A DUAL-PROCESS MEMORY ACCOUNT OF HOW TO MAKE AN EVALUATION FROM COMPLEX AND COMPLETE INFORMATION

Ismaël Rafaï, Sébastien Duchêne, Eric Guerci, Ariane Lambert-Mogiliansky et Fabien Mathy

Presses de Sciences Po | « Revue économique »

2019/6 Vol. 70 | pages 1079 à 1093

ISSN 0035-2764
ISBN 9782724636000

Article disponible en ligne à l'adresse :

https://www.cairn.info/revue-economique-2019-6-page-1079.htm

A Dual-Process Memory Account of How to Make an Evaluation from Complex and Complete Information

Ismaël Rafaï*
Sébastien Duchêne**
Eric Guerci*
Ariane Lambert-Mogiliansky***
Fabien Mathy‡

Individuals are required to cope with uncertain, dispersed, incomplete, and incompatible sources of information in real life. We devised an experiment to reveal empirical "anomalies" in the process of acquisition, elaboration and retrieval of economic related information. Our results support the existence of a dual process in memory that is posited by the fuzzy-trace theory: acquisition of information leads to the formation of a gist representation which may be incompatible with the exact verbatim information stored in memory. We gave participants complex and complete information and then measured their cognitive ability. We conclude that individuals used their gist representation rather than processing verbatim information appropriately to make an evaluation. Finally, we provide evidence that subjects with low cognitive abilities tend to demonstrate more often this specific behavior.

UN PROCESSUS DE MÉMOIRE DUAL PERMET D'EXPLIQUER DES ÉVALUATIONS BASÉES SUR UNE INFORMATION COMPLEXE ET COMPLÈTE

Dans la vie réelle, les individus font face à des sources d'informations incertaines, dispersées, incomplètes et incompatibles. Nous proposons une expérience visant à révéler des « anomalies » dans le processus d'acquisition, d'élaboration et de récupération d'informations économiques. Nos résultats corroborent l'existence d'un processus de mémorisation dual proposé par la fuzzy-trace theory *: l'acquisition d'informations conduit à la formation de représentations «* gist *» qui peuvent être incompatibles avec l'exact verbatim des informations rencontrées et stockées en mémoire. Nous avons donné aux participants des informations*

* Département d'économie, GREDEG (UMR 7321, CNRS, Université de Nice Sophia Antipolis). *Correspondence:* 250 rue Albert Einstein, 06560 Valbonne, France. *Emails:* ismael.rafai@gredeg.cnrs.fr, eric.guerci@gredeg.cnrs.fr

** CEE-M, Université de Montpellier, CNRS, INRA, SupAgro, Montpellier. *Correspondence:* Faculté d'économie, Espace Richter, avenue Raymond Dugrand, CS 79606, 34960 Montpellier Cedex 2, France. *Email:* sebastien.duchene@umontpellier.fr

*** Paris School of Economics. *Correspondence:* 48 boulevard Jourdain, 75014 Paris, France. *Email:* alambert@pse.ens.fr

‡ Département de psychologie, Laboratoire BCL: Bases, Corpus, Langage (UMR 7320, CNRS, Université de Nice Sophia Antipolis). *Correspondence:* Campus SJA3, 24 avenue des diables bleus, 06300 Nice, France. *Email:* fabien.mathy@unice.fr

The authors declare no conflict of interest.

complexes et complètes, puis mesuré leurs capacités cognitives. Nous concluons qu'afin de procéder à une évaluation, les participants ont préféré l'usage de leurs représentations gist *à un traitement approprié de l'information verbatim. Enfin, nous montrons que ce comportement est plus présent chez les participants ayant des capacités cognitives moins élevées.*

Keywords: fuzzy-trace theory, memory, dual process, cognitive reflection test, bounded rationality

Mots clés: fuzzy-trace theory*, mémoire, processus duaux,* cognitive reflection test*, rationalité limitée*

JEL Codes: C91, D83, D89.

INTRODUCTION

Economic agents are overwhelmed by information, either as textual or visual content, which is easily and cheaply available via the Internet or social media.[1] Criteria and descriptions by which individuals use to make decisions are often complex in real life. Indeed, they need to cope with uncertain, dispersed, incomplete, and incompatible sources of information. People memorize each source of information and combine them by creating a mental picture that eventually has the virtue of simplifying the mass of information. In this paper, we investigated how people combine such heterogeneous and complex information to make an evaluation. We here call *evaluation*, the processes involved in the construction of beliefs about the characteristics of an object, based on available information. Such evaluation are constantly needed to make decisions (e.g., evaluate the state of wear of a used car before buying it or not). In the context of abundant information, the memory-reasoning relation plays a crucial role in the evaluation process. The issue is whether or not, each source of information simply adds a piece to the mental picture (as in a puzzle, the more data, the more precise the picture). If not, the pieces of information might interact and interfere, thus disrupting the mental picture.

