Can an Individual Cognitive

Thinking Style and Decision Making Ability Predict

Performance on CRT?

Abstract

Dual processing Theory is an important part of rational decision making (Stanovich & West, 2000). This involves two systems; *System one* and *System two* that work together to solve simple and complex problems. This study looked at Actively Open-Minded Thinking Scale, Temporal Discounting and The Cognitive Failures Questionnaire and if they would predict scores on the Cognitive Reflection Test (CRT; Frederick, 2005). The participants are 167 university students aged between 19-67. The results show that the model is significant. However, only one predictor significantly predicted scores on CRT. This provides added support to past research that has found similar findings. Future research could expand this study by re-evaluating these predictors and look at new predictors that show a relationship CRT and cognitive reasoning tools.

Introduction

The ability to think logically is an important part of being human. There are two different systems at work, *System one* and *System two* this theory also known as dual-processing (Stanovich & West, 2000). Dual processing works on the basis that system one is unconscious, implicit, requiring very little cognitive expenditure (Mishra, Mishra & Nayakankuppam, 2007) and the individual is not aware of this process. This intuitive based system has shown that the response that a participant gives to an answer is one that the participant expects to be correct, even if the answer is wrong. This can happen if the participants unconsciously reword or skip words (attribute-substitution) to make a problem or question easier to understand (Cosmides & Tooby, 1996). Although, this is not always the case and would depend on the complexity of the situation and the problem. If the problem is too complex then system two would be used in replace of system one. System two is conscious, explicit; however, this system requires more concentration and effort which has a limited capacity (Evans, 2008) this would make it difficult to carry out multiple tasks (Toplak, West & Stanovich, 2014).

This analytical system allows a person to answer questions that system one in most cases gets wrong. However, repetition may allow an individual to become better at the task therefore less attention is required (Evans, 2008), allowing for an automated response using system one instead of system two. Therefore, learned behaviour can decrease cognitive load and dependency on system two. For instance an individual solving a math problem the first time without a calculator will rely on system two, but completing this problem multiple times the individual will already know the answer and

the process would become automated like a person knowing their address or phone number and not having to think about it first.

Evolutionary psychologists have suggested that system one is much older then system two; and is responsible for processing information that is frequently (Cosmides & Tooby, 1996) instead of using system two that is much slower; this evolutionary processes serve the purpose *"to think on ones feet"*, if the person does not have time to think logically about a problem and just makes a snap judgement (De Neys, 2012) in a life of death situation it would be better making a decision then not making one regardless of whether it is right or wrong.

Researchers have looked at this ability to make decisions using system one and two. Frederick (2005) developed the Cognitive Reflection Test (CRT) to test system one and system two. This is done by asking logic-based problems that appear to have a simple (intuitive) answer but instead has a more complex (analytical) answer that requires system two to solve. He studied participants at universities in America. The scores from the CRT were then compared to other standardised logic and intelligence tests, to see if CRT could predict performance on these tests. The participants who scored high on the CRT made more logical decisions when confronted with other cognitive problems, participants who scored low made poorer logical decisions. One area that did show a difference in genders; men scored (1.47) on average and women scored (1.03) on the three problems CRT.

Toplak, West, & Stanovich (2014) carried out a similar study to Frederick (2005) although Toplak et al., (2014) increased the number of questions from three in the CRT

to seven, as the three problems were well known. Participants took the CRT and then these scores were compared to predict how participants would perform on other logical tests (Cognitive ability, Thinking disposition measure, Superstitious thinking, to name few). The researcher found that their results were very similar to Frederick (2005) study. They found that certain questions on the CRT were better predictors than others. This may be down to the fact that the first three questions were well known, however, the new questions were better predictors of logic and reasoning.

