CIRANO
Knowledge into action

CREATIVE COGNITION AS A BANDIT PROBLEM

NOÉMI **BERLIN**
JAN **DUL**
MARCO **GAZEL**
LOUIS **LÉVY-GARBOUA**
TODD **LUBART**

2023s-11
WORKING PAPER

CS

ISSN 2292-0838 (online version)

Creative Cognition as a Bandit Problem

Noémi Berlin[*], Jan Dul[†], Marco Gazel[‡], Louis Lévy-Garboua[§] and Todd Lubart[**]

Abstract/Résumé

This paper characterizes creative cognition as a multi-armed bandit problem involving a trade-off between exploration and exploitation in sequential decisions from experience taking place in novel uncertain environments. Creative cognition implements an efficient learning process in this kind of dynamic decision. Special emphasis is put on the optimal sequencing of divergent and convergent behavior by showing that divergence must be inhibited at one point to converge toward creative behavior so that excessive divergence is counterproductive. We test this hypothesis in two behavioral experiments, using both novel and well-known tasks and precise measures of individual differences in creative potential in middle and high school students. Results in both studies confirmed that a task-dependent mix of divergence and convergence predicted high performance in a production task and better satisfaction in a consumption task, but exclusively in novel uncertain environments. These predictions were maintained after controlling for gender, personality, incentives, and other factors. As hypothesized, creative cognition was shown to be necessary for high performance under the appropriate conditions. However, it was not necessary for getting high grades in a traditional school system.

Cet article caractérise la cognition créative comme un problème de bandit-manchot à bras multiples impliquant un compromis entre exploration et exploitation dans les décisions séquentielles basées sur l'expérience et se déroulant dans de nouveaux environnements incertains. La cognition créative met en œuvre un processus d'apprentissage efficace dans ce type de décision dynamique. L'accent est mis sur l'enchaînement optimal des comportements divergents et convergents en montrant que la divergence doit être inhibée à un moment donné pour converger vers un comportement créatif, de sorte qu'une divergence excessive est contre-productive. Nous testons cette hypothèse dans deux expériences comportementales, en utilisant à la fois des tâches nouvelles et bien connues et des mesures précises des différences individuelles dans le potentiel créatif d'élèves de collège et de lycée. Les résultats des deux études ont confirmé qu'un mélange de divergence et de convergence dépendant de la tâche permettait de prédire une performance élevée dans une tâche de production et une meilleure satisfaction dans une tâche de consommation, mais exclusivement dans des

[*] CNRS, EconomiX, Université Paris Nanterre, noemi.berlin@parisnanterre.fr, Université Paris Nanterre, Bât. G - Bureau G605a, 200 Avenue de la République 92000 Nanterre, France
[†] Rotterdam School of Management, Erasmus University, the Netherlands, jdul@rsm.nl
[‡] NEOMA Business School, marco.gazel@neoma-bs.fr
[§] **Corresponding author**: Université Paris 1 Panthéon-Sorbonne and Paris School of Economics, Centre d'Economie de la Sorbonne, and CIRANO (Montreal), louis.levy-garboua@univ-paris1.fr 106-112, bd. de l'Hôpital 75013 Paris.
[**] Université de Paris Cité, LaPEA (Laboratoire de Psychologie et d'Ergonomie Appliquées), todd.lubart@u-paris.fr

environnements nouveaux et incertains. Ces prédictions ont été maintenues après contrôle du sexe, de la personnalité, des incitations et d'autres facteurs. Comme nous l'avions supposé, la cognition créative s'est avérée nécessaire pour obtenir des performances élevées dans les conditions appropriées. Cependant, elle n'est pas nécessaire pour obtenir de bonnes notes dans un système scolaire traditionnel.

Keywords/Mots-clés: Creative cognition; Multi-armed bandit problem; Exploration-exploitation trade-off; Individual differences in creative potential; Adolescents' behavior; Necessary Condition Analysis / Créativité cognitive ; Problème du bandit à plusieurs bras ; Compromis exploration-exploitation ; Différences individuelles de potentiel créatif ; Comportement des adolescents ; Analyse des conditions nécessaires.

Requests for the data or materials can be sent to the corresponding author at: : louis.levy-garboua@univ-paris1.fr

The research was conducted with French participants who provided informed consent with appropriate school and parental authorization. The protocol was designed with respect to the Helsinki convention and the French national research center (CNRS) ethical committee guidelines.

Declarations of interest: none.

Pour citer ce document / To quote this document

Berlin, N., Dul, J., Gazel, M., Lévy-Garboua, L., & Lubart, T. (2023). Creative Cognition as a Bandit Problem (2023s-11, Working Papers, CIRANO.) https://doi.org/10.54932/ANRE7929

1. Introduction

Creativity is consensually defined as the ability to produce work that is both novel and appropriate (e.g., Runco & Jaeger, 2012; Lubart et al., 2015). In popular representations, it is commonly attributed to exceptional thinkers, artists and scientists who "created" something new and useful out of apparently nothing, as opposed to ordinary people who routinely "produce" outputs out of existing inputs. However, creative cognition is now recognized as a basic mode of thinking that individuals may engage in various activities and levels (see, for example, the four-c model of creativity of Kaufman & Beghetto, 2009). Creativity is currently considered one of the four main "21st century skills", together with critical thinking, collaboration, and communication (http://p21.org), and was further identified in 2020 by the World Economic Forum as one of the top three abilities for future employability.

The creative process has been conceptualized in terms of an interplay between divergent and convergent thinking (see Lubart (2000) for a review of creativity models). Although divergence has typically been emphasized in the creativity literature (Guilford 1967), several process models in the creative cognition approach and related work (see Mumford et al., 1991; Finke et al, 1992; Ward, 2007; Smith et al., 2009) have identified a set of micro processes, underlying divergence and convergence, which combine in dynamic sequences. The divergent phase provides an extensive search of the conceptual space, or costly exploration of yet unknown solutions. The convergent phase involves elaboration and integration of collected information toward an appropriate solution to a problem. This process requires a synthesis of ideas but also pruning opportunities to exploit the better ideas and pursue an optimized path. It involves an implementation of the exploration-exploitation tradeoff describing optimal sequential decisions from experience in uncertain environments.

Our specific source of inspiration in this article is the *multi-armed bandit problem* in which a player who is facing a number of slot machines, each with an unknown distribution of payoffs, must repeatedly select one machine at a time to eventually identify the wealth-maximizing arm. This problem of sequential decision-making under uncertainty, combining exploration and exploitation, characterizes many problem-solving situations, be they novel or habitual (e.g., Erdem, T., & Keane, M. P. 1996, Lipshitz and Strauss 1997). Our contribution in this article will be to characterize creative cognition as a kind of multi-armed bandit problem encountered in novel and uncertain environments.

The *sequence* of divergence followed by convergence is an essential part of the exploration-exploitation trade-off and of the creative process. A major, though somewhat neglected, implication of the sequential nature of creative cognition is that divergence must be inhibited at one point to converge toward creative behavior. Consequently, excessive divergence should be counterproductive, notably in a simple and stable environment in which there is nothing for which to search. By contrast, intensive exploration would be needed to find solutions to novel issues or to complex problems in uncertain environments; and subsequent exploitation would also be needed to adopt these innovations early.

The optimal sequencing of divergence-exploration and convergence-exploitation will be a central issue of our research that will find confirmation in two unrelated studies involving different contexts of production and consumption. To show that an optimal trade-off between divergence and convergence is *necessary* to reach high performance in a novel and uncertain environment, we will exploit a new statistical tool, Necessary Condition Analysis (NCA, Dul 2016, 2020). Relating to the goal of educators and public policies in the 21st century to make children and adolescents more creative, and schools more receptive to creative thinking (ADOBE, 2013; Lucas et al., 2013), the current studies focus on the creative potential and behavior of middle and high school students.

2. Creative cognition as a trade-off between exploration and exploitation: an introduction

Burtini et al., (2015) provide a good survey of the technical literature on the trade-off between exploration and exploitation, and Cohen et al. (2007) discuss how the human brain manages this trade-off. Mekern et al. (2019) present a review of recent computational approaches "to demystify creative cognition". Hills et al. (2015) illustrate the generality of trade-offs between exploration and exploitation across diverse cognitive domains like animal foraging, visual search, problem solving, or social learning.

There is no known generally optimal solution to the bandit problem in stationary (stable) or changing environments, but a number of particular solutions (like the Gittins index under stationary conditions, infinite horizon and exponential discounting) and approximate algorithms have been proposed. However, the common feature of all the solutions is simple: it is rational to explore first and reap the rewards, i.e., exploit, when sufficient information has been gathered. In an uncertain and changing environment, some exploration is needed all the time, though often at a decreasing rate.

Navarro et al. (2015) designed an experiment in which subjects had to decide when to "observe", i.e., get information with no reward, and when to "bet", i.e., receive a reward with no immediate information on this reward. By separating information from reward, this task offers a way to compare the exploration-exploitation trade-off in environments that are known to be either stationary or changing. An important lesson from their experiment is that people's behavior is close to being optimal and most people adopt the optimal strategy after minimal experience.

How is the trade-off between exploration and exploitation of the information, depicted by the multi-armed bandit problem, implemented in the brain? There is clear evidence that divergent thinking and convergent thinking respond to different, and sometimes conflictual, cognitive processes. Radel et al. (2015) examined the role of impaired fronto-executive functioning on the enhancement of some types of creativity through a reduced capacity to exert inhibition. They tested this hypothesis by exhausting inhibition efficiency through prolonged and intensive practice of either the Simon or the Eriksen Flanker task. Subsequent creativity tests revealed that the impairment of inhibition enhanced fluency in a divergent thinking task (Alternate Uses Task, Guilford 1967), but no such changes occurred in a convergent task (Remote Associate Task, Mednick 1962). In another study, Chermahini and Hommel (2010) examined the effect of dopamine on divergent and convergent thinking by exploiting the fact that spontaneous eye blink rate (EBR) is a clinical marker of dopaminergic functioning. EBR predicted flexibility in divergent thinking and convergent thinking, but in different ways. The relationship with flexibility was independent of intelligence and followed an inverted U-shape function with medium EBR being associated with greatest flexibility. Convergent thinking was positively correlated with intelligence but negatively correlated with EBR, suggesting that higher dopamine levels impair convergent thinking. Beaty et al. (2016) contend that creative cognition involves dynamic interaction of two networks: a default network that influences the generation of candidate ideas, and a control network that constrains and directs this process to meet task-specific goals via top-down monitoring and executive control. These studies support the view that creativity emerges from the interaction of two mutually exclusive cognitive processes, implying a sequence of divergent and convergent behavior which requires some inhibition of the exploration process at one point to let the integration-exploitation process operate efficiently.