Fuzzy-trace theory (hereinafter FTT) provides a psychological theory to deal with the cognitive processes involved in the construction of mental pictures (Reyna and Brainerd [1995]; Brainerd and Reyna [2001]; Liberali et al. [2012]; Reyna et al. [2016]). FTT is based on a dual cognitive process (gist and verbatim), which rests on the assumption that people form representations of an event both by identifying semantic features (gist traces, e.g., "this car is cheap") and by storing surface details (verbatim traces, e.g., "this car cost 2,000 euros"). The psychological literature reports that people tend to reason with gist rather than verbatim traces, that is, by attempting to create meaningful mental pictures which do not correspond to the sum of an event's surface details (Reyna [2012]). The theory has been proposed as an explanation of various phenomena such as the generation of false memory, risk perception and estimation, as well as general biases and fallacies in decision making.

1. For example, a weekday edition of *The New York Times* contains more information than the average person was likely to come across in a lifetime in 17th century England (Wurman [1989]).

For example, FTT predicts a tendency to prefer a sure outcome over a lottery in the gain domain and to prefer a lottery over a sure outcome in the loss domain. Indeed, Reyna [2012] considers that the simplest gist representation needed to compare a sure outcome with a lottery under the gain (loss) frame, consists in cognitively interpreting the sure outcome as "some positive (negative) outcome," and the lottery as "some positive (negative) outcome or a null outcome." If we consider the comparison in this way, with this wording, it is obvious that the sure outcome seems more (less) attractive than the lottery in the gain (loss) frame. Those predictions are consistent with the preference reversal observed in the "Asian Disease problem," where individuals are risk-adverse (risk-seeker) when the choice is presented in the gain (loss) frame (Tversky and Kahneman [1981]). However, the predictions offered by FFT, are not consistent with the well-known tendency individuals have to prefer sure small losses over unlikely high losses (e.g., paying an insurance premium) and to prefer unlikely high gains over sure small gains (e.g., buying a lottery ticket) (Kahneman and Tversky [1979]).[2]

In this paper, we let aside the question of comparing FTT with more traditional theories of decisions under risk and uncertainty that do take into account outcomes and probabilities magnitudes through probability weighting functions (see Kühberger and Tanner [2010] for a comparison of prospect theory and FTT applied to framing effects). Instead, we report and discuss the results of an experiment in which we tested the role of verbatim and gist memory traces in making an evaluation. It is worth noting that in our experiment, the description of the event is complete, in that we provide sufficient information to make a rational and objective evaluation and does not involve choices under risk or uncertainty. First, we tested whether the participants had stored the verbatim information required to provide the rational answer and then see if they use this information—as they should—to answer correctly. We found that only a small proportion of the participants answered correctly, knowing that a correct answer required a proper combination of the relevant memorized pieces of information. Instead, most of them seemed to rely more on gist representation rather than combining their verbatim traces. Finally, our results provide evidence that cognitive ability explains, to some extent, the misalignment between correct memory storage and the ability to answer correctly.

This paper is organized as follows. The experimental design is described in the second section. Behavioral hypotheses are presented in the third section. The results are analyzed and discussed in the fourth section, and the fifth section concludes.

EXPERIMENTAL DESIGN

The experiment was conducted from December 14 to 18, 2015 at the Experimental Economics Laboratory of Nice (LEEN, France). Three hundred and twenty-one students from a broad range of disciplines (45.65% were students

2. We thank an anonymous referee for pointing out what we consider as an important issue in FTT. This issue is due to the conjunction of 1) FTT assumes that individuals "have a preference for fuzzy processing at the lowest possible level" (Kühberger and Tanner [2010]) and 2) Reyna [2012] argues that the lowest gist representation needed to compare a certain outcome with a lottery does not depend on the outcomes and probabilities' magnitudes. However, those questions deserve to be investigated deeper than in the present article which does not focus on decision under risk and uncertainty.

in economics) at the University Côte d'Azur were recruited with a web-based online recruitment system (ORSEE, Greiner [2015]), and 19 one-hour sessions were run. Tasks were implemented on computers using the z-Tree software (Fischbacher [2007]). We provided calculators to each participant. Participants were remunerated for participating in the experiment. The payment included a fixed amount of €10, plus a performance-related amount of up to €6.[3]

The task consisted in reading a description of a country (Australia) without giving its name to participants. The information was presented in three or four consecutive screens, each providing a different perspective of the country:

– an *introductory perspective* (*IP*) screen, made up of 131 words, giving a kind of "touristic description" of the country;

– a *demographic perspective* (*DP*) screen, consisting of 92 words and one graph, providing demographic data of the country, including the number of inhabitants;

– an *economic perspective* (*EP*) screen with 81 words and two graphs focusing on the definition and value of the gross national income (*GNI*) of the country;

– and, depending on the treatment: either a *climate perspective* (*CP*) screen, containing climate (e.g., temperature) and geographical (desert, etc.) information of 122 words[4], or an *empty information* screen (*NP*) with a written message asking participants to wait.