Toplak et al., (2014) and Frederick (2005) have both looked at similar predictors and have found inconsistent results, on one predictor in particular temporal discounting (TD), Frederick found this predictor to be overall significant but with some question showing a non-significant relationship between CRT and TD. Toplak et al., (2014) found similar issues with this predictor and the relationship with certain questions on the CRT. Campitelli & Gerrans (2014) study looked at CRT and how well this tool works in predicting logical and rational thinking and mathematical ability. They used the scores on the CRT and compared them to performance on three problems and decision questionnaires (Numeracy (NUM), Syllogistic reasoning with belief Bias (SRBB) and actively open-minded thinking scale; AOT). They found that there was a stronger relationship between NUM mathematical ability and CRT. The relationship between AOT was slightly weaker but was still significant overall. An issue with this study is that the CRT used with the three-question test, which was an issue stated in Toplak et al., (2014) study. The questions have been widely published in journals and books. So these questions may be known to some participants already, however, this may not have been

the case as participants scored similarly to Frederick (2005) study.

The issues with these studies are the fact that they rely on logic-based problems, that does not necessarily mean that everyone that scored average or low on the CRT is below average intelligence or will make poor or irrational decision on another task, also the opposite may be true individuals that score high may not have better decision-making skills and may still make poor choices. They may be better at predicting risk & reward but overall there is still a risk of a poor outcome (Reyna & Brainerd, 2011).

Both Frederick and Toplak studies have compared TD as predictors and both have shown significant results although with some issue concerning the questions. Toplak et al., (2014) and Campitelli et al., (2014) studies look at another predictor AOT and both showed significant results. This study will reexamine the relationship between CRT and its relationship between AOT and TD. This study will also introduce a new predictor Cognitive failure questionnaire (CFQ) that has had very little research showing whether there is a relationship between CRT and CFQ.

H1: Can Actively Open-Minded Thinking Scale, Temporal Discounting and Cognitive Failures Questionnaire significantly predict performance on Cognitive Reflection Test (CRT)

Method

Design

The study was a repeated measure design, all the participants completed all six of the cognitive tests. The DV is the CRT and the predictors are Actively Open-Minded Thinking Scale, Temporal Discounting and The Cognitive Failures Questionnaire. The question was counterbalanced with all participants answering the questionnaires in different orders.

Participants

The sample consisted of 147 participants 21 (14%) male and 126 (86%) female. The age range is 19 to 67 with the (M_{age} = 23.44) and (SD = 8.51). The participants were all 2nd year psychology students, studying at the University of Bedfordshire. Participants were recruited during their research method tutorial.

Materials

The materials used in the study consisted of six tests; **The Cognitive Reflection Test** (CRT) (Frederick, 2005) an example from the test is *"If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?"* (8 questions). **Actively Open-Minded Thinking Scale** (Stanovich & West, 2007) An example is *"Abandoning a previous belief is a sign of strong character"* participants were given responses to choose from (*"Disagree strongly, Disagree moderately,*

Disagree slightly, Agree slightly, Agree moderately and *Agree strongly"*) (41 questions).

Temporal Discounting (Toplak, West, & Stanovich, 2014) an example is *"Would you*

prefer £3400 this month or £3800 next month", with the responses *"I strongly prefer*

£3400 this month, I slightly prefer £3400 this month, I prefer £3400 this month, I prefer

£3800 next month, I slightly prefer £3800 next month and *I strongly prefer £3800 next*

month" (5 question) and **The Cognitive Failures Questionnaire** (Broadbent et al., 1982)

an example is *"Do you fail to notice signposts on the road?",* the responses *"very often,*

quite often, Occasionally, very rarely and *never"* (25 questions).

Procedure

The survey was done on a computer using online survey software qualtrics. The

ethical consideration was given to the participants on the first page of the computer

screen. These were all data will be kept confidentially, the data will be stored on a

password-protected computer for a minimum of 3 years. This gave the participant a

choice to complete the experiment but not have the result used in the study, the

participants could choose to have the result included in the study or not to take part in the

study. One of the options must be chosen before the participant moves past the page and

before taking part in the experiment. If the participant chose to partake in the study they

would then answer all four of the questionnaires, the participants were then debriefed

after the study had finished.

Result

A multiple regression analysis was conducted to predict AOT, TD and CFQ
performance on the CRT. The result shows that 3.4 percent of males and 26.53 percent of
females failed to correctly answer a question on the CRT. Table 1 shows the relationship
between the mean and standard deviation for each of the predictors and the DV in
relation to males and female scores on each of the predictors. The range for CRT is 0 to
8, this suggests that participants are not very analytical when answering logic and
math-based questions. The mean for TD shows that the majority of participants choose to
delay a financial reward as the range for this task was 5 to 30, the SD shows that fewer
participant scores at the low end of the task but closer to the midpoint or just over.