Our contribution in the present article will be to demonstrate that the multi-armed bandit captures an essential mechanism of creative cognition by means of behavioral experiments. In the next sections, we relate bi-dimensional measures of the creative potential of adolescents in the way they are usually captured in creativity research with their outcomes in a number of tasks and their use of an optimal exploration-exploitation trade-off in these tasks. It will be shown in the first experiment that high-school students with high creative potential are better decision-makers in a bandit problem task thanks to their adoption of an optimal sequence of exploration and convergence, but may be worse than their less creative fellows in a simple repetitive task. Because a major goal of educational policy in the 21st century should be to develop the creative potential of students, we also collected the grades obtained by the participants at the *Brevet*, a national exam taken in France by the vast majority of students after 9th grade, to examine the extent to which academic performance in French middle schools relates to divergent and convergent dimensions of adolescents' creative potential.

The second experiment controls exploration by imposing a radical change of consumption to middle-school students in the form of listening to unfamiliar music. Under these conditions, an excess of divergent thinking would be counterproductive but convergence would be especially needed to enhance immediate satisfaction with the novel product and probability of its future adoption.

3. Study 1: Creative adolescents are better decision-makers in a novel uncertain environment

3.1. Methodology

3.1.1 Participants

The sample was composed of Tenth graders (n=169) who attended a one-day educational conference about economics including participation in a laboratory experiment. Participants came from nine mixed classes of high school students from all academic districts of the Paris region (Paris, Créteil and Versailles, i.e., the center of Paris and suburbs), thereby enhancing its representative nature. Nine research sessions over three days (one per academic district) were conducted in the Paris Lab for Experimental Economics (LEEP), each session lasting 90 minutes. 76.3% of the students are aged 14 or 15, and 23.7% are aged 16 or 17; 46% are girls. This sample of teenagers controls for a number of common factors and is not contaminated by other unobservable factors that might appear later in life and cause additional differentiation between the subjects such as professional skills. A power analysis (g-power for group comparisons, using the Wilcoxon test) indicated that power was .71, given an estimated effect size of .2, two-tailed alpha of .05, and a sample size of 169. For regression analyses, power was greater than .90.

3.1.2 Material

The measures consisted of questionnaires for Big 5 personality traits and the willingness to take risks, two creativity tasks, two independent real-effort tasks including comprehension and satisfaction questions, and a general questionnaire including scores obtained on the *Brevet*- a nationwide exam taken after 9th grade. The incentives related to the two real-effort tasks, a number of candies or chocolate proportional to the gains, were well accepted. Some of the students won an entire bag of candies or chocolates, whereas others did not win anything.

Personality measures. Big 5 personality traits were measured with the French version of Ten Item Personality Inventory (TIPI), developed by Storme et al. (2016). The risk attitude was also captured by the question "How willing are you to take risks, in general?", rated on a scale from 0 to 10, that Dohmen et al (2011) have shown to be a reliable predictor of actual risky behavior, even after controlling for a large number of observables.

Creativity measures. Building on the wealth of literature, a modern approach to measuring creative potential was developed and formalized in the Evaluation of Potential Creativity battery (EPoC, Lubart et al., 2011). The measurement approach adopted in the EPoC battery is to solicit creative thinking in simulated situations (verbal creation, graphic creation). The individual is thus invited to engage in creative work and an estimate of a child's or adolescent's potential is obtained by comparing an individual's performance to others in the same task.

EPoC is composed of four tasks in each creative domain and can be used with children, adolescents or young adults. Two tasks engage divergent-exploratory thinking and two tasks involve convergent-integrative thinking. Both modes of creative thinking are evaluated in a given domain, notably the graphic/visual and verbal/literary domains. For example, in the graphic-visual field, we present a simple graphic stimulus (such as a banana shape) from which the participant must generate many drawings that integrate the stimulus, in a limited time. This task requires divergent-exploratory thinking. Also, in the graphic-visual domain, a task involving convergent-integrative processes involves presenting a series of photographs of objects from which the participant must produce a single integrated drawing, using at least four of the eight objects presented. These same kinds of graphic tasks using abstract stimuli are also part of the battery.

For convergent-integrative tasks, the scoring relies on judges, adults who are trained to recognize the convergent-integrative quality of a work (Storme et al., 2014). This part of EPoC focuses on creative synthesis and there are two main aspects that serve as the basis for scoring. First, the production must bring together, and truly synthesize, the various elements provided. The second criterion for convergent-integrative quality is the extent to which the synthesis is original, novel and meaningful. This novelty ranges from very typical, banal ideas to highly unique ones. It is appreciated by judges who have knowledge of the typical productions that children and adolescents provide for the given EPoC task, with the specific stimuli proposed in the task.

The profile of task performance (and thus the potential) of an individual will reveal his/her strengths and weaknesses in each creative domain. Inspired by the multivariate model of creativity (Sternberg & Lubart, 1995), EPoC provides a holistic approach to measuring creative potential. Thus, the tasks in EPoC refer explicitly to the divergent-exploratory or the convergent-integrative process modes. Of course, both divergent-exploratory and convergent-integrative modes are needed for creative potential.

The two graphical subtests of creativity from the EPoC battery (Lubart et al., 2011) were used in the first study for two main reasons. First, drawing is less culturally-dependent than writing and not a high-school-relevant skill; thus, it may be considered as fair and exogenous. Second, drawing is unexpected and generally engaging for middle-school and high-school students.

3.1.3 Performance tasks.

Two types of incentivized real-effort tasks were designed to observe the impact of adolescents' creativity on performance: typing tasks and buttons tasks. The "typing tasks" are purely repetitive whereas the "buttons tasks" aim at assessing the agents' search behavior. They are described in the following sections.

Buttons task: a multi-armed bandit problem. The buttons tasks operationalize the multi-armed bandit problem. They consist of two series of a hundred clicks on four or eight buttons respectively[1]. Each click and each button yield a number of points that can be positive, null or negative; each button holds a hidden pre-set distribution of values, the same for all participants. To control for order effects, the ordering of buttons is counterbalanced across participants. The cumulative distribution of points across buttons is presented in Tables A.2 and A.3 in the supplementary materials for the two buttons tasks. The goal assigned to players in these tasks is to obtain the highest cumulated number of points. One of the buttons tasks is randomly selected at the end of the session and its outcome determines the reward for this part. Since incentive differences might have an influence on creative behavior, participants were randomly assigned to two incentive conditions: "performance" and "competition". Players get one candy

[1] A third buttons task was also considered but we will not refer to it in this paper.

for 30 points in the performance condition for all buttons tasks. In the competition condition, the two players with the highest scores in the session receive a pack of 39 candies each whereas other players get nothing. Expected rewards are approximately the same in the two incentive conditions. On average, participants in the first task received 13.33 candies (max: 17.83), and in the second task they received 11.87 candies (max: 24.73)[2].

The first "basic" series of the buttons tasks is called "equal solution" because the same score can be reached with any button, and with a random or systematic choice of buttons. No unique optimal strategy should be observed here as any buttons combination over the 100 clicks will lead to the same expected score. In contrast, the second series was designed to reveal a link with the two facets of creative behavior. It is called "best button" because, after a number of trials, one can eventually discover the best button and understand that it will permanently remain so. To make this task more complex and search more valuable, we introduced eight buttons instead of four. The 96th click on the best button yields a "jackpot" of 200 points. The latter was unpredictable, but a participant who recognized the best button and its permanence would be more likely to win this prize. It is important to note that the best button was still the best if the player missed the jackpot. This series has been designed to elicit behavior relevant to creativity: there are lots of paths (buttons) to explore, and one is definitely the best but it takes time to discover and integrate this truth, and eventually win the jackpot.

Typing task: a simple repetitive task. In the typing tasks, participants retype as many codes as they can in three minutes' time. The list of codes is displayed on each participant's screen. A code consists of five random letters that do not form a word. Once correctly retyped, the code is highlighted in the list. The typing task was performed under a 2x2 factorial within-subject design: (alone, pair) x (performance, competition). In the "pair" conditions, participants were

[2] The number of candies was rounded up at the moment of reward distribution.

matched with a randomly assigned partner from the same session whose picture appeared on the screen but with whom they could not communicate[3]. Each participant played the four conditions which appeared in variable order across sessions. The "performance" condition is a "piece-rate scheme" (i.e., each participant gets a reward which is proportional to the number of codes correctly retyped) whereas the "competition" condition is a tournament (i.e., everybody participates but only the best performers are rewarded). Charness & Grieco (2019) suggested that incentives matter in tasks with pre-defined goals and constraints like ours. The detailed payoff structure per treatment is provided in Table A1 in the Appendix. The average number of correct codes in the treatment Alone*Performance was 21, which corresponds to a payoff of 7 chocolates. Expected rewards are approximately the same in the four incentive conditions.

The two selected scores of creative potential are graphical divergent thinking (GDT) and graphical integrative thinking (GIT). Correlations between them (r=.13) and with personality traits are not significant. The fact that GIT and GDT are not very correlated justifies our choice to consider them separately in statistical treatments.