Each screen lasted 70 seconds. We introduced an empty (*NP*) screen to compare the effect of additional climate information (*CP*) while keeping the total time constant. Apart from the *IP* screen, which was always displayed in the first position, each of the 12 possible screen orders was presented to participants.[5]

After reading all of the information, participants were asked to answer several questions successively. First, we asked participants to evaluate the welfare of the country described compared to their own country (France):

Q1: "According to you, on average, does an inhabitant of this country earn more or less money than an inhabitant of France?"[6]

Then, we asked them to reveal the country which fits the best the description, according to them:

Q2: "According to you, which country was described in the text?"

Question 1 (Q1) was asked to measure the evaluation of the country's wealth described in the experiment. Such evaluation can be carried out by the participant, either from the verbatim traces she has acquired during the experiment, or from the gist representation she has built, or by using both of them. Question 2 (Q2) allows us to measure specifically a proxy of the gist representation, while verbatim traces were elicited by administrating a set of memory *control questions*, in particular, about the total *gross national income* (GNI_1) and the *number of inhabitants* (POP_1) of the country.

3. Correct answers were paid €2 for Q1, €1 for Q2, and €3, for the control questions.

4. See Online Appendix V, DOI : 10.3917/reco.706.1079 for the text of the climate information in French.

5. See Online Appendix I for the complete list of screen order. Screen shots of the different screens are available in the Online Appendices III, IV and VI.

6. Participants may answer either "More" or "Less" by clicking on the respecting button. We counterbalanced the spatial location of the two buttons between participants.

We also elicited the participant's belief about the "*French average income*" (*F*), which was not included in the description. Nevertheless, we argue that students should know it approximately, and rational individuals should thus tend to use the verbatim information to answer Q1 correctly, while gist representation may interfere with the evaluation of non-rational participants. Indeed participants could combine the memorized verbatim information and their belief to answer Q1 in a rational way: 1) The *GNI* was introduced in *EP*; 2) The country's number of inhabitants was presented in *DP*. Using the calculator, and given the definition of the *GNI* as "the sum of all incomes earned in a year by the inhabitants of a country," participants should divide the *GNI* by the population to calculate *GNI* per capita, which indicates how much an inhabitant of the country earns in average.

Potential mistakes in Q1 might be explained by the fact that participants either did not memorize, had wrong beliefs about French average income, or did not combine the verbatim information given previously. Since we elicited verbatim memories of the gross national income (GNI_1) and population (POP_1), we could assess the impact of verbatim memory on the evaluation of the welfare of the country's inhabitants (through the probability of answering Q1 correctly). As a control of verbatim memory, we introduce the variable *MEM*, equal to 1 if the participant recalled both the population and *GNI*, and 0 otherwise.[7] Besides, to control for the heterogeneity about participants' beliefs of the French average income, we introduce and detail in the next section ("Behavioral Hypotheses") two variables (*R*1 and *R*2) that indicate respectively 1) if the participant answered Q1 correctly based on a correct combination of its verbatim traces and her belief about the French average income (*R*1) and 2) if the *GNI* per capita of the country guessed in Q2 is larger than the French average income (*R*2).

Finally, at the end of the experiment, we tested the participants' cognitive reflection abilities by means of the "Cognitive Reflection Test" (CRT hereafter), proposed by Frederick [2005].[8] We adopted the modified version of the CRT offered by Finucane and Gullion [2010] to ensure that students had never been exposed to the questions before.[9] We introduced the CRT since it is highly correlated with many robust and famous assessments of individual's cognitive abilities (e.g., the Wechsler abbreviated scales of intelligence (WASI) or the Scholastic Achievement Test (SAT), for more details, see Toplak, West and Stanovich [2014]). The CRT mainly evaluates the ability to resist the instinctive and wrong answers that come to our mind at first and choose the right and reflexive answer some time after. Participants with higher CRT scores tend to have many skills, like good numeracy, patience or rational reasoning skills. Therefore, we conjecture that high CRT participants, who are more reflexive and patient, with a good calculation capacity, answer Q1 based more frequently on their verbatim information rather than on their gist representation.

7. Participants were informed that they were entitled to a maximum 10% margin of error for these two control questions.

8. See Online Appendix VII for the CRT administered to our participants.

9. Indeed, the CRT is a well-known test. The modified CRT consists of three consecutive questions, each of them has an intuitive but false answer and a reflexive true answer. Scores were given by the number of correct answers, between 0 and 3.

During the whole experiment, participants were not allowed to take any notes, but had access to a basic calculator if they wished to process the numerical information that they could memorize verbally.

BEHAVIORAL HYPOTHESES

The correct answer to Q1 is "More" because the average income in Australia ($64,540) is greater than in France ($42,960). We score this answer as $Q1 = 1$ (respectively $Q1 = 0$, if the participant answered "Less").

