Contributed Paper prepared for presentation at the 96th Annual Conference of the
, Agricultural Economics Society, K U Leuven, Belgium
4 – 6 April 2022

Investigating inconsistencies in complex lotteries: The role of cognitive skills of low-numeracy subjects

Selina Bruns*† Daniel Hermann* Oliver Mußhoff*

Disclaimer: This manuscript has been recently published in the *Journal of Behavioral and Experimental Economics https://doi.org/10.1016/j.socec.2022.101840*

Abstract

Comprehension in risk elicitation tasks is crucial, as otherwise the results are rather noisy than reliable. One prominent risk-elicitation tool, the Holt and Laury task (HL-task), is particularly prone to a noisy outcome - indicated by high inconsistency levels - when used among low-literacy subjects. Yet, it is unclear what drives inconsistencies. In this note we investigate the HL-task inconsistency levels of 247 smallholder farmers from rural Cambodia. Cognitive skills, measured through Raven's Progressive Matrices (RPM), are a statistically significant determinant of inconsistency levels. A second step in the analysis reveals that cognitive skills are a statistically significant explanation for inconsistency levels for men, but not for women. Our results suggest that researchers should conduct a comprehensive pre-test when aiming at using abstract risk-elicitation methods among low-numeracy subjects in the field.

Keywords: risk measurement, risk attitude, Raven's Progressive Matrices, binary lottery, South-East Asia

*Department of Agricultural Economics and Rural Development, University of Goettingen
†Email: selina.bruns@uni-goettingen.de.

Investigating inconsistencies in complex lotteries: The role of cognitive skills of low-numeracy subjects

Abstract

Comprehension in risk elicitation tasks is crucial, as otherwise the results run the risk of being rather noisy than reliable. One prominent risk-elicitation tool, the Holt and Laury task (HL-task), is particularly prone to a noisy outcome - indicated by high inconsistency levels - when used among low-literacy subjects. Yet, it is unclear what drives inconsistencies. In this note we investigate the HL-task inconsistency levels of 247 smallholder farmers from rural Cambodia. Cognitive skills, measured through Raven's Progressive Matrices (RPM), are a statistically significant determinant of inconsistency levels. A second step in the analysis reveals that cognitive skills are a statistically significant explanation for inconsistency levels for men, but not for women. Our results suggest that researchers should conduct a comprehensive pre-test when aiming at using abstract risk-elicitation methods among low-numeracy subjects in the field.

Keywords: risk measurement, risk attitude, Raven's Progressive Matrices, binary lottery, South-East Asia

1 Introduction

According to the most recent statistics, around 767 million people worldwide live in poverty. The vast majority of these global poor reside in rural areas, depending to some degree on agriculture as their source of livelihood - in particular as smallholder farmers (World Bank, 2018). These smallholders operate under particularly risky conditions, e.g. due to the unstable environmental, market and household conditions (World Bank, 2016, 2020). Furthermore, the risk involved in economic decision making is extraordinary for poor small farm managers, as investments with respect to e.g. crop cultivation and input use can be pivotal for the economic survival of the entire household. However, to interrupt the downward spiral of poverty, individuals need to take the risk of investing into human, physical or social capital, which can be a difficult endeavor without any safety net. Individuals that are risk averse to the extent that they refuse to invest into new technologies will be less able to cope with shocks, and hence, might end up in chronic poverty (Mosley & Verschoor, 2005). Thus, risk attitude analysis is a crucial step in contributing to combating poverty, as supporting custom-fit policy interventions can be derived.

Critical to this analysis is having a reliable instrument to elicit individual risk attitudes. In recent years, various experimental methods such as playing lotteries, investment games, or rather intuitive games, have been tested in the lab or the field to understand how individuals decide under uncertainty (Charness et al., 2013). Nevertheless, the methods being used to elicit risk attitudes often involve complex and abstract tasks. For example, praised as the *gold standard*, the HL-task (a very precise, yet complex binary lottery (Charness et al., 2013)) is often used by researchers, even when measuring risk attitudes among low-numeracy subjects (Verschoor et al., 2016). However, the HL-task incorporates an indicator for irrational choice-patterns: the inconsistency rate. While the inconsistency rate among high-literacy subjects is between 10-15 percent (Charness & Viceisza, 2016), the rate tends to be crucially higher (up to 75 percent) among low-numeracy subjects from low-income countries (Charness & Viceisza, 2011; Jacobson & Petrie, 2009; Galarza, 2009). Due to this divergence in inconsistency rates, the literature suggests that while the HL-task is suitable for highly educated individuals (Vollmer et al., 2017), such a complex task might not be appropriate for those less educated as the data can be flawed (Brick et al., 2012).

However, if complex, yet finer, tasks yield flawed data from low-numeracy subjects and simpler, yet courser, methods might be more applicable but less precise - then which method should be used for low-numeracy subjects? Many researchers have identified and approached this question by either developing easier tasks (Lejuez et al., 2002; Crosetto & Filippin, 2013) or by adjusting the design of complex tasks to decrease inconsistencies. Examples for this strategy include employing better illustration of complex tasks (Ihli et al., 2018) or coining a task in a domain specific manner by framing a complex lottery in a real life and context specific scenario (Rommel et al., 2019) or by relying on contextual framing means such as bills to illustrate rewards (Estepa-Mohedano & Espinosa, 2021). However, the preceding yet open question is: Why do rural poor participants not understand complex tasks? Dave et al. (2010) investigate why inconsistencies are so high and conduct a simple math test before running the actual experiment. They find a statistically significant relationship between math results and level of inconsistency, with lower math results translating into higher inconsistency levels in the HL-task. He et al. (2018) also approach this question, hypothesizing that it is the level of cognitive skills that drives inconsistencies within different risk-elicitation methods. Taking education as a proxy for cognitive skills, they find no relationship and suggest that inconsistency levels rise due to measuring ambiguity rather than risk attitude and put forward that they "(...) leave it for future research to formally test for cognitive skills as a potential underlying reason of the inconsistency (...)" (He et al. (2018), p.1968). While this relationship between cognitive skills and errors within complex lotteries has been investigated in western societies (see for example Amador-Hidalgo et al. (2021); Andersson et al. (2016)), to the best of our knowledge there is no evidence stemming from low-numaracy subjects, especially not rural farmers.