3.2 Results (1): Buttons task

3.2.1 Optimal strategies.

The purpose of the two buttons tasks was to highlight the interplay of the two components of creativity by letting individuals make good dynamic decisions from experience. The optimal strategies are radically different in the "equal solution" task and in the "best button" task. No best behavior emerges in the "equal solution" task, because the expected performance is the same whether one switches constantly or does not switch at all. In sharp contrast, there exists an optimal strategy in the "best button" task, which is to search for the best button in the beginning of the task and to stick to the best button once it has been found. This simple strategy,

[3] Participants to a session were students of the same class who knew each other. By showing the picture, we reproduce conditions prevailing in a team of workers.

called “ϵ-first” in multi-armed bandit literature, is optimal in a stationary environment with a fixed and known horizon H (100 rounds in our experiment). During the first $\epsilon.H$ plays, which form the exploration phase, the individual picks arms uniformly randomly, producing an estimate of each arm’s payoff. In the remaining $(1-\epsilon).H$ plays, which form the exploitation phase, she makes systematically the choice of the best empirically estimated arm. Table 2 confirms the ϵ-first optimal strategy of *early exploration-subsequent convergence* for the best button task, with $\epsilon = 0.2$. Such behavior conforms to individuals with a high creative potential.

A group of high (low) “early explorers” is composed of subjects who switched strictly more (no more) than average (10.13 switches) across buttons in the first 20 clicks, and a group of high (low) “late convergers” is composed of subjects who chose intensively strictly more (no more) than average (7.6 clicks) the best button in the last 20 clicks. Table 2 compares the average number of points obtained in the best button task as a function of the respective timing of early exploration and late convergence toward the best button. The high “early explorers/late convergers” obtained 3.2 times more points in the best button task than the low “early explorers/late convergers”. However, low explorers did not perform significantly less than high explorers if they were also high convergers because lucky individuals who found the best button quickly and kept clicking on it subsequently performed obviously very well. Comparatively, the relative performance of the two extreme groups in Table 2 was only 0.95 in the equal solution task, showing that creativity may be harmful for simple tasks[4]. Moreover, it is shown in Table 3 that the number of switches - a measure of exploration - is positively and significantly correlated with the number of late clicks on the best button - a measure of convergence - when switches are decided early; but the correlation turns negative and significant when switches are decided late.

[4]The high early explorers/late convergers obtained 375.91 points on average versus 397.44 points for the low “early explorers/late convergers”. This small difference is significant (p=0.000), however, because the scores of the former are more dispersed (Sd=35.60) than those of the latter (Sd=19.69).

Table 2 Average cumulated outcomes in the best button task as a function of the respective timing of exploration and convergence toward the best button

Number of switches (first 20 clicks)

Number of clicks on the best button (last 20 clicks)

Explorer / Converger	Low	High	p-value
Low	184.56 (15.46) N=50	249.17 (17.00) N=53	0.003
High	574.81 (159.22) N=32	591.17 (118.46) N=34	0.903
p-value	<0.001	<0.001	

Notes: Standard deviation in parentheses. Statistical test: Wilcoxon sign rank test. High explorers switched 11 times at least in the first period of 20 clicks. High convergers chose the best button 8 times at least in the last period of 20 clicks.

Table 3 Correlation between numbers of switches over 20-click periods and of late clicks on the best button

	number of switches	r	p
clicks	1-20	0.1858	0.0000
	21-40	-0.0133	0.7118
	41-60	-0.2048	0.0000
	61-80	-0.2536	0.0000
	81-100	-0.3505	0.0000

Note: r: Pearson's correlation coefficient.

Do adolescents adopt an optimal strategy? Figure 1 shows that they do on average, at least to some extent. In line with expectations, the frequency of switches is maintained around 40% during the equal solution series, over the 100 clicks. Participants search for a solution until the end of the task not knowing with certainty that there is in fact no single best solution. On the other hand, in the best button task the frequency of switches falls monotonously from 50% to 33% from the beginning to the end. This occurs because some participants eventually identify the best button, and then keep clicking on it frequently. The frequency of clicks on the best button is multiplied by three on average between the beginning and the end of the task.

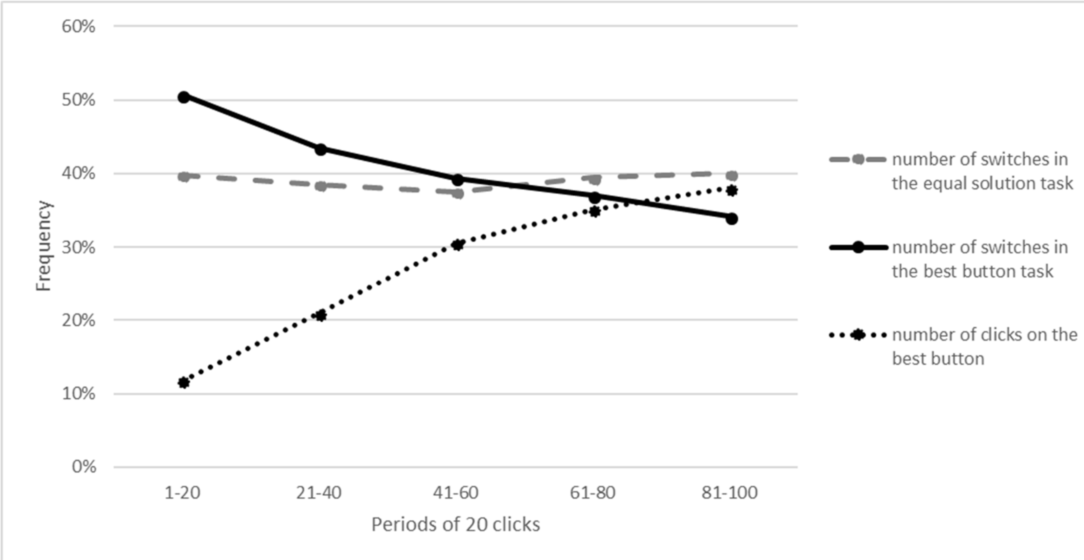
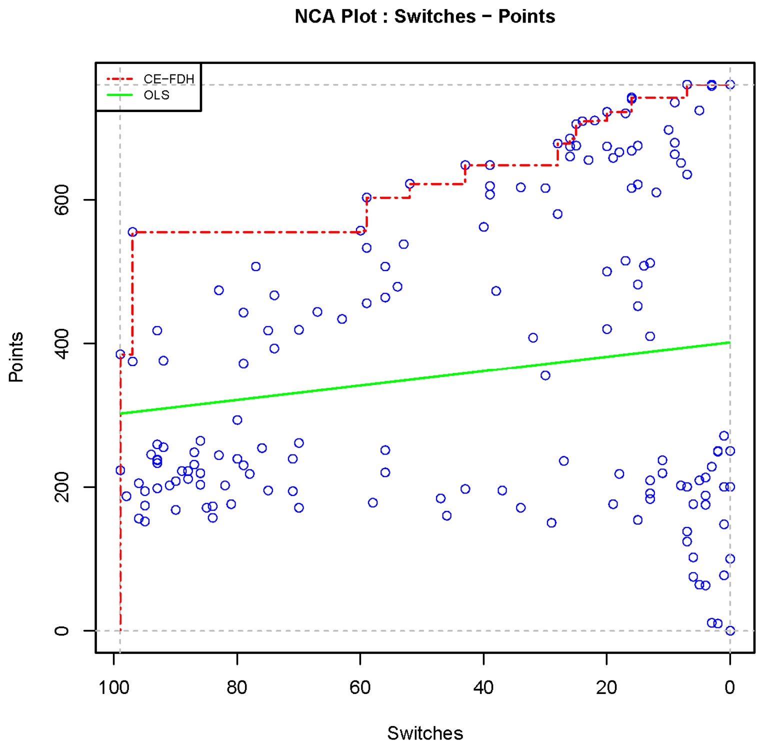


Figure 1 Exploration and convergence in the two buttons tasks

3.2.2 *Is creativity necessary for high performance in the best button task?*

Having identified the optimal strategy *on average*, one can wonder whether this strategy of early exploration and subsequent convergence is *necessary*, hence required for each individual to obtain high performance. To answer this question, we performed a Necessary Condition Analysis (NCA, Dul, 2016). NCA is a novel data analysis method based on necessity logic that has been applied in several disciplines, including the psychology of academic success (Tynan et al., 2020) and the psychology of creativity (Dul et al., 2020). For example, scholars demonstrated with NCA that intelligence is a necessary condition for creativity (Karwowski et al., 2016, 2017; Shi, et al., 2017).

We applied NCA with R package NCA 3.2.1 to evaluate whether all high performing individuals use the optimal strategy, or that some individuals could achieve high performance without this strategy. Hence, we answer the question: "Is early exploration and subsequent convergence *necessary* for high performance"? The results of our necessary condition analysis are shown in Figure 2.