H1: *Participants recalling the verbatim information had a higher probability of answering Q1 correctly.*

Since the information provided can be used to correctly compute the average income per inhabitant of the described country (Australia), we postulate that participants who are able to recall perfectly the *GNI* information and the number of inhabitants are more likely to answer Q1 correctly. Therefore, we expect a positive correlation between Q1 and *MEM*, the control dummy for verbatim memory.

Moreover, we conjecture that this correlation will be larger for participants with a higher CRT score. Besides, Toplak, West and Stanovich [2011] have shown that the CRT score is correlated with the probability of recalling verbatim information. We argue that this effect may play a role when retrieving the stored information to make an evaluation.

H2: *Participants rationally combined their verbatim memory and their beliefs to answer Q1.*

Another source of potential mistake in evaluation is the heterogeneity of participants' beliefs about French average income or their capacity to evaluate it. For example, someone with exact verbatim memory of both *GNI* and population of the described country, should rationally answer "Less" to Q1 if she overestimates French average income.

Thus, a rational individual should compare the ratio between Australian's GNI_1 and POP_1 she recorded in memory, with her belief about French average income (F). Based on the belief elicited about French average income and the information revealed, we construct the variable $R1$ which indicates the sense of the inequality:

$$R1 = 1 \Leftrightarrow \frac{GNI_1}{POP_1} > F \text{ and } R1 = 0 \Leftrightarrow \frac{GNI_1}{POP_1} < F.$$

Thanks to this method, we were able to distinguish if the suject's answer is consistent with a rational combination of the memorized information and her belief. We therefore expected a positive correlation between $R1$ and Q1.

H3: *Participants with a "rich-country" gist representation had a higher probability of answering Q1 correctly.*

FTT suggests that individuals are using a gist representation rather than the verbatim memory to make an evaluation. To estimate the information contained in participants' gist representation, we used the answers to Q2 about the country described in the experiment. We are aware that even if the answer to this

question did not exactly reveal the participants' gist representations, in many cases it should indicate the most likely representation.

In particular, we used World Bank data to extract GNI_2 and POP_2 from the country revealed by participants. Then we compared the $\frac{GNI_2}{POP_2}$ ratio with participants' beliefs about the French average income. The variable $R2$ shows the sense of this inequality:[10]

$$R2 = 1 \Leftrightarrow \frac{GNI_2}{POP_2} > F \text{ and } R2 = 0 \Leftrightarrow \frac{GNI_2}{POP_2} < F.$$

Indeed, it is impossible to be sure that the gist representation of the inhabitants' welfare of the described country corresponds to the per capita *GNI* (using World Bank data) of the country revealed in Q2. For example, 1) one can think that South Africa is most likely the described country in many aspects but not concerning inhabitants' welfare; 2) gist representation of South Africa welfare may be lower or greater than it is in reality. Therefore, these two phenomena generate noise and reduce the probability to detect a potential effect, leading to conservative results, due to a noisy proxy of gist representation.

Despite this, if the probability of answering Q1 correctly, given that $R2 = 1$, is greater than the probability of answering it correctly, given that $R2 = 0$, then it suggests that participants were using their gist representation in the evaluation of inhabitant's welfare of the described country.[11]

H4: *Climate information impacts gist and verbatim memory.*

Verbatim reasoning about *GNI* and *POP* should not be affected by any other information. In contrast, the gist representation of inhabitants' welfare is built from all the received information (demographic, economic, climatic, and touristic descriptions). For example, the climate information, that describes the country as hot ("up to 50 degrees Celsius") and "mostly desertic," can interfere with the gist memory of the country, leading to a representation of a country of a lower income. Indeed, we deliberately chose a country with these characteristics because hot temperatures and deserts could be associated with poorer countries.

To understand the formation of this dual process, we investigated the impact of additional climatic information on both verbatim and gist representations. According to the treatments, participants received either a blank screen (*NP*) or one with climate information (*CP*). Since participants with the *CP* screen received more information, and specifically climate information, it should be

10. Therefore, GNI_2 corresponds to the gross national income of the country answered in Q2, and POP_2 to the population of this country. For example, if someone answered "South Africa" to Q2 and thought that the average income in France was $25,400 per year, we extracted the GNI_2 and POP_2 (respectively 3.233×10^{11} and 50.52×10^6) using the World Bank Data API and compared the ratio with her belief about French average income as follows: $\frac{GNI_2}{POP_2} = 6,399 < 25,400 = F \Rightarrow R2 = 0.$

With the same answer to Q2 but with a lower belief about French welfare, one would obtain $R2 = 1$. Since the answer to Q2 could be freely input, we excluded from our analysis the 41 participants who did not report a real country (e.g., "a country in Africa," or "Brazil or China," etc.).

11. It is equivalent to looking at whether the probability of not responding correctly to Q1, given $R2 = 1$, is lower than the probability of not responding correctly to Q1, given $R2 = 0$.

more difficult for them to remember economic and demographic verbatim information, compared to participants with the blank screen.

