

Self-employment, management and cognitive reflection revisited: mixed evidence from the UK

Chris Hand and Pawan Tamvada
Department of Strategy Marketing and Innovation
Kingston Business School

c.hand@kingston.ac.uk

Summary

The study revisits and attempts to replicate recent findings on how entrepreneurs and managers make decisions – specifically whether entrepreneurs are more intuitive decision-makers and managers more contemplative. Based on data from the UK Household Longitudinal Study we find support for managers being more contemplative, but not for entrepreneurs being more intuitive. We use a different method of capturing decision-making style to those used previously. Our results suggest that the earlier findings in the literature are dependent on the way decision-making style is measured. In particular, it is possible that the measures used before are capturing numeracy alongside decision-making style.

Track: Entrepreneurship

Self-employment, management and cognitive reflection revisited: mixed evidence from the UK

Introduction

When faced with a decision, we typically have two options: to base it on “gut feeling” and let our intuition guide our judgement or to take a more reasoned approach and analyse the available information before deciding. The first involves the use of heuristics – short cuts that allow for quick decision-making, but ones which might be subject to biases. Entrepreneurs are typically thought of as being more intuitive decision-makers, in particular where information is limited and the business environment is uncertain (e.g. Busenitz and Barney, 1997). Two recent studies directly investigate this association between employment type (self-employed/entrepreneur and hired manager) and decision-making. The first, Koudstaal, Sloof and van Praag (2019), provides evidence that entrepreneurs are more predisposed to intuitive decision-making. The second, Fossen and Neyse (2024) replicates and extends Koudstaal et al’s (2019) findings. This study builds on and extends the earlier findings in two ways. First, we use an alternative measure of cognitive reflection to that used by Fossen and Neyse (2024); second, we extend the findings from the Netherlands (Koudstaal et al, 2019) and Germany (Fossen and Neyse 2024) to the UK. Hence our aim is to test the validity and the generalisability of the earlier findings. Our results are rather less clear cut than those in the previous literature. We are able to replicate the finding that managers are more contemplative decision-makers (compared to non-managers). However, in contrast to Koudstaal et al (2019) and Fossen and Neyse (2024), we are unable to replicate the conclusion that entrepreneurs are less contemplative (and so more intuitive) than managers.

Our paper contributes to this emerging body of scholarship on cognitive reflection of entrepreneurs. Firstly, we replicate and test the existing models using an alternative measure of cognitive reflection (one designed to address some of the limitations of the original measure). Secondly, by examining a new context—the UK—we check if the results of extant studies are robust to alternative geographic contexts. More generally, we follow Fossen and Neyse (2024) in highlighting the benefits of establishing the stability of empirical findings through independent and extended replications. An extended replication is one which goes beyond a direct replication by testing a result using a new method or a new sample.

Intuition, cognition and entrepreneurship

Entrepreneurs operate in an environment where decisions are made without full information, especially as such decisions are innovative by nature and so there will be limited data or experience to fall back on (e.g. Baldacchino et al., 2015). Such decisions will tend to involve intuitive decision-making. Decision-making and thinking styles can be divided into a more intuitive and heuristic-based approach (labelled system I) and slower, contemplative but more cognitively demanding approach (system II). These are typified in the title of Daniel Kahneman’s (2011) book as “Thinking, fast and slow” and are more formally known as the dual-process approach. The switch from system I to system II has been labelled cognitive reflection and requires an initial intuitive answer to be held back allowing for a more considered response. Such system II thinking imposes a heavier cognitive burden and a delay in reaching a decision. However, decisions based on intuition, especially judgements made on incomplete information, are subject to biases. Interest in the biases arising from heuristic-based decisions arguably dates back to Tversky and Kahneman (1974) who described the anchoring, availability and representativeness biases. Anchoring refers to when a (potentially

irrelevant) piece of information is used as a reference point for a subsequent decision. Availability bias is where a decision is based on immediately available examples, such as the first example to come to mind. Representativeness bias can arise where we rely on a stereotype to inform a decision. Entrepreneurship research has explored range of biases affecting entrepreneurs' decisions, as reviewed by for example Zhang and Cueto (2017), including those arising from the effects of goals or desires (such as overconfidence or over-optimism); those arising from erroneously focusing on one attribute when others are more relevant (such as illusion of control and the availability and representativeness biases mentioned above). Somewhat less attention has been given to the dual-process system in entrepreneurship research (Fossen and Neyse, 2024). Koudstaal et al (2019) provide evidence from a large-scale survey study of a relationship between decision-making style and entrepreneurship. Their findings have since been replicated by Fossen and Neyse (2024).