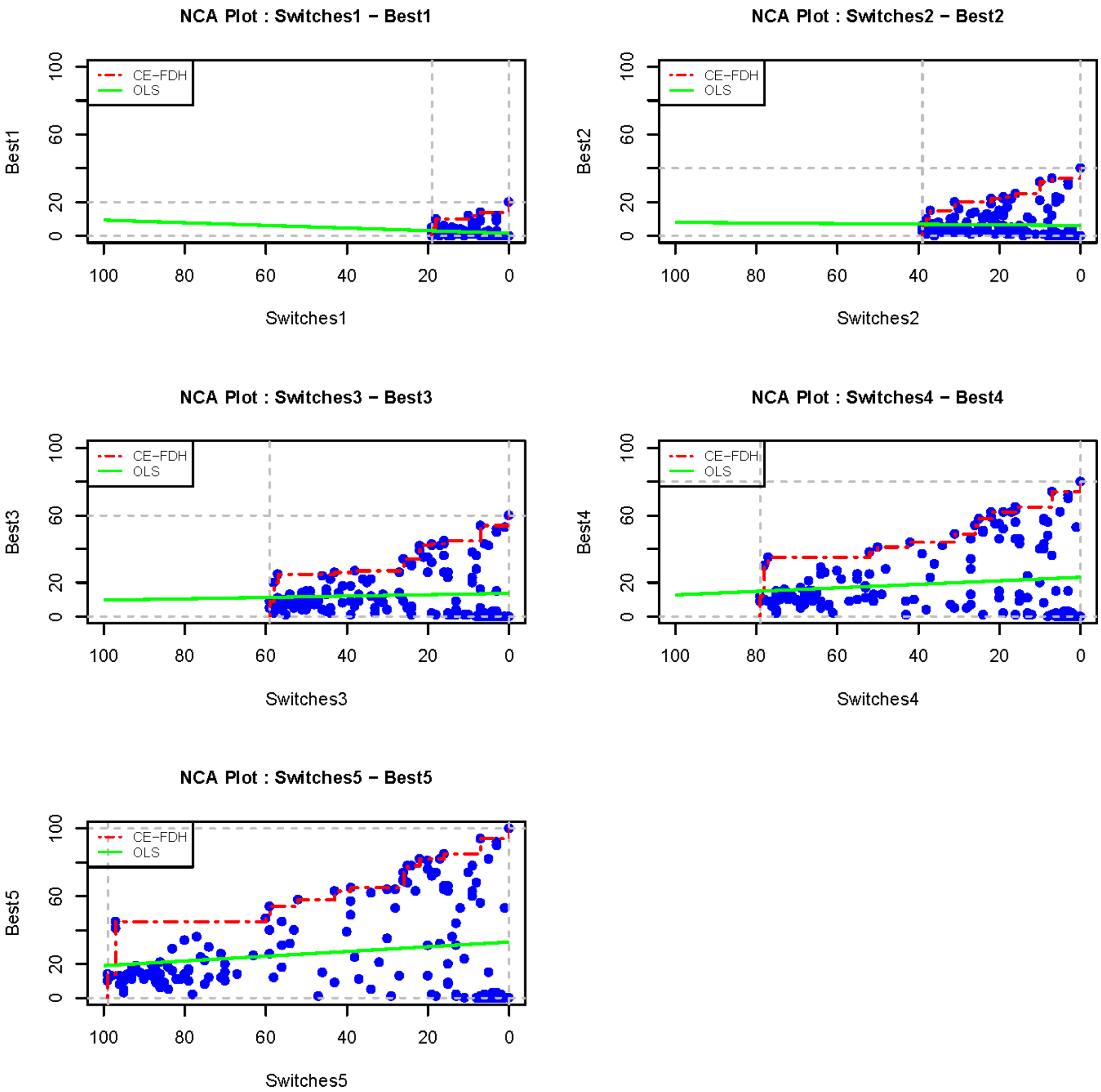
Note: The line through the middle is the OLS regression line for comparison, which shows the *average* outcome in the best button task as a function of total number of switches.

Figure 2. Necessary Condition Analysis (NCA) showing the *maximum* outcome in the best button task as a function of number of switches.

In this figure the horizontal axis represents the number of switches during a session of 100 clicks (Note that the horizontal axis is reversed and runs from left to right: from 100 to 0). The vertical axis represents performance, i.e., the number of points in the best button task after 100 clicks. NCA draws a *ceiling line* on top of the data of an XY scatter plot[5]. This line is different from a central line such as the OLS regression line (also shown in Figure 2 for reference). The ceiling line represents the level of $X = X_c$ that is necessary (but not sufficient) for a given level of performance $Y = Y_c$. The necessity effect size is the empty space above the

[5] For all NCA analyses we selected the CE-FDH ceiling line because of the non-linearity of the border between the empty and full area. The p-values were estimated with 10000 permutations.

ceiling line as a fraction of the total possible space of observations, given by the minimum and maximum observed values of X and Y. The results support that a low total number of switches is necessary for high performance. The effect size is 0.18 ($p < 0.001$). This means that with a high number of switches it is impossible to have a high-performance score. For example, when an individual switches more than 60 times, the performance score cannot be above approximately 550. For reaching a performance score of more than 700 the number of switches should not be higher than about 30. The highest scores are reached with a few switches only. The results also show that a few individuals reached the highest scores without switching at all. Apparently, these individuals were lucky that they pushed on the best button at the first time without any further exploration. Thus, we can state that *early exploration or luck is necessary for high performance in conjunction with the necessity to stop exploring after the best option is discovered*. The ability to stop on time while exploring is a crucial part of creative performance. The necessity of convergence is illustrated also in Figure 3, which shows the cumulative number of clicks on the best button for the five successive periods of 20 clicks. Individuals who continue exploring will not have a high number of clicks on the best button, and consequently will not gain a high number of points.



Note: The ceiling line represents the *maximum* number of clicks on the best button as a function of the number of switches. The line through the middle is the OLS regression line for comparison, which shows the *average* number of clicks on the best button as a function of number of switches.

Figure 3. Necessary Condition Analysis (NCA) showing how the ceiling line (CE-FDH) develops over time in 5 successive periods of 20 clicks.

3.2.3 Creativity and task performance

Having identified the optimal strategy in the best button task and its inexistence in the equal solution task, it will now be demonstrated that exploration behavior is enhanced by divergent thinking (GDT) and subsequent convergence in the best button task is enhanced by integrative thinking (GIT). This is shown in Tables 4a, 4b, and 4c. The results confirm the theoretical predictions. Table 4a shows that high-performing adolescents endowed with a high GDT score switch significantly more than others in *all* 20-click periods for the "equal solution" task. They maintain their exploration intensity all along the task in search of the best move. For the "best button" task, Table 4b shows that the significantly positive impact of the GDT score on switching is limited to the first two periods of 20-clicks, which conforms to the optimal strategy. Finally, Table 4c demonstrates that GDT has no effect on the frequency of clicks on the best button whereas GIT has a sizable effect on the latter essentially in the last period of 20 clicks, as predicted by the optimal strategy. The results contained in these three tables are a clear indication that creativity is useful and even necessary for making good decisions in a complex dynamic environment. Divergent thinking is useful and necessary for uncovering potential moves[6] and integrative thinking is useful and necessary for discovering the best move and reaping the benefits from it in the future. Further proof of this last claim is given in Table 5, showing that winning the 200-point jackpot by clicking the best button on the 96th click is not just a matter of luck. Integrative thinking significantly raises (at the 5% level) the probability of winning the jackpot and is the only significant variable in the Probit analysis.

[6] Lee et al. (2017) theorizes the effect of switching on creativity by arguing that switching reduces cognitive fixation.

Tables 4 a,b,c Exploration-exploitation trade-off and creativity (OLS[7])

		(1)	(2)	(3)	(4)	(5)
Table 4 (equal solution)	**Number of switches**	1-20	21-40	41-60	61-80	81-100
	GDT	1.226**	1.964***	2.100***	1.790***	1.347**
		(0.507)	(0.657)	(0.526)	(0.545)	(0.619)
	GIT	-0.176	-0.482	-0.202	-0.164	0.122
		(0.563)	(0.612)	(0.628)	(0.662)	(0.675)
	Controls	Yes	Yes	Yes	Yes	Yes
	Constant	14.45	14.78	21.67	19.50	27.18
		(13.42)	(18.05)	(18.30)	(18.90)	(18.96)
	Observations	169	169	169	169	169
	R-squared	0.159	0.147	0.187	0.126	0.116
Table 4b (best button)	**Number of switches**	1-20	21-40	41-60	61-80	81-100
	GDT	1.371***	1.282**	0.991	1.113	1.069
		(0.523)	(0.621)	(0.664)	(0.695)	(0.681)
	GIT	-0.0458	-0.00702	-0.124	-0.0579	-0.184
		(0.601)	(0.638)	(0.661)	(0.669)	(0.635)
	Controls	Yes	Yes	Yes	Yes	Yes
	Constant	25.77	5.804	32.88*	32.28*	28.42
		(16.16)	(18.30)	(18.09)	(18.76)	(18.69)
	Observations	169	169	169	169	169
	R-squared	0.159	0.153	0.139	0.124	0.100
Table 4c (best button)	**Number of clicks on the best button**	1-20	21-40	41-60	61-80	81-100
	GDT	0.0533	0.691	1.080	0.550	0.852
		(0.256)	(0.592)	(0.676)	(0.712)	(0.690)
	GIT	0.242	0.481	0.0739	0.819	1.317**
		(0.330)	(0.502)	(0.587)	(0.650)	(0.652)
	Controls	Yes	Yes	Yes	Yes	Yes
	Constant	-13.12	-20.46*	-37.02**	-26.76	-22.74
		(8.609)	(12.15)	(15.35)	(17.04)	(18.60)
	Observations	169	169	169	169	169
	R-squared	0.125	0.187	0.184	0.169	0.157

Notes: Standard errors in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.10. Control variables: academic district, session, order of the sub-task, comprehension errors, Big Five personality traits, girl, age, general risk seeking, competition treatment (dummy).

[7] Results are very similar when performing a negative binomial regression model.

Table 5 Probit on finding the jackpot

	Prob (jackpot)
GDT	0.0804
	(0.107)
GIT	0.233**
	(0.114)
Controls	Yes
Constant	-5.170
	(3.186)
Observations	169
Pseudo R-squared	0.0830

Notes: Standard errors in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.10. Control variables: academic district, session, order of the sub-task, comprehension errors, Big five personality traits, girl, age, general risk seeking, and competition treatment (dummy variable).

3.3 Results (2): Typing task

The purpose of the typing task was to test our hypothesis that creative individuals are no more productive than others in simple well-known tasks bearing no uncertainty. The adolescents' productivity in our typing tasks is measured by the number of correct codes typed in three minutes. An OLS regression in Table 1, column (1), relates productivity in typing to the two creativity scores, the four conditions, gender, general risk-seeking, and a number of controls listed below the table. It is worth noticing that free riding occurred when the teenagers worked in pairs, but competition forced them to cooperate and work harder when they formed pairs. However, if subjects work alone, a tournament scheme, *i.e.*, the competition condition, does not make the adolescents work harder than a piece-rate offering the same expected return[8]. Girls are no more productive than boys but more willing to cooperate with a partner; and general risk-seekers are both more productive and more cooperative than others.