Table 1:

The mean, Standard Deviation, Range and Skew for the three predictors.

	Mean			Std. Deviation			Range	skew
	Total	Male	Female	Total	Male	Female		
CRT[A]	1.43	1.71	1.38	1.60	1.79	1.53	0 - 7	1.58
AOT[B]	162.00	163.47	161.60	20.21	20.47	20.24	112 - 215	-0.12
TD[C]	18.10	16.90	18.29	5.20	2.89	5.43	5 - 30	-0.19
CFQ[D]	78.22	82.19	77.55	12.00	11.42	11.97	43 - 107	-0.32

A - Cognitive Reflection Test
B - Actively Open-Minded Thinking Scale
C - Temporal Discounting
D - Cognitive Failures Questionnaire

Table two shows that means and standard deviation for participants and score one or more and participants that score zero, the result suggests that participants that scored one or more scored higher on each of the three variables compared to participants that scored zero. This suggests that participants that scored one or more reported higher levels of Temporal Discounting and were more open minded compared to participants that scored zero. However, participants that scored zero reported lower cognitive failures.
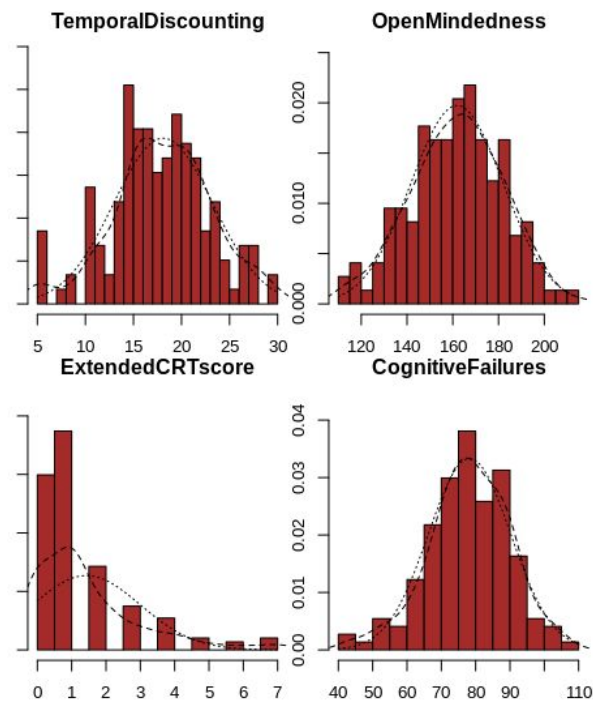
Table 2:
This table shows the means and standard deviation for the participants that score zero and the participants that scored 1 or more on the CRT test.

	Participants scored Zero		Participants scored One or More	
	Mean	SD	Mean	SD
TD	17.56	4.49	18.32	5.43
AOT	158.09	21.94	163.48	19.32
CFQ	76.36	13.66	79.00	11.14

Figure one shows the distribution of each of the three variables, three of the variables show a normal distribution. However, CRT shows a distinct positive skew towards the lower end of the scores, suggesting that participants struggled with this type of problem.

Figure 1:

This figure shows the distribution for each of the four variables used in the study.

Correlations were carried out to examine the relationship between each of the variables used in the study. Table two shows that there was a correlation between Actively Open-Minded Thinking Scale and Temporal Discounting with a correlation of $r_{(145)} = .14$, $p = .043$, this shows that there is a weak but positive correlation between participant's ability to think openly and to delay gratification. Another correlation that was significant was Cognitive Reflection Test and Actively Open-Minded Thinking Scale with a correlation of $r_{(145)} = .20$, $p = .002$, this was also a weak positive correlation, suggesting that there is a relationship between participants ability to think openly and think rationally.

Table 2:

Correlation showing the relationships between each of the variables.

	TD	AOT	CRT	CFQ
TD	-			
AOT	0.14*	-		
CRT	0.07	0.20**	-	
CFQ	-0.01	0.06	0.02	-

CRT - Cognitive Reflection Test. AOT - Actively Open-Minded Thinking Scale. TD - Temporal Discounting. CFQ - Cognitive Failures Questionnaire.