Koudstaal et al (2019) aim to test if entrepreneurs differ from managers and employees in their decision-making style. Decision-making style was captured using a measure based on Rubinstein's (2016) Contemplative Index (CI), along with self-report measures of faith in intuition and need for cognition. The original CI measure used 10 questions, referred to as strategic games, where the respondent has to make a decision based in part on how they expect a competitor to react) each of which has intuitive and contemplative answers. The CI has the advantage of being phrased in terms of strategic decisions, which arguably makes it more appropriate for studies with entrepreneurs and managers than other more general measures of decision-making style. Koudstaal et al (2019) used a reduced form of the CI, selecting three items to reduce the length of the survey administered to a sample of entrepreneurs, managers and employees. They find that managers score significantly higher than employees on CI, while entrepreneurs show no significant difference. They also find that the regression coefficients for entrepreneurs and managers are significantly different, with the entrepreneur coefficient being significantly smaller. Similarly, they find entrepreneurs score significantly higher on the faith in intuition scale than managers. The picture for need for cognition is less clear cut; when controls are introduced to the regression models, the entrepreneur variable stops being significant and managers remains significant only at the 10% level. A potential limitation of Koudstaal's et al.'s (2019) approach is the comparatively low response rate for their survey, and the differences in response rates between employment types

Fossen and Neyse (2024) present results of an extended replication of Koudstaal et al's (2019) findings which confirm the original result. They extend the study in a number of ways: by employing an alternative measure of decision-making style introduced by Frederick (2005): the Cognitive Reflection Test (CRT); controlling for effects of personality dimension and testing for effects of overconfidence. In choosing the CRT in preference to CI, Fossen and Neyse (2024) opted for a measure which was discounted by Koudstaal et al (2019). The CRT is a bank of three questions which have an intuitive but wrong answer and a correct answer that requires the initial intuitive answer to be suppressed. However, some concerns have been raised about how effective the CRT questions are now they are so widely used (e.g. Toplak et al, 2014) and how far it captures other characteristics alongside cognitive style (e.g. Thomson and Oppenheimer, 2016). The CRT questions take the form of trick questions, which lead people an intuitive but wrong answer. Familiarity with the questions and with the right answers could potentially undermine the validity of the measure. Fossen and Neyse (2024) address this by filtering out respondents who were familiar with the CRT items. Their data was obtained from the 2020 wave of a German household panel survey, SOEP-IS (Sozio-oekonomischen Panels Innovationsample or in English, the Socio Economic Panel Innovation Sample). SOEP-IS is a panel used for short-term experiments alongside the larger, longer running, SOEP survey. The CRT questions were administered to a subset of the

SOEP-IS sample, yielding 1961 responses. They replicate Koudstaal et al.'s (2019) findings, with managers scoring significantly higher on the CRT than non-managers. They also find a significant difference between entrepreneurs and managers, again implying that managers are more contemplative whilst entrepreneurs tend towards intuitive decision-making. This relationship holds after they control for dimensions of personality. Furthermore, they find evidence that overconfidence also plays a role in explaining decision-making style. When a measure of overconfidence is included in their regressions the impact of employment types is reduced. However, as they note, overconfidence is related to CRT by definition.

