenously imposing time pressure, which discourages subjects from exploring the strategy space, and therefore favors routinization.

We find that routinization and learning are mostly independent of idiosyncratic features of individuals, with the only exceptions being a slightly more pronounced tendency to routinize in older subjects, and a faster learning for males.

As for the efficiency of routinization, we find that the effects of routinization on performance depend on whether subjects are expert or not. Notably, we find that for expert subjects there is no difference between the performance of routinized and creative subjects. A difference is instead present in those subjects who were slow at learning the correct strategy in the training phase. Among inexpert subjects, we find that routinized subjects are faster, so that their average number of completed hands is higher than the creative subjects' one, but so is the moves/hands ratio. Overall, this reduces their likelihood of winning the game, as most winners tied in terms of number of hands (completing the maximum number of 32), so that the winner was selected according to the lower moves/hands ratio criterion.

In the paper, we also look at three treatment manipulations of the TTT game: type of suggested

[8] As one could expect, it never occurred that two subjects completing the 32 hands in the same session also had the same moves/hands ratio, hence a lower time/hands ratio was never useful to win the game.

strategy in the training phase, fixed vs. changing goal, and covered vs. uncovered opponent role's cards. Among the treatment effects, there is one which represents a particularly interesting test of creativity, namely the fixed vs. changing goal treatment. As a matter of fact, in the presence of a changing goal, both in the learning phase and in the routinization phase subjects, in order to replicate the same strategy for a different goal, need to change the representation of the problem and of the strategy itself. This is in itself a higher form of creativity, which creates a new and richer structure of rules in the mind of subjects, in the spirit of Simon (1969, pp. 94-98). In this interpretation, learning in the training phase of the changing goal treatment may itself be considered creative, and is therefore more difficult than learning in other treatments, which is consistent with our findings.

References

[1] Attanasi, G., Casoria, F., Centorrino, S., and Urso, G. (2013). Cultural investment, local development and instantaneous social capital: A case study of a gathering festival in the South of Italy. Journal of Socio-Economics, 47, 228-247.

[2] Attanasi, G., Bortolotti, S., Cicognani, S., Filippin, A. (2017). The drunk side of trust: Social capital generation at gathering events (No. 2017-21). Bureau d'Economie Théorique et Appliquée, UDS, Strasbourg.

[3] Attanasi, G., Passarelli, F., Urso, G., and Cosic, H. (2019). Privatization of a tourism event: Do attendees perceive it as a risky cultural lottery? Sustainability, 11, 2553.

[4] Attanasi, G., Chessa, M., Gil Gallen, S., and Llerena, P. (2021a). A survey on experimental elicitation of creativity in economics. Revue d'Economie Industrielle, 174, 273-324.

[5] Attanasi, G., Maffioletti, A., Shalukhina, T., Bel, C., and Cherikh, F. (2021b). Gender differences in the impact of COVID-19 lockdown on potentially addictive behaviors: An emotion-mediated analysis. Frontiers in Psychology, 12, 703897.

[6] Attanasi, G., Cox, J., and Sadiraj, V. (2021c). Festival Games: Inebriated and Sober Altruists. Typescript.

[7] Becker, M. C. (2001). Empirical research on routines - The state of the art and its integration into the routines. Nelson and Winter Conference.

[8] Becker, M. C. (2004). Organizational routines: a review of the literature. Industrial and corporate change, 13, 643-678.

[9] Charness, G., and Grieco, D. (2019). Creativity and financial incentives. Journal of European Economic Association, 17, 454-496.

[10] Charness, G., Gneezy, U., Imas, A. (2013). Experimental methods: Eliciting risk preferences. Journal of Economic Behavior and Organization, 87, 43-51.

[11] Chae, H., and Choi, J. N. (2019). Routinization, free cognitive resources and creativity: The role of individual and contextual contingencies. Human Relations, 72, 420-443.

[12] Cohen, M. D., and Bacdayan, P. (1994). Organizational routines are stored as procedural memory: Evidence from a laboratory study. Organization Science, 5, 554-568.

[13] Egidi, M. (1996). Routines, Hierarchies of Problems, Procedural Behaviour: Some Evidence from Experiments. In: The rational foundations of economic behaviour , Arrow, K.J., et al., editors, St. New York: Martin's Press, London: Macmillan Press, in association with the International Economic Association, 303-333.

[14] Egidi, M. (2001). Biases in Organizational Behavior. In: The Economics of Choice, Change and Organization: Essays in Memory of Richard M. Cyert, Augier M., March J. J., editors. Aldershot U.K.: Elgar, 190-242.

[15] Egidi, M. (2016). Organizational Decisions in the Lab. Cap 11 in Frantz R., Marsh L. Minds Models and Milieux, Commemorating the Centennial of the Birth of Herbert Simon, Palgrave Macmillan, pp. 186-206.

[16] Egidi, M., and Narduzzo, A. (1997). The emergence of path-dependent behaviors in cooperative contexts. International Journal of Industrial Organization, 15, 677-709.

[17] Gneezy, U., and Imas, A. (2017). Lab in the field: Measuring preferences in the wild. In Handbook of economic field experiments (Vol. 1, pp. 439-464). North-Holland.

[18] Gneexy U., Niederle, M., and Rustichini, A. (2003) Performance in competitive environments: gender differences. Quarterly Journal of Economics, 118, 1049-1074.

[19] Gneezy, U., and Rustichini, A. (2004) Gender and competition at a young age. American Economic Review, 94, 377-381.

[20] Gustafson, R. and Norlander, T. (1994). Effects of alcohol on persistent effort and deductive thinking during the preparation phase of the creative process. Journal of Creative Behavior, 28, 124-132.

[21] Harrison, G. W., Lau, M. I., and Williams, M. B. (2002). Estimating individual discount rates in Denmark: A field experiment. American Economic Review, 92, 1606-1617.

[22] Harrison, G. W., and List, J. A. (2004). Field experiments. Journal of Economic literature, 42, 1009-1055.

[23] Laird, J. E., Newell, A., and Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. Artificial intelligence, 33, 1-64.

[24] Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. Psychological Monographs, 54(6), i–95.

[25] Luchins, A. S., and Luchins, E.H. (1950). New experimental attempts at preventing mechanization in problem solving. Journal of General Psychology, 42, 279–297.

[26] Luchins, A. S., and Luchins, E. H. (1959). Rigidity of behavior: A variational approach to the effect of Einstellung. Oxford, England: University of Oregon Press.

[27] March, J. G. and Simon, H. A. (1958). Organizations. Wiley.

[28] Newell, A. (1990). Unified theories of cognition. Harvard University Press.

[29] Newell, A., and Simon, H. A. (1972). Human problem solving (Vol. 104, No. 9). Englewood Cliffs, NJ: Prentice-hall.

[30] Norlander, T. (1999). Inebriation and Inspiration? A Review of the Research on Alcohol and Creativity. Journal of Creative Behavior, 33, 22-44.

[31] Ohly, S., Sonnentag, S., and Pluntke, F. (2006). Routinization, work characteristics and their relationships with creative and proactive behaviors. Journal of Organizational Behavior, 27, 257-279.

[32] Shiffrin, R. M., and Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. Psychological Review, 84, 127.

[33] Simon, H. A. (1969). The sciences of the artificial. The MIT Press.

[34] Simonton, D. K. (2000). Creativity: Cognitive, personal, developmental, and social aspects. American Psychologist, 55, 151–158.