[8] The impact of economic factors like incentives, competition, cooperation, and strategic behavior on creativity is discussed in the recent survey of Attanasi et al. (2021).

Our main result is that high creativity scores do not help typing fast as neither GDT nor GIT show significant relationships to task performance.

The probability of typing faster than the partner is captured by a Probit in column (2) of Table 1. Interestingly, this regression shows that the more integrative thinkers were more eager to cooperate with their partner when they formed a pair, whether pairs competed or not, as the probability of working harder than the partner increases significantly (at 5% level) with GIT. Increasing the GIT score by one standard deviation raises the probability of being the more productive partner by 10.1%[9] (p = 0.014). A plausible reason for more integrative thinkers being cooperative is that they are intelligent (Karwowski et al. 2016), and understand the benefits from cooperation in teamwork.

Table 1 Absolute and relative output in the typing task

	(1) OLS Number of correct codes	(2) Probit Prob (typing faster than partner)
GDT	-0.299	0.0512
	(0.683)	(0.112)
GIT	1.216	0.289**
	(0.781)	(0.123)
Competition	-0.589	0.0548
	(0.543)	(0.223)
Pair	-1.308***	
	(0.499)	
Competition x Pair	3.140***	
	(0.651)	
Female	0.477	0.685***
	(1.336)	(0.239)
General risk seeking	0.656**	0.114**
	(0.269)	(0.0481)
Constant term	31.76	-2.778
	(22.06)	(3.854)
Controls	Yes	Yes
Observations	436	218
R-squared/Pseudo R-squared	0.202	0.113

Notes: Standard errors in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.10. Competition is a dummy variable=1 for the competition treatment, 0 otherwise. Pair=1 if pair treatment, 0 otherwise. General risk seeking: a 10-point Likert Scale. The control variables included

[9] Average marginal effects.

are the following: school district, session, order of the tasks, number of comprehension errors, scores of the Big Five personality traits, age.

3.4 Results (3): School grades and creativity

The results indicated that divergent-exploratory thinking and convergent-integrative thinking, the two components of creativity, were related essentially to better sequential decisions from experience in a bandit problem, which is the archetype of a novel, uncertain, and dynamic environment. Given that the enhancement of creative behavior is becoming an educational goal, we took the opportunity to examine relationships between scores on creative potential and school grades on the national middle school exam at the end of middle school (9th grade), the *Brevet,* as reported by the 10th graders. It is important to note that this is an anonymous national exam in which grades are not delivered by the students' own teachers.

The French grading system is based on a continuous 20-point scale. We regress grades in Math, French literature, and the general grade[10] for the *Brevet* on individual measures of creativity, personality, general risk tolerance, gender, controlling for age and academic districts. As the error terms can be correlated across all the different grades contributing to the general grade as the result of common unobservable factors, we estimated the four grades with the SUR (Seemingly Unrelated Regression) method. Indeed, the error terms were correlated (Breusch-Pagan test of independence: chi2= 62.415, $p < 0.0001$).

As the two components of creativity may be complements or substitutes in scholastic performance, we introduced them with an interaction term. We disaggregated three of the five personality variables (openness, agreeableness and emotional stability) into their positive and negative components because these are not strictly opposite within the condensed Ten Item Personality Inventory (TIPI) score and do not necessarily exert opposite effects. We present the regressions in Table 6.

[10] History/Geography and Art history also contribute to the general grade with Math and French literature.

Table 6 SUR regression of school grades of 9th graders

Grades (SUR)	(1) Math	(2) French literature	(3) General grade
GDT	0.196	0.245	0.205
	(0.274)	(0.217)	(0.175)
GIT	0.720**	0.528**	0.240
	(0.295)	(0.234)	(0.189)
GDT x GIT	0.0177	-0.541**	-0.483***
	(0.292)	(0.232)	(0.182)
Risk-seeking	-0.116	0.321***	-0.0111
	(0.155)	(0.123)	(0.0961)
Girl	-1.199**	1.671***	0.0830
	(0.576)	(0.457)	(0.359)
Open to new experiences, complex	-0.0902	0.186	0.270**
	(0.186)	(0.148)	(0.115)
Conventional, uncreative	-0.0638	0.327**	0.235**
	(0.178)	(0.141)	(0.116)
Conscientious	0.292	0.127	0.440***
	(0.269)	(0.213)	(0.166)
Extravert	0.0392	-0.0454	0.0349
	(0.278)	(0.221)	(0.164)
Critical, quarrelsome	0.490**	-0.107	0.238*
	(0.221)	(0.175)	(0.137)
Sympathetic, warm	-0.0286	-0.415**	-0.281*
	(0.230)	(0.183)	(0.143)
Anxious, easily upset	-0.369**	-0.111	-0.134
	(0.161)	(0.128)	(0.0970)
Calm, emotionally stable	-0.00352	0.197	0.0707
	(0.201)	(0.159)	(0.115)
Controls	Yes	Yes	Yes
Constant	12.23	3.283	8.371
	(8.271)	(6.567)	(5.328)
Observations	113	113	119
R-squared	0.262	0.330	0.263

Notes: Standard errors in parentheses. Significance level: *p < 0:10, **p < 0:05, ***p < 0:01. Control variables: age and academic district. Gender=1 for female Personality traits: Openness to Experiences: Open to new experiences, complex + Conventional, uncreative (reverse-scored item), Agreeableness: Critical, quarrelsome (reverse-scored item) + Sympathetic, warm; Emotional Stability: Anxious, easily upset (reverse-scored item) + Calm, emotionally stable.

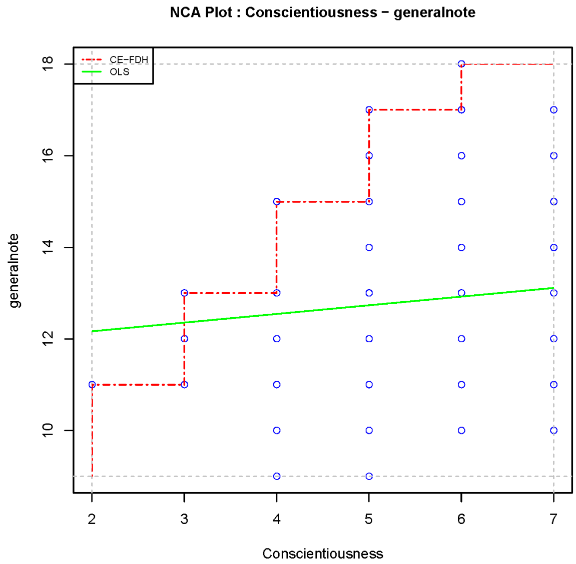
The relationship of creativity to the three grades reported here is contrasted. GIT has a positive, significant relationship for both Math and French literature[11]. This is not surprising because integrative thinking should help students synthesize knowledge and learn efficiently. More surprisingly, GDT has a significantly negative relationship to grades in French literature.

[11] Also note that anxiety is negatively related to mathematical ability and girls have lower grades than boys in Math and higher grades in French literature even after controlling for the set of creativity and personality measures.

Convergent and divergent thinking show distinct relationships to Math and French literature grades. However, they appear to be substitutes for the general grade (negative interaction term). The opposite effects of GDT and GIT on grades suggest that there is little room for divergent thinking in French middle school; a conclusion confirmed by the result that individuals with relatively low scores on divergence are better graded in French literature than those with high scores.

The finding of a negative attitude of French middle schools toward divergent students may reflect the general expectation of teachers that young adolescents should learn existing knowledge before solving novel issues. Therefore, divergent thinking of students would appear excessive to anonymous teachers at this stage of the child's education while convergence would be valued.

We performed additionally a necessary condition analysis on these data that shows no necessity of creativity for high grades. In contrast, conscientiousness is needed by a student to be successful in middle school (Figure 4). Interestingly, conscientiousness has no significant effect on each specific grade, but the attention uniformly devoted to all fields makes a difference on the general grade. Conscientiousness not only contributes to a high general grade *on average* as the results of the previous analysis shows, it is also *necessary* for students to have a high general grade. The necessity effect size is 0.30 ($p = 0.004$).



Note: The line through the middle is the OLS regression line for comparison, which shows the *average* grade as a function of conscientiousness.

Figure 4. Necessary Condition Analysis (NCA) showing the ceiling line (CE-FDH) for general grade as a function of conscientiousness.

4. Study 2: Creative adolescents adopt innovations earlier

The continuous adoption of innovations in a changing world requires decision making in a novel and uncertain environment. Thus, the ability to solve bandit problems provided by creative cognition should lead the more creative adolescents to adopt innovations faster than others. This hypothesis will be tested experimentally on the ease of consuming new products. Music was selected as an experimental good because consuming music in the lab is no different from consuming music outside; and music can be listened to repeatedly in short spells without causing excessive satiety, repulsion, and boredom.

We offered the opportunity to middle-school students to listen repeatedly to unfamiliar music (Classical and Jazz/Blues) and report their satisfaction after each listening episode. As is well-known in marketing analysis, the reported satisfaction from exposure to a product is a good predictor of further consumption of this product in the future (also see Lévy-Garboua et al., 2006). Therefore, a high level of reported satisfaction for unfamiliar music signals a relatively high propensity for adolescents to adopt this style in the future. The more creative adolescents are expected to appreciate unfamiliar music more than average and to be the potential adopters of the novel music.

Furthermore, by forcing students to listen to unfamiliar music, we prevented exploration of other musical styles and thus restricted the need for divergent thinking. Under these conditions, a high potential for divergence might be counterproductive and would not enhance average satisfaction of listening to unfamiliar music unless it is controlled by a high potential for convergence.