To sum up, two recent empirical studies have found evidence for a relationship between decision-making style and type of employment (entrepreneurs v managers). This provides supporting evidence for the view that entrepreneurs' decision-making tends to be more intuitive whilst managers tend to be more contemplative. Earlier literature (e.g. Busenitz and Barney, 1997) suggests that this arises in part from the differences in the environment entrepreneurs and managers operate in. However, as Koudstaal et al (2019) note entrepreneurs are more likely to use some decision-making heuristics (e.g. representativeness heuristic) than managers, but are less likely to use other (e.g. the status quo heuristic).

Research Aim

Our aim in this study is to provide a further replication of results reported by Koudstaal et al (2019) and Fossen and Neyse (2024). Replication of findings is important for theory building and theory testing (e.g. Crawford et al. 2022). In particular, testing the boundary conditions of a finding through, for example, re-examining it in a different context or using a different measure adds weight to the previous results. The entrepreneurship field has seen a number of rallying calls for replication research, including a special issue of the *Journal of Business Venturing Insights* (van Witteloostuijn et al. 2021) and Crawford et al's (2022) call in *Entrepreneurship Theory and Practice* coming against a backdrop of a replication crisis in social sciences (e.g. Freese and Peterson, 2017) and a lack of replications in Business and Management research (Ryan and Tipu, 2022). Ryan and Tipu (2022) also highlight that the majority of replications in Business and Management are in-study replications (i.e. studies replicated within a publication by the same research team) and hence there is a lack of independence between original research and replications. If the original results have already been replicated, one might ask why a further replication would be useful. Thus far we have two studies which draw the same conclusion using two different ways of capturing decision-making and two different national samples: one Dutch and one German. Certainly, a successful replication adds additional weight to the conclusions. A further replication in a new national context would provide even further evidence. Secondly, the cognitive style measure used by Fossen and Neyse (2024) has limitations. In particular, it may to an extent conflate thinking style with mathematical (in particular arithmetical) ability. Our aim in this study is to investigate whether previous findings hold when an alternative measure of approach to decision making is used with a sample from a different country.

Our hypotheses, based on the discussion above are:

H1: Managers score significantly higher than non-managers on a cognitive reflection measure.

In H1 we expect to find that managers are associated more strongly with a contemplative approach to decision-making, measured through a cognitive reflection measure.

H2: Entrepreneurs do not score significantly higher on cognitive reflection than non-managers.

Perhaps rather unusually, we expect and hypothesise a non-significant result in H2. In doing so we are expecting to replicate the non-significant result reported in earlier findings.

H3: Entrepreneurs score significantly lower than managers on cognitive reflection.

Following the previous literature, we expect entrepreneurs to score significantly lower than managers on cognitive reflection. Doing so we adopt an indirect way of assessing if entrepreneurs are more associated with intuitive rather than contemplative decision-making.

Measures and Data

To investigate our research questions we use data from Wave 16 of the UK Household Longitudinal Survey's Innovation Panel (UKHLS IP). The UKHLS is a major longitudinal household panel survey which started in 2009. The Innovation Panel is drawn from a separate sample from the main UKHLS survey consisting of approximately 1,500 households from across the UK (except Northern Ireland). Both the UKHLS and the UKHLS IP are managed by the Institute for Economic and Social Research at the University of Essex with the data made available via the UK Data Service (University of Essex, Institute for Social and Economic Research, 2024). Wave 16 of the Innovation Panel data was released in November 2024 and included cognitive reflection questions as part of a study on political decision-making.

The dependent variable, following Koudstaal et al (2019) and Fossen and Neyse (2024) is a measure of cognitive reflection. Rather than use the original version of CRT (as Fossen and Neyse, 2024 use) we use an alternative form, CRT2 (Thomson and Oppenheimer, 2016). As our study is based on secondary data this was not our choice of measure. However, CRT2 has a number of advantages over the original. The CRT items have been widely used in research on decision-making and appeared in media discussion (including Kahneman's, 2011, book *Thinking, Fast and Slow*). CRT2 was developed to provide an alternative measure using new items to reduce the risk of research samples being over-exposed to the CRT items. A criticism of the original CRT is that correctly answering the questions requires mathematical ability (as the questions involve calculations). A number of studies have identified strong correlations between CRT answers and measures of numeracy (e.g. Campitelli and Gerrans, 2014; Liberali et. al., 2012). This has led to development of verbal tests of cognitive reflection (e.g. Sirota et al., 2021). CRT2 employs a similar set of questions to CRT which require less mental arithmetic from the respondent. As a result, CRT2 scores are less strongly correlated with numeracy (Thomson and Oppenheimer, 2016). The CRT and CRT2 measures are shown in Table 1 below.