RESULTS

In our sample, only 28% correctly answered Q1 and 9% found the right country.[12] These low rates of correct answers for both Q1 and Q2 can be explained by 1) the large number of countries in the world (193) and 2) our experimental design that voluntarily describes a "hot and desert country," so as to make participants think of a poorer country than Australia, and therefore create a gap between the gist representation and verbatim information.

According to the CRT score, we divided participants among two groups: those with a CRT score of 0 or 1 are hereafter called the "LOW" group $(n = 266)$ and those with a CRT score of 2 or 3 the "HIGH" group $(n = 55)$.

56.1% of the participants recalled correctly the *GNI* of the country and 51.4% recalled its number of inhabitants. 32.4% of the participants recalled both pieces of information.[13]

H1: *Participants recalling the verbatim information had a higher probability of answering Q1 correctly.*

We reported in Table 1, the cross tabulation between verbatim memorization and answer to Q1, depending on CRT score.

Table 1. *Q1 and MEM contingency*

	LOW		HIGH		TOTAL	
	$MEM = 0$	$MEM = 1$	$MEM = 0$	$MEM = 1$	$MEM = 0$	$MEM = 1$
$Q1 = 0$	135	59	24	12	159	71
$Q1 = 1$	52	20	6	13	58	33
Mean	27.81%	25.32%	20.00%	52.00%	26.73%	31.73%

12. Top answered countries were Brazil (30), Australia (29), "Africa" (18), China (14), Morocco (13), Egypt (12), South Africa (12), USA (12), and India (11).

13. An economics student might have a potential advantage, since she might be more able to understand, recall and combine economic information such as *GNI* if compared to students from other disciplines. We do not collect individual data, but checked at the session level if the percentage of economics students impacts the proportion of correct answers to Q1 and the proportion of participants being able to correctly recall economic-related information. We did not find any impact of the proportion of economics students in a session (45% in the whole sample), 1) on the proportion of participants correctly answering Q1 (OLS regression: coeff. $= 0.0751$, $t = 0.564$, $p = 0.580$) nor 2) on the proportion of participants correctly recalling the verbatim information related to the economy, more precisely concerning the recall of the industry's share of *GNI* (OLS regression: coeff. $= -0.12185$, $t = -0.869$, $p = 0.397$) and the recall of the *GNI* (OLS regression: coeff. $= -0.02916$, $t = -0.149$, $p = 0.883$). Regarding the low percentages of participants from other disciplines in a session, we unfortunately cannot investigate the effect of studying other disciplines.

We did not find any significant difference in the probability to answer Q1 correctly, depending on the correct recall of verbatim information (*t*-test: $t = -0.911$, $p = 0.363$). We conducted a logistic regression (Table 4, Models 4-6) on the probability of answering Q1 correctly, as a function of *MEM* and CRT score. We did not find any effect of recalling verbatim information on the correctness of Q1 for LOW participants (cf. Model 3, coeff. $= -0.01$, $p = 0.975$) but found a statistically significant one for HIGH participants (cf. Model 2, coeff. $= 2.01$, $p = 0.0027$). Indeed, HIGH participants who recall the right verbatim information have a higher probability to answer Q1 correctly.[14]

Since a high CRT measures the individuals' ability not to respond intuitively but in a thoughtful and rational way, this result may suggest that individuals with high CRT scores tend to resist an initial intuitive response (which may be based on a gist representation) to use the stored verbatim information instead and then make the calculation required to answer Q1. Another hypothesis, the effects of which could be added to the previous one, although less likely in our view, would be that students with low CRT scores would have more difficulty dividing *GNI* by the number of inhabitants, since CRT is also correlated to the individual's overall mathematical level. However, since many students are enrolled in economics or quantitative disciplines, we believe that the difficulty of calculating average income, with a calculator, is not the main cause of failure to answer Q1 correctly.

H2: *Participants rationally combined their verbatim memory and their beliefs to answer Q1.*

We reported in Table 2, the cross tabulation between $R1$ and answer to Q1, depending on CRT score.

Table 2. *Q1 and R1 contingency, conditional to CRT*

	LOW		HIGH		TOTAL	
	$R1 = 0$	$R1 = 1$	$R1 = 0$	$R1 = 1$	$R1 = 0$	$R1 = 1$
$Q1 = 0$	92	102	18	18	110	120
$Q1 = 1$	34	38	5	14	39	52
Mean	26.98%	27.14%	21.73%	43.75%	26.17%	30.23%

If individuals rationally processed their verbatim memory to answer Q1, we should expect that the probability of answering Q1 correctly, given that $R1 = 1$, should be greater than the probability of answering Q1 correctly, given that $R1 = 0$.