4.1 Methodology

4.1.1 Participants

Study 2 was conducted in a middle school from a low socioeconomic Parisian suburban area (Rosny-sous-Bois). Participants were 9th graders (n=99, mean age=14.4 years, Sd=0.62). Given the focus on adoption of novelty, the choice of this sample enhanced the tendency to listen to popular music (Rap-RnB, Pop-Rock), allowing classical music to be a novelty genre. A power analysis (g-power for group comparisons, using the Wilcoxon test) indicated that power was .99, given a medium effect size, two-tailed alpha of .05, and, for regression analyses, with a medium effect size, power was greater than .60.

4.1.2 Materials

Brief Big Five (BB5, Barbot 2012). This is a short version of the Big Five (McCrae & Costa, 1987) targeted for teenagers. This questionnaire uses a list of 100 adjectives to get a score on the Big Five (Openness, Conscientiousness, Extroversion, Agreeableness, Emotional stability). For each adjective, individuals report how much it corresponds to them on a 5-point Likert scale (from "totally" to "not at all") and these scores are then aggregated along the same five dimensions of the Big Five.

Evaluation of Creative Potential (EPoC). The graphic tasks were described in study 1. In the verbal domain, the divergent thinking tasks require the individual to produce many endings to a story, in which the beginning is provided. Alternatively, in a second task, many beginnings of the story must be generated for a certain story ending. In the verbal convergent-integrative task, a title, or a set of characters is provided. The individual needs to generate a single, elaborated story that integrates the title or characters. These tasks are presented on two occasions separated by one week so that the stability of their creative potential can be assessed.

4.1.3 Procedure

There were three sessions, each one organized as a regular class in presence of the teacher. In the first and second sessions, we implemented the EPoC measure (Lubart et al. 2011) in order to elicit divergent and convergent *graphical* thinking (GDT and GIT), and *verbal* thinking (VDT and VIT) respectively. The in-class measurement of graphic and verbal creative potential required two sessions of 50 minutes using paper-and-pencil assessments. We recorded those scores for 81 pupils.

In the third session, we used notebooks and tested student's propensity to like a musical style they were not familiar with over repeated exposures to that style. We used four musical genres: Rap/RnB, Pop/Rock, Jazz/Blues and Classical music. Each pupil reported the frequency of listening to each genre on a four-point scale. The two genres they were listening most were then considered as their « usual » music, the two others as « novel » music. 95% were used to listen to Pop/Rock and Rap/RnB music. Only 5 pupils were used to listening to the other genres. They were excluded from the analysis. We let participants listen to their familiar music during four periods of one minute, and then they listened exclusively to Jazz/Blues and Classical music during 20 periods of one minute and rated their experience using three measures: satisfaction (10-point scale), pleasure (5-point scale) and arousal (5-point scale). Self-assessment manikins (Bradley & Lang, 1994) were used for the pleasure and arousal measures.

During the third session, the teenagers were asked to fill in a demographic questionnaire (gender, age), an adapted Holt & Laury (2002) test of risk aversion, and a Brief Big Five (BB5) personality test for teenagers.

Word-of-mouth treatments were introduced in the experiment to control for the malleability of adolescents' preferences. In recent generations, teenagers tend to be in opposition with the preferences of parents and adults. Thus, middle school students might take high school students as their model and adopt their musical prescriptions. Two "word-of-mouth" treatments were introduced to test this hypothesis, called Low (n=35) and High (n=34) respectively, in addition to a control group (n=30). Students in the word-of-mouth groups were shown the mean satisfaction reported by the low-satisfaction and high-satisfaction quartiles drawn from a sample of high school students (10th graders) on the same musical pieces. The control group received no such information.

As the experiment took place over three different weeks, some pupils did not attend the three classes. The sample of students with complete and consistent data contains 59 observations out of 99 who listened to music. Subjects from the restricted sample have higher grades on average than those who skipped one or two classes. The greater homogeneity of our sample might reduce the significance of the measured effects of creativity or personality on the satisfaction of listening to music.

4.2 Results

Our students gave low ratings to the novel music to which they had to listen. On average, they reported a satisfaction score of 5.02 (out of 10) for their familiar music versus 2.98 for the novel music (p=0.00 for a Wilcoxon sign rank test). Girls, however, were less negative than boys. Indeed, the average satisfaction level for familiar and novel music was respectively 5.48 and 3.30 for girls, and 4.65 and 2.72 for boys (Wilcoxon sign rank test on the difference of average satisfaction by gender yielded a p-value=0.00 for both the familiar and novel music).

Surprisingly, middle-school pupils strongly contradicted the judgment of their high-school "models": those exposed to the lowest satisfaction scores gave above-average ratings, whereas those exposed to the highest satisfaction scores gave below-average ratings. By comparing the average over the 20 periods, pupils in the "Treatment High" (TH) reported significantly lower satisfaction than in the control group (T0) (Wilcoxon sign rank test, p=0.00). Similar trends and differences between treatments were found for the pleasure-component of overall satisfaction (Wilcoxon sign rank test, p=0.00). For arousal, differences between treatments were not significant. With only one grade difference, the younger class identified the older class as an out-group from which they ought to differentiate, suggesting that middle-school pupils build their musical identity in opposition, not just with parents or adults, but also with high-school pupils.

4.2.2 Creative types and music appreciation

If creative cognition is a bandit problem, we expect creative adolescents to appreciate novel music, but not familiar music, more than others. Moreover, in the context of our experiment that prevents deliberate choice of music, we demonstrate more specifically that a high potential for divergence will be counterproductive if it is not counterbalanced by a high potential for convergence.

To test these predictions, we distinguish two creative dimensions (Verbal, Graphical) and two creativity types (Divergent/Explorer, Convergent/Integrative) that we designate as Verbal-Divergent (VDT), Verbal-Integrative (VIT), Graphical-Divergent (GDT), and Graphical-Integrative (GIT). Then, we divide the working sample (N=81) into high and low-ranked individuals relative to each dimension and type-specific mean creativity score[12]. Table 7a shows

[12] The four sub-groups in each creative dimension are rather equally distributed, except for relatively few high explorers in the verbal dimension (26 high vs. 55 low verbal explorers) reflecting the low socioeconomic background in the selected school.

the average satisfaction reported in each of the sub-groups for familiar music and Table 7b does the same for novel music.

Looking first at Table 7a, in line with our hypothesis, creativity does not significantly affect average satisfactions from listening to familiar music, with one exception concerning verbal creativity: high verbal divergence is bad with low convergence but it is good with high convergence. (p=0.092, t-test). This last result corroborates our central assumption that high divergence must be inhibited at one point to converge toward creative behavior.

Table 7a: Creative types and average satisfaction derived from familiar music

Verbal

Explorer \ Converger	Low	High	p-value
Low	4.84 (2.97) N=100	5.22 (2.72) N=120	ns
High	4.63 (2.83) N=40	5.58 (2.74) N=64	0.092
p-value	ns	ns	

Graphical

Explorer \ Converger	Low	High	p-value
Low	4.95 (3.11) N=112	5.01 (2.43) N=68	ns
High	5.38 (3.02) N=56	5.18 (2.60) N=88	ns
p-value	ns	ns	

Notes: Standard deviation in parentheses. Statistical test: T-test. High verbal (graphical) convergers' scores are strictly above mean VIT (GIT). High verbal (graphical) explorers' scores are strictly above mean VDT (GDT). Number of observations is based on the number of pieces of music*number of individuals per subgroup. The total number of observations is 4*81=324.

Table 7b: Creative types and average satisfaction derived from novel music

Verbal

Explorer \ Converger	Low	High	p-value
Low	2.74 (2.18) N=500	3.40 (2.30) N=600	<0.001
High	1.96 (1.73) N=200	3.16 (2.47) N=320	<0.001
p-value	<0.001	ns	

Graphical

Explorer \ Converger	Low	High	p-value
Low	3,3 (2.46) N=560	2,68 (2.20) N=340	<0.001
High	2,55 (2.06) N=280	3,04 (2.17) N=440	0.003
p-value	<0.001	0.020	

Notes: Standard deviation in parentheses. Statistical test: T-test. High verbal (graphical) convergers' scores are strictly above mean VIT (GIT). High verbal (graphical) explorers' scores are strictly above mean VDT (GDT). Number of observations is based on the number of pieces of music*number of individuals per subgroup. The total number of observations is 20*81=1620.

Table 7b concerns novel music and offers a different picture. Since exploration opportunities have been ruled out in this experiment, we hypothesized that a high potential for divergence would be counterproductive if it was not counterbalanced by a high potential for convergence. Indeed, looking down the columns of table 7b shows that high divergence leads to a low appreciation of novel music by low convergers but not by high convergers, which corroborates our central assumption. And looking at the second row of this table shows that highly divergent subjects benefit from high convergence to enhance their appreciation of unfamiliar music. These results are wholly consistent with those of the button tasks and with our theoretical interpretation of creative cognition: individuals appreciate a novel piece of music if they are able to perceive the diversity/richness of sounds as high explorers would do *and* integrate it subsequently into a coherent picture or story as high convergers would do.

4.2.3 Regression and necessary condition analysis

We complete this analysis in Table 8 with an OLS regression of the three scores of overall satisfaction, pleasure, and arousal after having listened to the novel music. The explanatory variables are the four scores of creative potential that we measured (GDT, GIT, VDT, VIT), the five personality traits (BB5)[13], two dummies for treatment effects (TH, TL), gender, risk

[13] Among the 99 pupils who filled in the BB5, three did not finish the questionnaire and five did not have valid scores. We thus have 91 observations for the BB5. We use raw scores of the BB5 traits given by a specific computation.

aversion dummies (elicited from a Holt & Laury (2002) task adapted to teenagers[14]), and impatience[15].