Table 1. Comparison of CRT and CRT2 items

CRT items (Frederick, 2005)	CRT2 (Thomson and Oppenheimer, 2016)
A bat and ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost? (intuitive answer \$0.1 or 10 cents; correct answer: \$0.05)	If you're running a race and you pass the person in second place, what place are you in? (intuitive answer: first; correct answer second)
If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? (intuitive	A farmer had 15 sheep and all but 8 died. How many are left? (intuitive answer: 7; correct answer 8)

answer: 100 minutes; correct answer: 5 minutes)	
In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? (Intuitive answer: 24 days; correct answer 47 days)	Emily's mother has three daughters. The first two are named April and May. What is the third daughter called? (intuitive answer: June; correct answer: Emily)
	How many cubic feet of dirt are there in a hole that is 3' deep x 3' wide x 3' long? (intuitive answer: 27, correct answer: none)

A correct answer is scored as 1 and number of correct answers is summed, giving each respondents a score of between 0 (none correct) and 4 (all correct). A higher CRT/CRT2 score is thus associated with a more reflective / contemplative thinking style.

The full sample contains 2847 respondents. In the study we only use responses from those who are self-employed or in paid employment, based on the job status variable included in the Innovation Panel data (variable name: *jbstat*). Following prior studies, we use self-employed as a proxy for being an entrepreneur. A second variable identifies if a respondent in paid employment is a manager, a foreman or supervisor or lastly not a manager or supervisor (variable name: *jbmgr*). This gives 1,326 respondents in either paid employment or self-employment, with 174 self-employed, 305 managers, 116 foreman / supervisors and 731 non-managers. From this, three dummy variables were created to indicate self-employed (as a proxy for entrepreneur), manager and lastly foreman / supervisor, with non-manager used as the base category.

Similar control variables to previous studies are used. Respondent's age in years, whether the respondent is female and highest level of education qualification obtained: A levels, Degree or Higher Degree with other qualification below these as the base category. All of these have been found to be related to decision-making style. Koudstaal et al. (2019) argue that education is likely to be an important influence on decision-making style and so introduce controls for it. A similar argument could be made for age, acting as a proxy for experience. A finding which has held across multiple studies is that males tend to score higher on the original CRT items (Frederick, 2009; Branas-Garza et al, 2019). CRT2 shows less of a difference between male and female response patterns (Thomson and Oppenheimer, 2016). However, we retain respondent's sex as a control variable.

Findings

As a preliminary step, the responses to the CRT2 questions are reviewed. Responses to two of the CRT2 items (item 1 Race and item 3 Daughters) were provided pre-coded into correct, intuitive and wrong in the data set. Numerical responses to items 2 (sheep) and Item 4 (hole) were coded into the same categories. Item 4 showed a very different distribution of responses with half of the respondents giving the incorrect intuitive answer and a further 30% giving a wrong answer. The patterns of responses is shown in table 2.

Table 2. Distribution of responses to CRT2 items

	Race	Sheep	Daughters	Hole
Correct	1088	1778	1564	258
Intuitive	1345	951	1067	1410
Wrong	380	49	182	857

Refusal	-	22	-	50
Don't know	-	13	-	238
Total	2813	2813	2813	2813

A crosstabulation of answers against respondent sex showed a different pattern of answers for item 4, with male respondents likely to give either the correct answer or an intuitive answer and female respondents likely to either give a wrong answer or answer 'don't know'.

Because of this, we drop responses to item 4 and base our measure on response to items 1 to 3, giving us a variable with four possible values, from 0 (none correct) to 3 (all correct).