We did not find any significant difference in the probability to answer Q1 correctly, depending on $R1$ (t-test: $t = -0.805$, $p = 0.421$). We have estimated logistic regressions (Table 4, Models 4-6) on the probability of answering Q1, as a function of $R1$ and CRT score. We found that the results obtained in H1 are robust to heterogeneity of beliefs about French average income. LOW

14. Models 2 and 3 are equivalent and statistically better than Model 1 (Likelihood ratio test: $p = 0.012$).

participants did not seem to process verbatim information to answer Q1 (Model 6, coeff. = 0.01, $p = 0.972$). Conversely, HIGH individuals did so (Model 5, coeff. = 1.68, $p = 0.0157$), confirming the first hypothesis H1, regarding the significant correlation between the CRT scores and the students' ability to answer Q1 consistently with a correct combination of the verbatim information.[15]

H3: *Participants with a "rich-country" gist representation had a higher probability of answering Q1 correctly.*

We reported in Table 3, the cross tabulation between *R*2 and answer to Q1.

Table 3. *Q1 and R2 contingency*

	LOW		HIGH		TOTAL	
	$R2 = 0$	$R2 = 1$	$R2 = 0$	$R2 = 1$	$R2 = 0$	$R2 = 1$
Ql = 0	114	48	26	6	140	54
Ql = 1	27	41	8	10	35	50
Mean	19.15%	46.07%	23.52%	62.50%	20.00%	48.08%

We have found a significant difference in the probability to answer Q1 correctly, depending on having a "rich-country" gist representation (*t*-test: $t = -4.958$, $p = 2.6 \times 10^{-6}$). Since our measure *R*2 is a noisy proxy of the wealth measure of the gist representation, it is likely that those results are conservative and that our study does underestimate the effect of the gist representation in the evaluation of the country.

We have estimated a logistic regression (Table 4, Models 7-10) on the probability of answering Q1, as a function of *R*2, *R*1 and CRT score. We have found that the gist representation impacted the evaluation for both LOW (Model 8: coeff. = 1.30, $p < 0.001$) and HIGH participants (Model 7: coeff. = 1.76, $p = 0.021$). However, we haven't found any statistical difference of this effect between the two groups ($p = 0.58$).[16]

Moreover, we have analyzed the answers of the 41 participants who failed to report a real country (e.g., they reported a whole continent as Africa, South America, or more than one country, etc.). The probability to report such inconsistent answers does not depend on the CRT (9.09% for HIGH vs. 13.53% for LOW [*t*-test: $t = -1.000$, $p = 0.320$]) nor on the fact that participants recall the Economic and Demographic verbatim information (13.36% for $MEM = 0$ vs. 11.54% for $MEM = 1$ [*t*-test: $t = 0.467$, $p = 0.641$]). Participants reporting an inconsistent answer to Q2 have a lower probability to answer Q1 correctly (30.71% vs 12.20% [*t*-test: $t = 3.158$, $p = 0.002$]).

15. We ran a likelihood ratio test, and Model 6 is statistically better than Model 4 ($p = 0.038$).

16. The latter regression also confirms H2. These results are robust for the use of *MEM* (see Online Appendix II) instead of *R*1 as a control of verbatim memory, which provides support to results presented in H1.

Table 4. *Impact of verbatim and gist memory on Q1*

	$P(Q1=1)$									
	H1			H2			H3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
CRT Ref. Group	All	HIGH	LOW	All	HIGH	LOW	HIGH	LOW	HIGH	LOW
(Intercept)	– 0.93**	– 1.29*	– 0.60	– 0.94*	– 1.26	– 0.59	– 0.56	– 0.86*	– 1.76*	– 0.86
	(0.35)	(0.62)	(0.39)	(0.37)	(0.67)	(0.41)	(0.56)	(0.44)	(0.82)	(0.48)
MEM	0.45	2.01*	– 0.01							
	(0.28)	(0.67)	(0.33)							
*R*1				0.30	1.68*	0.01			1.87*	– 0.08
				(0.27)	(0.69)	(0.30)			(0.79)	(0.33)
*R*2							1.70*	1.30***	1.76*	1.30***
							(0.73)	(0.33)	(0.76)	(0.34)
LOW		0.69			0.67		– 0.30		0.90	
		(0.54)			(0.61)		(0.49)		(0.78)	
HIGH			– 0.69			– 0.67		0.30		– 0.90
			(0.54)			(0.61)		(0.49)		(0.78)
MEM × LOW		– 2.02**								
		(0.74)								
MEM × HIGH			2.02**							
			(0.74)							
*R*1 × LOW					– 1.67*				– 1.95*	
					(0.75)				(0.85)	
*R*1 × HIGH						1.67*				1.95*
						(0.75)				(0.85)
*R*2 × LOW							– 0.39		– 0.46	
							(0.81)		(0.84)	
*R*2 × HIGH								0.39		0.46
								(0.81)		(0.84)
Logistic regressions. Treatment controls are masked.										
Log Likelihood	– 176.59	– 170.86	– 170.86	– 177.22	– 172.60	– 172.60	– 145.37	– 145.37	– 142.20	– 142.20
No. obs.	321	321	321	321	321	321	280	280	280	280

Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

H4: *Climate information impacts gist and verbatim memory.*

Table 5 displays the frequencies of recalling both *GNI* and *POP* verbatim information (i.e., $MEM = 1$), as well as the probability of recalling a "richer country than France" (i.e., $R2 = 1$), according to *CP*.