Table 8 OLS regression on satisfaction, pleasure and arousal over the 20 periods of listening to the unfamiliar music

VARIABLES	(1) Satisfaction	(2) Pleasure	(3) Arousal
TH	-0.455	-0.228	0.411
	(0.557)	(0.260)	(0.288)
TL	0.586	0.179	0.415
	(0.505)	(0.214)	(0.334)
Girl	-0.851*	-0.414*	-0.286
	(0.444)	(0.222)	(0.378)
Openness	0.020	0.016	0.013
	(0.026)	(0.011)	(0.016)
Conscientiousness	-0.029	-0.019*	-0.006
	(0.023)	(0.010)	(0.012)
Extroversion	-0.045**	-0.015	-0.025*
	(0.021)	(0.009)	(0.014)
Agreeableness	0.054***	0.020**	0.023*
	(0.020)	(0.009)	(0.013)
Emotional Stability	-0.019	-0.004	-0.003
	(0.022)	(0.010)	(0.013)
GDT	0.012	-0.021	-0.003
	(0.167)	(0.066)	(0.052)
VDT	0.118	0.042	0.065
	(0.132)	(0.051)	(0.098)
GIT	- 0.325***	-0.106**	-0.122**
	(0.110)	(0.052)	(0.055)
VIT	0.520***	0.237***	0.111
	(0.162)	(0.069)	(0.112)
Risk aversion	0.365	0.242	0.219

[14] In the original Holt & Laury (2002) design, the subjects are confronted to ten choices among two bets yielding positive outcomes: *A* is a safe bet with payoffs \$2 and \$1.60 and *B* is a risky bet with payoffs \$3.85 and \$0.10. Probabilities of the higher payoffs are equal for the two bets and vary by steps of 0.10 from 0.10 to 1.00. Normally, subjects should switch only once from *A* to *B* for an intermediate value of this probability and the latter determines their risk aversion in a simple way. The crossover probability is a discrete index of risk aversion. In our experiment, the payoffs were converted in Euros and multiplied by ten.
The instructions were adapted to our sample of teenagers in the following way: "We consider two scratch games with different outcomes. The chances of winning each outcome vary. There are 10 scratch games to be played. With the first scratch game you can either win 20€ or 16€. In a second scratch game you can win either 38.50€ or 1.00€. You will have then to choose between these two scratch games 10 times (game A or game B)". The instructions for lottery choices were read aloud and the pupils could also read them on the screen. Then, an example with different payoffs was given to make sure they understood the task and they were allowed to ask questions. The rate of inconsistent choices (with more than one switch point) was no different from what has been observed on adults. In our sample, based on the number of safe choices and by assuming expected utility, we characterize our sample as risk averse or risk seeker. We exclude pupils who made multiple switches and did not seem to understand the task (n=19). We are hence left with 80 observations, among which 65% are strictly risk averse. In our analysis, we use a dummy variable equal to 1 if the pupil is strictly risk averse and 0 otherwise.

[15] Pupils were asked a simple question to measure their impatience: What would you prefer? A. Receive 10euros today B. Receive 11euros tomorrow. 43.43% (n=99) who answered A are described as impatient.

	(0.442)	(0.181)	(0.178)
Impatience	-1.177***	-0.625***	-0.593**
	(0.399)	(0.185)	(0.233)
Constant term	4.546**	2.288**	1.595
	(2.182)	(0.975)	(1.266)
R-squared	0.459	0.440	0.265
Observations	1,180	1,180	1,180
Number of Individuals	59	59	59

Notes: Robust standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1. TH and TL are respectively the "treatment high" and "treatment low" dummies. Gender=1 for female. Risk aversion=1 if the pupil is risk averse, 0 if risk seeker or risk neutral. Impatience=1 if the pupil is impatient, 0 otherwise.

A comparison of the three columns in table 8 shows very similar effects on overall satisfaction and its two components, pleasure and arousal. Introducing the two creative dimensions and types separately in the regression[16], we find no effect of divergence and contrasted effects of convergence on average satisfaction. The statistical insignificance of divergence is consistent with our hypothesis about the limitations imposed on exploration in this experiment. The average effect of creative potential appears to be dimension-specific as more verbal convergence favors the adoption of novelty in music consumption whereas more graphical convergence has a negative effect. As demonstrated by the comparison of rows 1 in Table 7b for the verbal and graphical creative dimensions, this result is essentially due to subjects with a low potential of divergence who benefit from high convergence in the verbal dimension but suffer from it in the graphical dimension. This finding may be related to the fact that, in this sample, verbal divergence is more skewed to the right than graphical divergence and graphical convergence is more skewed to the left than verbal convergence (see note 12 in this paper, and fig.1 of Berlin et al. 2016). It does not contradict our central assumption that excessive divergence is harmful because that assumption does not concern low divergence.

[16] The interaction effects between the creative potential have been examined in Table 7b, and we have too few individuals to introduce them in the regression.

Since creative cognition is dimension-specific, we achieved the necessary condition analysis (NCA) by searching for a difference in the optimal mix of divergence and convergence for novel music appreciation between verbal and graphical competencies. NCA was implemented on a linear combination of divergence and convergence: $Creative\ potential = \alpha V(G)DT + (1-\alpha)V(G)IT$. This analysis was replicated on decimal values of $\alpha = (0, 0.1, 0.2, \ldots, 0.9, 1.0)$ and the value of α maximizing the effect size was elicited. We found optimal values of 0.3 for the verbal dimension with an effect size 0.127 (p<0.001), and 0 for the graphical dimension with an effect size of 0.162 (p=0.003) in the restricted sample of 59 students with complete and consistent data. However, the gap between the two creative dimensions disappeared on the larger sample of 99 students in which the prevalence of verbal divergence is very limited. Thus, a student certainly *needs* convergence in a task that strictly limits exploration opportunities. Nonetheless, NCA points to the danger of a blind adoption of a common all-purpose creativity score, in line with lessons from the literature on bandit problems.

A replication of the necessary condition analysis (NCA) on convergence alone for both creative dimensions yielded significant and very similar effect sizes of integrative thinking on the restricted sample of 59 students with complete and consistent data and on the larger sample of 99 students containing all available observations: 0.162 (p=0.004) for graphical creativity and 0.111 (p < 0.001) for verbal creativity. More precisely, subjects with a low score of integrative thinking *never* gave a high score (i.e., 7 to 10) for unfamiliar music, which means that they were always unable to overcome their prior distaste for novel music.

Going back to Table 8, it is important to see that the creative potential remains highly significant after addition of many controls like social influence, gender, personality, risk aversion, and impatience. Extroversion and impatience drive adolescents into conformism and partly offset the beneficial impact of creative potential upon unfamiliar music appreciation.

When controlling for all individual variables, treatment effects, shown by TH and TL estimates, keep the right sign but are no longer significant; and girls appear to be less satisfied and pleased than boys. This last result may be attributed to observable differences in verbal creative potential and personality among adolescent boys and girls, as demonstrated in table A4 of the appendix which reports substantial gaps in the BB5 scores between boys and girls at middle-school age.

General discussion and Conclusion

The first study served the purpose of demonstrating that adolescents with a higher-than-average creative potential made better decisions in a multi-armed bandit problem rather than in simple repetitive tasks like our typing or equal solution task. We found strong evidence for the optimality of early exploration followed by exploitation of the best button after finding it, and for the inefficiency of pursuing exploration when it is no longer needed. Luck may substitute for intentional exploration, but convergent-integrative thinking is still needed for exploiting this luck (serendipity). Because best behavior is known or obvious in simple repetitive tasks like our typing task, exploration by a creative mind is counterproductive for these tasks, at least in the short run. According to our data, grades in French middle schools are aligned with these rather straightforward tasks.

The purpose of the second study was twofold. In the first place, it provided a different way of testing that the sequence of divergent and convergent behavior that characterizes optimal learning in multi-armed bandit problems is an essential component of creative thinking. In this study, exploration opportunities were controlled by forcing adolescents to listen repeatedly to a selected musical genre so that high divergence of an individual would be counterproductive in the absence of a high ability for convergence. We did find, as expected, that highly divergent participants benefit from high convergence to enhance their appreciation of unfamiliar music. High convergence appears to be necessary for early adoption of innovations. Furthermore, by observing the same behavior in our two studies with very different production and consumption tasks, the theoretical association of creative cognition with multi-armed bandit problems, combining a task-specific sequence of divergence/exploration followed by convergence/exploitation, appears to be fairly general.

The returns to creative potential appear thus to be positive and high in novel settings but low and possibly negative in simple stationary environments. They remained significant after controlling for gender, personality, incentives, and other factors. Showing, for the first time, the *necessity* of a high creative potential for successful adaptation and decision making in novel uncertain environments provides a compelling policy argument in support of the screening and development by the educational sector of the creative potential of students. Therefore, our results enable us to provide an evidence-based answer to an important question in this 21st century: Do we *need* creativity? The results suggest that the answer is no for unqualified jobs and familiar activities; but certainly, yes for the growing proportion of qualified jobs and for successful adaptation to a rapidly changing world.

Funding: This work was supported by the French Research National Agency (ANR) under Grant number ANR-10-CREA-008.

Acknowledgements: We thank Gabrielle Tallon for her excellent contribution to experiment 1 and her participation in the first stages of this research, the Saint-Exupéry middle school in Rosny-sous-Bois for opening its doors to a time-consuming field experiment, Jean-Louis Tavani, Maud Besançon, and Baptiste Barbot for their essential collaboration to the realization of experiment 2 and many discussions, and Maxim Frolov for excellent programming.