Next, we explore differences in CRT2 score across employment types. We would expect managers to score higher on CRT2, suggesting a more contemplative decision-making process and entrepreneurs to score lower, implying a more intuitive approach. Mean CRT2 scores for each employment type are shown in table 3 below.

Table 3. CRT2 means by employment type

	Mean	N	Standard Deviation
Self-employed	1.62	174	1.04
Manager	1.88	305	1.05
Foreman or supervisor	1.76	116	1.03
Non-manager	1.71	731	1.06

Table 3 shows a similar, but less pronounced pattern than that reported by Fossen and Neyse (2024) with slightly higher mean for managers implying a more contemplative decision-making style. An ANOVA test ($F(2, 1325) = 2.806, p = 0.039$) suggests that there are significant (but small) differences in the means. A striking difference between the results in table 3 and Fossen and Neyse's (2024) is that self-employed and non-managers in the UK data score almost twice as high than in the German data. Some of that difference may be accounted for by the different versions of CRT used. Given that the German SOEP and UKHLS surveys adopt a similar structure, it seems less likely that differences in sample composition explain the result.

To test our hypotheses we run a series of regressions. The descriptive statistics for the regression data are shown in table 4 below.

Table 4 Descriptive Statistics

	%
Education (base category = other or none)	
A levels	20.1
Degree	41.9
Other higher degree	12.2
Female	51.7
Employment type (base category = non-manager)	
Self-employed	13.1
Manager	23.0
Foreman or supervisor	8.7

CRT number correct	
0	16.8
1	21.2
2	32.6
3	29.4
	Mean (Std Dev)
Age	45 (13.0)
N	1326

Our data comes from a household panel, so the data may contain multiple responses from the same household. To accommodate any potential household effects, we first report regression results with clustered standard errors. These allow the assumption that residuals are independent to be relaxed and instead to assume that the residuals are independent between clusters (but not within clusters). We report OLS estimates for consistency with Fossen Neyse (2024). However, as our dependent variable is limited to one of four values (0, 1, 2 or 3), a count data model would arguably be more appropriate than OLS regression. Alongside the OLS results we report results from a count data regression model: negative binomial regression. As a further robustness check we also employ a multilevel regression model. Multilevel regressions are designed to accommodate data containing hierarchies or clusters of responses; in our case households (for an accessible overview see e.g. Robson and Pevalin, 2016). OLS and Negative Binomial regression results using clustered standard errors are shown in tables 5 and 6 below.

Table 5. OLS regression results

	Coef.	Std. Err.	t	Sig.
Self-employed	-0.007	0.094	-0.080	0.939
Manager	0.117	0.073	1.620	0.106
Foreman / Supervisor	0.073	0.098	0.750	0.455
Age	-0.017	0.002	-7.600	0.000
a_levels	0.203	0.088	2.320	0.020
Degree	0.322	0.078	4.100	0.000
other higher degree	0.239	0.106	2.260	0.024
Female	-0.211	0.054	-3.910	0.000
Constant	2.394	0.129	18.610	0.000
R ²	0.079			
F(8, 930)	13.77			
Sig	0.001			

Notes: Dependent variable is a measure of Cognitive Reflection (CRT2). Base category of occupation variables is 'Non-Manager'.

Table 6. Negative Binomial results

	Coef.	Std. Err.	Z	Sig.
Self-employed	0.001	0.057	0.001	0.999
Manager	0.071	0.040	1.790	0.073
Foreman / Supervisor	0.049	0.055	0.880	0.379
age	-0.010	0.001	-7.440	0.000
A levels	0.129	0.054	2.380	0.017
degree	0.198	0.049	4.000	0.000
other higher degree	0.152	0.065	2.350	0.019
female	-0.120	0.031	-3.910	0.000
constant	0.906	0.073	12.410	0.000
Pseudo R ²	0.0168			
Wald chi2 (8)	96.17			
sig	0.001			

Notes: Dependent variable is a measure of Cognitive Reflection (CRT2). Base category of occupation variables is 'Non-Manager'.