Table 5. *MEM, R2 and CP contingency, conditional to CRT*

	LOW		HIGH		TOTAL	
	$CP = 0$	$CP = 1$	$CP = 0$	$CP = 1$	$CP = 0$	$CP = 1$
$MEM = 0$	48	139	11	19	59	158
$MEM = 1$	18	61	3	22	21	83
Mean	27.27%	30.5%	21.42%	53.66%	26.25%	34.44%
$R2 = 0$	34	107	6	28	40	135
$R2 = 1$	28	61	8	8	36	69
Mean	45.16%	36.31%	57.14%	22.22%	47.37%	33.82%

We did not find any significant difference in the probability to recall verbatim information, depending on receiving the climate information (*t*-test: $p = 0.162$). However we found a significant effect of receiving the climate information on the probability of recalling a "rich country" (*t*-test: $p = 0.044$).

We ran logistic regressions (Table 6, Models 11-13) on the probability of recalling both *GNI* and *POP* verbatim information (*MEM*), as a function of the CRT score and receiving the climate information (*CP*). We did not find that providing climate information to LOW participants impacted their probability of recalling verbatim economic and demographic information (Model 12: coeff. $= 0.16$, $p = 0.62$), as opposed to the HIGH group for which it increases their recall rate (Model 13: coeff. $= 1.45$, $p = 0.045$). We anticipated that giving more information would reduce the ability to remember demographic and economic data. Nevertheless, climate information (compared to a blank screen) may help HIGH participants to remain focused on the country data.

We also ran logistic regressions on the probability of identifying a richer country than France (*R2*), as a function of *CP* and CRT (Table 6, Models 14-16). We found that the effect of climate information on the gist representation is more pronounced for the HIGH group (Model 16: coeff. $= -1.54$, $p = 0.022$) than for the LOW one (Model 15: coeff. $= -0.37$, $p = 0.222$). This tends to support the idea that climate information more strongly impacts the gist representation of HIGH participants.

Moreover, the introduction of the climate information also increases the probability for a participant to give an inconsistent answer to Q2 (5.00% for $CP = 0$ vs. 15.35% for $CP = 1$ [*t*-test: $t = -3.062$, $p = 0.002$]). This latter result would suggest that the climate information creates some confusion in the gist representation.

Table 6. *Impacts of climate information on gist and verbatim memory*

	$P(MEM=1)$			$P(R2=1)$		
	(11)	(12)	(13)	(14)	(15)	(16)
(Intercept)	– 1.03***	– 0.98***	– 1.30*	– 0.11	– 0.19	0.29
	(0.25)	(0.28)	(0.65)	(0.23)	(0.26)	(0.54)
CP	0.39	0.16	1.45*	– 0.57*	– 0.37	– 1.54*
	(0.29)	(0.32)	(0.72)	(0.27)	(0.30)	(0.67)
HIGH		– 0.32			0.48	
		(0.71)			(0.60)	
CP × HIGH		1.29			– 1.17	
		(0.79)			(0.74)	
LOW			0.32			– 0.48
			(0.71)			(0.60)
CP × LOW			– 1.29			1.17
			(0.79)			(0.74)
Log Likelihood	– 201.23	– 197.26	– 197.26	– 183.10	– 181.38	– 181.38
No. obs.	321	321	321	280	280	280

Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Exploratory results: *Climate information reduces existing order effect.*

Since we proposed all different combinations of information order through 12 treatments, it is possible to investigate this effect on the participants' evaluations. We first run a Fisher's exact test on all the treatments and found a significant treatment effect on the probability to correctly answer Q1 ($p = 0.009$). However, this effect is only present for the treatments in which the climate information is not given (Fisher's exact tests: $p = 0.006$ when *CP* is missing and $p = 0.187$ when *CP* is displayed). Adding climate information seems to lessen or cancel the order effect. This point could be the subject of further research. Indeed, in the treatments with no climate information, the participants are more likely to answer Q1 correctly when the economic information is given before the demographic information ([logistic regression: coeff. = 1.186, $z = 2.331$, $p = 0.020$]). While our design is not adapted to understand why this order effect occurred, one explanation could be that it is easier to divide the first information (number) received by the last one, than the other way around. Therefore, it would be interesting to test in subsequent research whether displaying the numerator first (*GNI*) and then the information from the denominator (population) helps participants to perform the calculation rather than the reverse order.