Appendix

Study 1

Table A1: Payoff structure in the four incentive conditions for the typing task

Treatment	**Payoff**
Alone*Performance	3 correct codes=1 chocolate
Alone*Competition	The 4 highest-earning players in the session receive one pack of 34 chocolates each. Other players get nothing
Pair*Performance	6 correct codes per pair=1 chocolate each
Pair*Competition	The 2 highest-earning pairs in the session receive one pack of 34 chocolates each. Other players get nothing

Table A2: Cumulated distribution of points for each button over the 100 periods
"Equal solution" task

"Equal solution" task	**Buttons**			
Click	**A**	**B**	**C**	**D**
1-10	40	40	40	40
11-20	80	80	80	80
21-30	120	120	120	120
31-40	160	160	160	160
41-50	200	200	200	200
51-60	240	240	240	240
61-70	280	280	280	280
71-80	320	320	320	320
81-90	360	360	360	360
91-100	400	400	400	400
Outcomes from each button	(3, 4, 5)	(0, 2, 8)	(-4, 0, 6, 10)	(-8, 0, 16)
Distribution of outcomes in each button (%)	(38, 24, 38)	(23, 36, 41)	(26, 14, 24, 36)	(32, 27, 41)

Table A3: Cumulated distribution of points for each button over the 100 periods
"Best button" task

"Best button" task	**Buttons**							
Click	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**
1-10	8	10	10	12	16	15	10	10
11-20	12	20	62	32	56	45	4	40
21-30	11	30	126	60	64	83	-2	80
31-40	6	40	190	80	80	99	-8	120
41-50	3	50	260	100	112	121	-2	140
51-60	3	60	318	120	136	155	-4	160
61-70	0	70	382	140	152	168	2	180
71-80	1	80	452	160	160	196	-6	220
81-90	2	90	522	180	176	230	8	240
91-100	0	100	760	200	200	250	0	250
Outcomes from each button	(-1, 0, 1)	(1)	(0, 1, 7,200)	(1, 2, 4)	(0, 1, 8)	(0, 1, 3, 7, 9)	(-2, 0, 1, 2)	(0, 1, 10)
Distribution of outcomes in each button (%)	(32, 36, 32)	(100)	(1, 21, 77, 1)	(8, 88, 4)	(68, 8, 24)	(38, 23, 14, 20, 5)	(35, 26, 8, 31)	(66, 10, 24)

Notes: Cumulated points for every 10 clicks for the two buttons tasks. "Outcomes" refers to the possible outcomes hidden behind each button, and distribution refers to the frequency of appearance of each outcome (in %) if the button is selected.

Study 2

Table A4: Mean scores of the BB5 personality traits by gender

	Girls	Boys	Difference
Agreeableness (A)	67.69 (8.08)	62.43 (9.01)	P<0.00
Conscientiousness (C)	64.35 (10.17)	60.98 (8.78)	p=0.09
Extraversion (E)	64.26 (9.38)	62.51 (8.18)	p=0.29
Emotional Stability (ES)	64.47 (11.43)	64.45 (9.63)	p=0.80
Openness (O)	62.73 (7.64)	59.45 (7.94)	p=0.02
N	42	49	

Notes: Standard deviations in parentheses; Wilcoxon sign rank test of difference.

Table A4 reports gender gaps of personality in our sample of middle- school students in a poor social environment and the p-values of a Wilcoxon test. This table shows that 14-year-old girls in our sample are significantly more agreeable (p=0.005), more open (p=0.02), and slightly more conscientious than boys (p=0.09). There are no significant differences for extroversion and emotional stability between genders[17]. 14-year-old girls are more satisfied and more likely to adopt unfamiliar music than boys because they are significantly more agreeable, open and verbally creative than boys of that age.

[17] Previous studies made on adults found that women were more agreeable and more conscientious, and men more emotionally stable. No constant results are really found for openness and extraversion (see Costa Jr et al., 2001).

References

ADOBE (2013). *Barriers to creativity in Education: Educators and parents grade the system*. ADOBE publications.

Attanasi, G., Chessa, M., Gil-Gallen, S., & P. Llerena (2021). A survey on experimental elicitation of creativity in economics. *Revue d'Economie Industrielle* 174 (2), 273-324.

Barbot, B. (2012). BB5 : *Test de personnalité pour Adolescents*. Editions Hogrefe France.

Beaty, R.E., Benedek, M., Silvia, P.J., & D.L. Schacter (2016). Creative cognition and brain network dynamics. *Trends in Cognitive Sciences* 20 (2), 87-95.

Berlin, N., Tavani, J.L., & Besançon, M. (2016). "An exploratory study of creativity, personality, and schooling achievement", *Education Economics*, 24 (5), 536-556.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1), 49-59.

Burtini, G., Loeppky, J., & Lawrence, R. (2015). A survey of online experiment design with the stochastic multi-armed bandit. *arXiv*: 1510:00757.

Charness, G., & Grieco, D. (2019). Creativity and Incentives. *Journal of the European Economic Association* 17, 454-496.

Chermahini, S.A., & Hommel, B. (2010). The (b)link between creativity and dopamine: Spontaneous eye blink rates predict and dissociate divergent and convergent thinking. *Cognition* 115, 458-465.

Cohen, J.D., McClure, S.M., & Yu, A.J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society. B: Biological Sciences*, 362, 933-942.

Costa Jr, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: robust and surprising findings. *Journal of personality and social psychology*, 81(2), 322.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522-550.

Dul, J. (2016) Necessary Condition Analysis (NCA): Logic and methodology of "necessary but not sufficient" causality. *Organizational Research Methods* 19(1), 10-52.

Dul, J. (2020) *Conducting Necessary Condition Analysis*. Mastering Business Research Methods series, London: Sage Publications.

Dul, J., Karwowski, M., & Kaufman, J.C. (2020). Necessary Condition Analysis in Creativity Research. In: *Handbook of Research Methods on Creativity* (V. Dörfler and M. Stierand, eds.). Edward Elgar publishers.

Erdem, T., & Keane, M. P. (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science* 15, 1–20.

Finke, R.A., War, T.B., & Smith, S.M. (1992). *Creative Cognition: Theory, Research, and Applications*. Cambridge, MA: MIT Press.

Guilford, J.P. (1967). Creativity: Yesterday, today, and tomorrow. *Journal of Creative Behavior* 1, 3-14.

Hills, T.T. Todd, P.M., Lazer, D., Redish, A.D., Couzin, I.D., & The Cognitive Search Research Group (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Science*, 19, 46-54.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644-1655.

Karwowski, M., Dul, J., Gralewski, J., Jauk, E., Jankowska, D.M., Gajda, A., Chruszczewski, M.H., & Benedek, M. (2016). Is creativity without intelligence possible? A Necessary Condition Analysis, *Intelligence* 57, 105-117.

Karwowski, M., Kaufman, J. C., Lebuda, I., Szumski, G., & Firkowska-Mankiewicz, A. (2017). Intelligence in childhood and creative achievements in middle-age: The necessary condition approach. *Intelligence*, 64, 36-44.

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. *Review of general psychology*, *13*(1), 1-12.

Lee, J.G., Akinola, M., & Mason, M.F. (2017). 'Switching On' creativity: Task switching can increase creativity by reducing cognitive fixation. *Organizational Behavior and Human Decision Processes*, 139, 63-75.

Lévy-Garboua, L., Lohéac, Y., & Fayolle, B. (2006). Preference formation, school dissatisfaction and risky behavior of adolescents. *Journal of Economic Psychology*, 27(1), 165-183.

Lipshitz, R., & Strauss, O. (1997). Coping with uncertainty: A naturalistic decision-making analysis. *Organizational behavior and human decision processes* 69, 149–163.

Lubart, T.I. (2000). Models of the creative process: Past, present and future. *Creativity Research Journal*, 13 (3-4) , 295-308.

Lubart, T.I., Besançon, M., & Barbot, B. (2011). Evaluation of Creative Potential: Test and Manual. Paris: Editions Hogrefe France.

Lubart, T., Mouchiroud, C., Tordjman, S., & Zenasni, F. (2015). *Psychology of creativity*. 2nd Edition. Paris: Armand Colin.

Lucas, B., G. Claxton, & Spencer, E. (2013), « Progression in Student Creativity in School: First Steps Towards New Forms of Formative Assessments », *OECD Working Paper on Education* no. 86, Paris, OECD Editions.

McCrae, R. R., & Costa Jr, P. T. (2008). The five-factor theory of personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 159–181). The Guilford Press.

Mednick, S. (1962). The associative basis of creative problem-solving process. *Psychological Review* 69, 200-232.

Mekern, V., Hommel, B., & Sjoerds (2019). Computational models of creativity: a review of single-process and multi-process recent approaches to demystify creative cognition. *Current Opinion in Behavioral Sciences* 27, 47-54.

Mumford, M. D., Mobley, M. I., Reiter-Palmon, R., Uhlman, C. E., & Doares, L. M. (1991). Process analytic models of creative capacities. *Creativity Research Journal*, *4*(2), 91-122.

Navarro, D.J., Newell, B.R, & Schulze, C. (2016). Learning and choosing in an uncertain world: An investigation of the explore-exploit dilemma in static and dynamic environments. *Cognitive Psychology*, 85, 43-77.

Radel, R., Davranche, K., Fournier, M., & Dietrich, A. (2015). The role of (dis)inhibition on creativity: Decreased inhibition improves idea generation. *Cognition* 134, 110-120.

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, 24(1), 92-96.

Shi, B., Wang, L., Yang, J., Zhang, M., & Xu, L. (2017). Relationship between Divergent Thinking and Intelligence: An Empirical Study of the Threshold Hypothesis with Chinese Children. *Frontiers in Psychology*, 8, 254.

Smith, SM., Ward, T.B., & Finke, R.A. (Eds.) (2009). *The Creative Cognition Approach*, Cambridge, MA: MIT Press.

Sternberg, R. J., & Lubart, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity*. New York: Free Press.

Storme, M., Myszkowski, N., Çelik, P., & Lubart, T. (2014). Learning to judge creativity: The underlying mechanisms in creativity training for non-expert judges. *Learning and Individual Differences*, 32, 19-25.

Storme, M., Tavani, J. L., & Myszkowski, N. (2016). Psychometric properties of the French ten-item personality inventory (TIPI). *Journal of Individual Differences*, 37(2), 81-87.

Tynan, M.C., Credé, M., & Harms, P.D. (2020). Are individual characteristics and behaviors necessary-but-not-sufficient conditions for academic success? A demonstration of Dul's (2016) necessary condition analysis. *Learning and Individual Differences*, 77, 1-11.

Ward, T.B. (2007). Creative cognition as a window on creativity. *Methods*, 42, 28-37.