In the OLS results (table 5), the coefficient for Managers has a significance value of 0.106, just above 10%. The pattern of coefficients is as expected though, with the coefficient for managers being greater than those for entrepreneurs and supervisors. The regression coefficients for entrepreneurs and supervisors are not significant, indicating no significant difference between them and the base category, non-managerial employees. All of the control variables have statistically significant coefficients.

The negative binomial regression results (table 6) show a very similar pattern, this time with the Manager coefficient being significant at the 10% level ($p = 0.071$) whilst the coefficients for self-employed and supervisors remaining non-significant. As with the OLS results, the control variable coefficients are all significant. That coefficients are smaller in this set of results does not imply that the relationship is weaker. In the OLS results the coefficient represents the effect of a unit change in the independent variable. If the coefficient is say 0.3 it shows that an increase of one in the independent variable is associated with an increase of 0.3 in the dependent variable. In Negative Binomial Regression, again with a coefficient of 0.3, a change of one unit in the independent variable is associated with an increase of 0.3 in the natural log of the dependent variable. Hence the results in tables 5 and 6 provide limited support for our hypotheses. Managers have higher cognitive reflection than non-managers (H1) but in neither regression do the results reach the standard level of significance. Entrepreneurs are not significantly different when compared to non-managers in their cognitive reflection, as reflected in H2. To test H3 we test if the regression coefficients for managers and for entrepreneurs are significantly different. For the OLS results, an F test fails to reject the null hypothesis that the manager and entrepreneur are equal: ($F(1, 930) = 1.47$, $\text{sig} = 0.23$) providing no support for H3. A similar test for the

Negative Binomial results also produces a non-significant result: $\text{Chi}^2(1) = 1.41$, $\text{sig} = 0.23$, again offering no support for H3.

That the significance value for the manager coefficient changes between regression models suggests that the household composition of our data has an effect. It may be that the clustered standard errors do not fully account for the household structure of our data. As a further check, a multilevel model OLS regression model is estimated with the results given in table 6 below (we attempted to run a multilevel negative binomial regression but the model failed to converge) .

Table 7. Multilevel regression model results

	Coef.	Std. Err.	z	Sig	[95% Conf.	Interval]
Self-employed	0.014	0.090	0.150	0.880	-0.163	0.190
Manager	0.145	0.071	2.050	0.040	0.007	0.284
Foreman / Supervisor	0.073	0.100	0.730	0.465	-0.123	0.270
age	-0.017	0.002	-7.620	0.000	-0.022	-0.013
A levels	0.199	0.086	2.310	0.021	0.030	0.367
degree	0.293	0.077	3.810	0.000	0.142	0.443
Other higher degree	0.242	0.097	2.490	0.013	0.051	0.434
Female	-0.207	0.053	-3.900	0.000	-0.310	-0.103
constant	2.388	0.129	18.520	0.000	2.135	2.641
Random-effects Parameters	Estimate	Std. Err.				
Household: Identity						
Var (Constant)	0.32232	0.05771				
Var (Residual)	0.704566	0.053736				
Wald Chi ² (8)	98.96					
Sig	0.001					
No. of obs	1269					
No of groups	931					
Obs per group						
Min	1					
Avg	1.4					
Max	4					

As table 6 shows, the coefficient for managers is significant at the 5% level this time ($\text{sig} = 0.040$), but the coefficients for foreman/supervisor and entrepreneur are not significant, in line with our expectations in H1 and H2. Again, all the control variable coefficients are all statistically significant.

The pattern of results across all three tables reflects the results presented by Fossen and Neyse (2024), with the results for entrepreneurs being much smaller than managers. However, in our results the coefficient for managers is not significantly different to that for entrepreneurs ($\text{Chi}^2(1) 1.97$, $\text{Sig} = 0.16$). Hence, we can conclude that whilst managers are more contemplative thinkers than non-managerial employees, but we cannot say that entrepreneurs are significantly less contemplative and more intuitive. This seems to arise

from a smaller coefficient for managers in our models than in Fossen and Neyse's. Whether this is a quirk of the CRT2 measure or of the sample is an open question, but one we consider further in the Discussion and Conclusion section below.