*p < .05. **p < .01.

All the predictors were entered in one block. The purposes of the model were to predict performance on the CRT scores using AOT, TD and CFQ. The regression was a poor fit with an effect size of (R^2 = .073) However, the overall model does predict [$F_{(3,143)}$ =3.76 p=.01] a relationship between the scores on AOT, TD and CFQ and the participant's performance on CRT.

Table three shows the relationship between each of the predators. The model showed that one of the predictors (Actively Open-Minded Thinking Scale) was significant β = .24, [$t_{(3,143)}$=2.91 p=.004] when predicting CRT scores. The model showed that for every one-unit increase in the AOT predictor there would be an increase of .24 in CRT score. The model showed that The Cognitive Failures questionnaire was non-significant β = -60, [$t_{(3,143)}$ = -.70, p = .50] and did not show a relationship between the predictor and CRT. The model also showed that Temporal Discounting predictor was also non-significant β = .09, [$t_{(3,143)}$ = 1.10, p = .27] and did not show relationship

between the predictor and CRT. Therefore, TD and CFQ could not predict the participant performance on CRT.

Table 3:

Predictors beta values, t-test and the significance rate

Variables	Beta values (β)	t	Sig
AOT*	.24	2.91	$p = .004$
Temporal Discounting	.09	1.10	$p = .27$
Cognitive Failures Questionnaire	-.60	-.70	$p = .50$

*Actively Open-Minded Thinking Scale

Table four shows the results from a second multiple regression. However, 44 participants that failed to answer a single question on the CRT were removed. The model showed that one of the predictors (Actively Open-Minded Thinking Scale) was still significant β = .28, $[t_{(3,100)}=2.85 \ p=.005]$ when predicting CRT scores. The model showed that for every one-unit increase in the AOT predictor there would be an increase of .28 in CRT score, however this time there was a greater increase between the two models. The model showed that The Cognitive Failures questionnaire was non-significant β = -.19, $[t_{(3,100)} = -1.96, p = .051]$ and did not show a relationship between the predictor and CRT. The model also showed that Temporal Discounting was also non-significant β = .06, $[t_{(3,100)} = 0.59, p = .55]$ and did not show the relationship between the predictor and CRT. Therefore, TD and CFQ could not predict the participant performance on CRT.

Table 4:

Predictors beta values, t-test and the significance rate

Variables	Beta values (β)	t	Sig
AOT*	.28	2.85	p =.005
Temporal Discounting	.06	0.59	p = .55
Cognitive Failures Questionnaire	-.19	-1.96	p =.051

*Actively Open-Minded Thinking Scale

Discussion

The purpose of this study was to examine whether Actively Open-Minded Thinking Scale, Temporal discounting and Cognitive failures questionnaire could predict performance on CRT. The findings show that the model is significant but two of the predictors do not significantly contribute to the model. Therefore, the hypothesis can be accepted. However, as the results show TD and CFQ did not significantly predict performance on two of the predictors.

The finding in this study showed what Toplak et al., (2014) and Campitelli et al., (2014) had both found in their research. That the data suggests that there is a significant relationship between CRT and AOT, so the AOT was able to predict the participants' ability to think rationally. The results of this study did theoretically replicate what had been found in previous research showing a significant model. Frederick (2005) study showed that participants that score high on the CRT made more logical decisions when it came to TD. However there were certain issues with some questions being non-significant within the study, this was also similar to what Toplak et al., (2014) found as well with TD. Similar issues were also present in this study with this predictor being non-significant. This may be explained looking at previous research, participants who are impatient when answering complex (CRT) questions tend to answer faster and are also more likely to choose the instant gratification options in TD (Hardisty & Weber, 2009). Whereas, participants who score low on the CRT still made more logical decisions choosing to delay financial reward to receive a larger reward later. The differences in response between Instant gratification or delayed gratification dissipates with increased

time (Joshi & Fast, 2013). So the participants are more likely to delay gratification by a few weeks or month if there is an incentive but are less likely over a greater period of time. The final predictor CFQ had very little past research investigating whether CRT could predict performance on CFQ. Based on the result and the fact it was the least significant predictor used, it would suggest that CFQ can not predict performance on CRT.