DISCUSSION

In this experiment, participants received several pieces of information on different aspects of an unknown country. After receiving the description, they were asked to evaluate the average income of the inhabitants of this country compared to that of their country of residence (France). Our experimental setting offered complete information since the description of the country included

all of the relevant information to make an exact evaluation. We estimated both verbatim and gist memories by asking participants to reveal: 1) recorded information about economics and demographics, and 2) the country that best fitted the description according to them.

To summarize, the results of the present experiment can be accounted for the dual memory-reasoning process posited by the fuzzy-trace theory (FTT), as participants mainly relied on their gist information to evaluate the described country's welfare. This is principally true for participants with a low cognitive reflection ability. However, participants with a higher cognitive reflection ability tended to make more often evaluations consistent with stored verbatim information. Our results provide evidence for a dual memory process (gist and verbatim) in the specific context of making an evaluation with complete and complex information.

While the FTT provides a convincing explanation of our specific experiment, we do not argue that the theory is superior in the context of decision under risk and uncertainty, to more traditional approaches based on probability theory, such as expected utility theory or prospect theory. Indeed, as mentioned in the Introduction, our study does not consider the weakness of the FTT in the specific context of "decision under risk or uncertainty." In this respect, it would be valuable to do some new research and experimentally test the various potential limitations of the FTT and the underlying vulnerabilities in relation to the prospect theory.

However, we believe our protocol makes it possible to distinguish verbatim and gist traces of economic and demographic data, so as to investigate which information was in line with the final evaluation. Furthermore, we controlled for participants' cognitive abilities using the Cognitive Reflection Test. We found that participants with low cognitive abilities did not combine the verbatim information to make the evaluation. Indeed, although they correctly recalled the gross national income and the population of the country, they did not make a consistent assessment with an appropriate combination of this information. Conversely, individuals with higher cognitive abilities tended to exhibit more consistency between their evaluation and verbatim memory. Furthermore, regardless of their cognitive abilities, participants relied mainly on their gist representation.

In addition, investigating the impact of additional climatic information on verbatim and gist memory leads to novel results. For individuals with low cognitive abilities, we did not find that climate information impacted gist and verbatim traces of economic and demographic information. However, it helped high cognitive participants to recall verbatim information and changed their gist representation about the inhabitants' welfare. These results suggest further research to better understand both the link and balance between cognitive abilities and the coexistence of gist and verbatim reasoning.

REFERENCES

Brainerd, C. J. and Reyna, V. F. [2001]. "Fuzzy-Trace Theory: Dual Processes in Memory, Reasoning, and Cognitive Neuroscience," *Advances in Child Development and Behavior*, 28: 42–100.

Finucane, M. L. and Gullion, C. M. [2010]. "Developing a Tool for Measuring the Decision-Making Competence of Older Adults," *Psychology and Aging*, 25 (2): 271–288.

Fischbacher, U. [2007]. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments," *Experimental Economics*, 10 (2): 171–178.

Frederick, S. [2005]. "Cognitive Reflection and Decision Making," *The Journal of Economic Perspectives*, 19 (4): 25–42.

Greiner, B. [2015]. "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE," *Journal of the Economic Science Association*, 1 (1): 114–125.

Kahneman, D. and Tversky, A. [1979]. "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47 (2): 263–292.

Kühberger, A. and Tanner, C. [2010]. "Risky Choice Framing: Task Versions and a Comparison of Prospect Theory and Fuzzy-Trace Theory," *Journal of Behavioral Decision Making*, 23 (3): 314–329.

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M. and Pardo, S. T. [2012]. "Individual Differences in Numeracy and Cognitive Reflection, with Implications for Biases and Fallacies in Probability Judgment," *Journal of Behavioral Decision Making*, 25 (4): 361–381.

Reyna, V. F. [2012]. "A New Intuitionism: Meaning, Memory, and Development in Fuzzy-Trace Theory," *Judgment and Decision Making*, 7 (3): 332–359.

Reyna, V. F. and Brainerd, C. J. [1995]. "Fuzzy-Trace Theory: An Interim Synthesis," *Learning and Individual Differences*, 7 (1): 1–75.

Reyna, V. F., Corbin, J. C., Weldon, R. B. and Brainerd, C. J. [2016]. "How Fuzzy-Trace Theory Predicts True and False Memories for Words, Sentences, and Narratives," *Journal of Applied Research in Memory and Cognition*, 5 (1): 1–9.

Toplak, M. E., West, R. F. and Stanovich, K. E. [2011]. "The Cognitive Reflection Test as a Predictor of Performance on Heuristics-and-Biases Tasks," *Memory & Cognition*, 39 (7): 1275–1289.

Toplak, M. E., West, R. F. and Stanovich, K. E. [2014]. "Assessing Miserly Information Processing: An Expansion of the Cognitive Reflection Test," *Thinking & Reasoning*, 20 (2): 147–168.

Tversky, A. and Kahneman, D. [1981]. "The Framing of Decisions and the Psychology of Choice," *Science*, 211 (4481): 453–458.

Wurman, R. S. [1989]. *Information Anxiety*. New York: Doubleday.