Discussion and Conclusion

The entrepreneurship literature suggests that entrepreneurs, proxied in this and previous studies by self-employed, to be more intuitive than contemplative decision-makers, and less contemplative than managers. The nature of the environment entrepreneurs operate in may even preclude or contemplative approaches. Decisions are made with incomplete and / or ambiguous information, with little historical data to draw on and under pressure of time. Two studies Koudstaal et al, (2019) and Fossen and Neyse (2024) have set out to test this empirically, following on from a range of studies exploring biases in entrepreneurial decision-making. Both studies found evidence that managers were more contemplative and indirect evidence that entrepreneurs were more intuitive (in that a higher CRT score suggests a contemplative decision-making style whilst a low CRT score could arise either from an intuitive style or from giving an incorrect answer). In our study we find additional evidence for the first of these findings but not the second.

It is perhaps somewhat counterintuitive that both studies whose results we seek to replicate were designed to test and report results in terms of contemplativeness of decision-making, particularly by managers, when the central question is are entrepreneurs more intuitive decision-makers. As we aimed to re-test this conclusion in an extended replication study, we have adopted the same approach. Future research could modify the approach in one of two ways. First using managers or entrepreneurs as the base category rather than non-managers would allow for a more direct test of differences between entrepreneurs and managers (and remove the need for post-regression tests for equality of coefficients). Secondly, if using a form of CRT, using an alternative way of coding the CRTS results could be useful. The conventional method is to count the number of correct responses given as a measure of contemplative decision-making. A lower score could arise from either intuitive thinking or from giving another wrong answer. Cueva et al (2016) propose and test measures of cognitive reflection and 'irreflection' which employ alternative ways of coding the CRT responses. This would allow for a direct test of association between being an entrepreneur and being either a contemplative or an intuitive decision-maker.

That we achieve different results using a different measure of decision-making style raises an intriguing possibility of whether the Fossen and Neyse (2024) results are due in part to other characteristics being captured by the CRT. As noted above, CRT has been found to be related to numerical ability (e.g. Morsanyi et al. 2014) whilst the evidence regarding its association with impulsiveness is more mixed (Littrell, et. al, 2020). As table 1 shows, the CRT items involve an element of calculation; the CRT2 items involve less calculation and more deduction. Thomson and Oppenheimer (2016) report that correlations between CRT2 and measures of numeracy are significantly smaller than for the original CRT. Hence the previously reported results could in part reflect an underlying comfort with using and interpreting numerical information in addition to decision-making style. We suggest that a fruitful avenue for future research could be to re-test the results presented above using non-numeric measures of CRT (such as that proposed by Sirota et al., 2021) or including measures of numeracy as additional control variables. There are two possible alternatives which we are yet to explore using the CRT2 measure. The first is omit the non-numeric 'daughters' question from CRT2. This would give a similar set of number-based questions to those in CRT and test if results more similar to Fossen and Neyse (2024) are obtained. The second is to use the CRT2 items individually and see if the numerical items show a different relationship to employment type than the non-numeric items. Doing so could introduce a

greater element of measurement error (as it would in effect entail replacing a multi-item scale with a single item scale), but would allow for differences between the numeric and non-numeric items to be identified.

We should note that, like all studies, our study has some limitations. We rely on a secondary data set and so have a different measure of cognitive reflection applied to a different national sample. Consequently, we cannot say for certain which of these two changes is responsible for the different results we obtain. We suspect that it is due to the way in which CRT and CRT2 capture cognitive reflection. Secondly, although our sample is of a similar size to previous studies, we note that it contains a relatively small number of entrepreneurs. This is in part due to the nature of the sample employed – a panel of the UK population rather than a survey of people in paid or self-employment. Lastly, like Fossen and Neyse (2024), our source data does not include measures of need for cognition or of faith in intuition, so we are unable to replicate those findings in Koudstaal et al.'s (2019) study.