This research shows the two different types of intuitive and analytical systems that other researchers have looked at in the past. This research shows the difference between participants giving the intuitive (scoring low on CRT) or the analytical response (scoring high). CRT does show these two systems at work but only in a research setting. It has been suggested that making decisions in modern society requires more cognitive ability than what is offered using type one processing (Toplak, West & Stanovich, 2011). However the opposite may also be true, that the vast majority of processing in modern society is type one, as a society is organised to make life simple. Whereas CRT has been designed around problems that are not part of everyday thinking unless individuals do this kind of problem-solving routinely (work or studying). This does not mean that CRT is pointless as it serves a purpose in understanding logic and reasoning ability but in reality may lack mundane realism.

One of the ways to improve this study would be to change the predictors that provided non-significant results or uses more than three predictors to have a greater range of data. The TD predictor used in this study was non-significant, however Frederick (2005) study using the same predictor found this to be significant this may

have been because this study used a shorter version of the test (five questions compared to ten). The sample used was predominantly female, and could have distorted the results. This could be improved by using a more equal gender sample compared to what was used.

Future research could look at examining CFQ, although within this study the data showed to be non-significant using a large and equal gender sample, may improve the usefulness of this predictor. Researchers could also look at intuitive and analytical systems in relation to the CRT. Then compare response time and the answer given to measure the true relationship between system one processing.

Previous research has shown how the CRT can be used as a measure for predicting cognitive ability and decision making. This study was able to replicate what had been found in previous research, although one of the issues was the effectiveness of the predictors. The CRT itself is deterministic in nature meaning that depending on how an individual performs on the CRT they will perform a certain way on other cognitive tests. Whereas extraneous variables one of which maybe intelligence may come into play. Scoring low on the CRT does not necessarily mean an individual will make poor, irrational or illogical decision, and scoring high does not mean an individual will make logical, rational decision. The CRT is a very good tool for predicting performance on other cognitive tasks and when paired with the right predictors it is possible to predict cognitive and reasoning styles.

References

Broadbent, D.E., Cooper, P.F., FitzGerald, P., & Parkes, K.R. (1982). The Cognitive

Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical*

Psychology, *21*, 1-16.

Campitelli, G., & Gerrans, P. (2014). Does the cognitive reflection test measure

cognitive reflection? A mathematical modeling approach. *Memory &*

cognition,*42*(3), 434-447.

Cosmides L, Tooby J. (1996). Are humans good intuitive statisticians after all?

Rethinking some conclusions from the literature on judgment under uncertainty.

Cognition 58:1-73

De Neys, W. (2012). Bias and conflict a case for logical intuitions. *Perspectives on*

Psychological Science, *7*(1), 28-38.

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social

cognition. *Annu. Rev. Psychol.*, *59*, 255-278.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic*

perspectives, 25-42.

Hardisty, D. J., & Weber, E. U. (2009). Discounting future green: money versus the

    environment. *Journal of Experimental Psychology: General*, *138*(3), 329.

Joshi, P. D., & Fast, N. J. (2013). Power and reduced temporal

    discounting.*Psychological science*, 0956797612457950.

Mishra, H., Mishra, A., & Nayakankuppam, D. (2007). Seeing through the heart's eye:

    The interference of system 1 in system 2. *Marketing Science*, *26*(5), 666-678.

Reyna, V. F., & Brainerd, C. J. (2011). Dual processes in decision making and

    developmental neuroscience: A fuzzy-trace model.*Developmental Review*,

    31(Special Issue: Dual-Process Theories of Cognitive Development), 180-206.

    doi:10.1016/j.dr.2011.07.004

Stanovich, K., & West, R. (2000). Individual differences in reasoning: Implications for

    the rationality debate?. *Behavioral And Brain Sciences*, *23*(5), 645-665.

Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive

    ability. *Thinking and Reasoning, 13,* 225-247.

Toplak, M. E., West, R. F., & 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., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflections Test. *Thinking & Reasoning, 20* (2), 147-168.