References

- Baldacchino, L., Ucbasaran, D., Cabantous, L. and Lockett, A. (2015) "Entrepreneurship research on intuition: A critical analysis and research agenda", *International Journal of Management Reviews*, 17(2), 212-231.
- Busenitz, L.W. and Barney, J.B. (2007) "Differences between entrepreneurs and managers in large organisations: biases and heuristics in strategic decision-making", *Journal of Business Venturing*, 22(3), 340-362.
- Campitelli, G. and Gerrans, P. (2014) "Does the cognitive reflection test measure cognitive reflection", *Memory and Cognition*, 42(3), 434-447.
- Cueva, C. Iturbe-Ormaetxe, I., Mata-Perez, E., Ponti, G., Sartarelli, M. Yu, H. and hukova, V. (2016) "Cognitive ir(reflection): New experimental evidence", *Journal of Behavioral and Experimental Economics*, 64, 81-93.
- Crawford, G. C., Skorodziyevskiy, V., Frid, C. J., Nelson, T. E., Booyavi, Z., Hechavarria, D. M., Li, X., Reynolds, P. D., & Teymourian, E. (2022) "Advancing Entrepreneurship Theory Through Replication: A Case Study on Contemporary Methodological Challenges, Future Best Practices, and an Entreaty for Communalit", *Entrepreneurship Theory and Practice*, 46(3), 779-799.
- Frederick, S. (2005) "Cognitive Reflection and Decision Making", *Journal of Economic Perspectives*, 19(4), 25-42.
- Freese, J. and Peterson, D. (2017) "Replication in Social Science", *Annual Review of Sociology*, 43, 147-165.
- Fossen, F.M. and Neyse, L. (2024) "Entrepreneurship, Management, and Cognitive Reflection: A preregistered replication study with extensions", *Entrepreneurship Theory and Practice*, 48(4), 1082-1109.
- Kahneman, D. (2011) *Thinking, fast and slow*, London: Penguin Books.

Koudstaal, M., Sloof, R. and van Praag, M. (2019) “Entrepreneurs: intuitive or contemplative decision-makers?”, *Small Business Economics*, 53, 901-920.

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 judgement”, *Journal of Behavioral Decision-Making*, 25(4), 361-381.

Litrell, S., Fugelsang, J. and Risko, E.F. (2020) “Not so fast: Individual differences in impulsiveness are only a modest predictor of cognitive reflection”, *Personality and Individual Differences*, 154, 109678.

Robson, K. and Pevalin, D. (2016) *Multilevel modelling in plain language*, London: Sage.

Rubinstein, A. (2016) “A typology of players: Between instinctive and contemplative”, *Quarterly Journal of Economics*, 131(2), 859-890.

Ryan, J.C. and Tipu, S.A.A. (2022) “Business and management research: Low instances of replication studies and a lack of author independence in replications”, *Research Policy*, 51(1), 104408.

Sirota, M., Dewberry, C., Juancich, M., Valus, L. and Marshall, A.C. (2021) “Measuring cognitive reflection without maths: Development and validation of the verbal cognitive reflection test”, *Journal of Behavioral Decision Making*, 34(3), 322-343.

Thomson, K.S. and Oppenheimer, D.M. (2016) “Investigating an alternate form of the cognitive reflection test”, *Judgement and Decision Making*, 11(10), 99-113

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. (1974) “Judgement under Uncertainty: Heuristics and biases”, *Science*, 185, 1124-1131.

University of Essex, Institute for Social and Economic Research. (2024) *Understanding Society: Innovation Panel, Waves 1-16, 2008-2023*. [data collection]. 13th Edition. UK Data Service. SN: 6849, DOI: <http://doi.org/10.5255/UKDA-SN-6849-16>

van Witteloostuijn, A., Dejardin, M., Pollack, J.P. (2021) “Replication in entrepreneurship”, *Journal of Business Venturing Insights*, 16, e00271.

Zhang, S. X. and Cueto, J. (2017) “The study of bias in entrepreneurship”, *Entrepreneurship Theory and Practice*, 41(3), 419-454.