This is the starting point of our study: We hypothesize that cognitive skills - measured through the RPM test - are a negative and statistically significant determinant of inconsistency levels in the HL-task for low-numeracy subject pools[1]. To investigate our hypothesis, we undertake a regression analysis and include the results from the HL-task and the RPM-test. Other characteristics of the individual had to be controlled for in the regression model

[1]Note that we do not aim to investigate a relationship between analytical or cognitive skills and risk attitude. Nor do we aim to find determinants for economic behavior besides risk taking. We solely focus on the noise within the risk elicitation method and its relationship with cognitive skills. For an investigation in the former, see for example Brañas-Garza et al. (2008) or Branas-Garza & Rustichini (2011).

to exclude them as confounding influences in the analysis of the effect of cognitive skills on inconsistencies. While doing so, the initially unwanted variation in sociodemographic characteristics also facilitated an exploratory search of associations. The results from our initial regression led to a second hypothesis, namely that the statistical significance of the effect of cognitive skills - measured through the RPM test - on inconsistency levels in the HL-task differs by gender.

We test both hypotheses with 247 smallholder farmers from rural Cambodia. Clustered as a least developed country (United Nations, 2018), Cambodia is among the poorest countries worldwide and thus presents itself as a highly fitting study region for our investigation. Our contribution to literature is straight forward: we test if cognitive capacities are a statistically significant determinant of inconsistencies in the HL-task. We further go beyond the mean effect of cognitive skills on inconsistency levels and test if gender drives the result. The results can aid researchers in identifying ex-ante weather as a complex lottery can be a suitable instrument for the given study group.

2 Method

2.1 Risk elicitation: Holt and Laury Task

Being praised as the 'gold standard' to elicit risk preferences throughout literature (Anderson & Mellor, 2009; Charness & Viceisza, 2016; Charness et al., 2018), the binary lottery developed by Charles A. Holt and Susan K. Laury in 2002 seems to be the first choice instrument for researchers interested in understanding risk attitudes Charness et al. (2018). For measurement, subjects are presented with paired lottery-choices among which they need to choose upon over ten rounds. As seen in Table 1 the payoffs in option A and option B stay unchanged throughout the ten rounds. As seen in Table 1 the payoffs of option A range in a narrow margin (2 USD-1.60 USD) while those of option B are more variable (3.85 USD-0.10 USD). What changes during the lottery are the probabilities with which the payoffs occur. Thus, within the ten rounds the player is expected to switch from option A to option B. The switching point is crucial as it determines the risk attitude of the subject. A risk neutral individual is expected to switch in round five, as the difference in expected payoffs from option A to option B turns negative (-0.18 USD). Even an extreme risk averse person would show a switching point, namely in round ten, as the probability for receiving the high payoff of 3.85 USD is equal to 100 percent. A risk seeking person would switch even before round five, with the most extreme starting off with option B in the first place and thus being the only exception of not switching at all (Holt & Laury, 2002). An individual is clustered to have an inconsistent choice pattern if they deviate from the utility maximizing strategy. While this is captured in different ways, a starting point for inconsistent behavior is if they i) switched more than once between the lotteries (Inconsistent), including also a backward switch from B to A and/or ii) chose lottery A in row 10 *(Dominated)* (see for example Charness et al. (2018) and Filippin & Crosetto (2016)).

TABLE 1: HL-task

	Lottery A			Lottery B			CRRA[b,c]
	chance of: 2 USD	chance of: 1.60 USD	EV[a] Lottery A	chance of: 3.85 USD	chance of: 0.10 USD	EV[a] Lottery B	Range of constant relative risk aversion
	odds in percent:		in USD[b]	odds in percent:		in USD[b]	
1	10	90	1.64	10	90	0.475	$-\infty \leq r \leq 1.71$
2	20	80	1.68	20	80	0.85	$1.71 < r \leq 0.95$
3	30	70	1.72	30	70	1.23	$0.95 < r \leq 0.49$
4	40	60	1.76	40	60	1.60	$0.49 < r \leq 0.15$
5	50	50	1.80	50	50	1.98	$0.15 < r \leq 0.15$
6	60	40	1.84	60	40	2.35	$0.15 < r \leq 0.41$
7	70	30	1.88	70	30	2.73	$0.41 < r \leq 0.68$
8	80	20	1.92	80	20	3.10	$0.68 < r \leq 0.97$
9	90	10	1.96	90	10	3.48	$0.97 < r \leq 1.37$
10	100	0	2.00	100	0	3.85	$1.37 < r \leq 2.21$

[a]expected value; [b]column was not shown; [c]a power utility function of the form U(x) =[$x^{(1r)}$/(1-r)] is assumed (modified from Holt & Laury (2002))

In our experiment, we aimed at capturing the complexity, while still using some tools to ensure proper explanation. Thus, we provided the enumerators with 20 cards. The cards had different colors, 10 red cards and 10 blue cards. The enumerators were extensively trained to use the cards to explain the logic of the HL-task and to later demonstrate the probabilities in every round (see Appendix A for a detailed illustration). They used the cards as a complementary to a printed version of Table 1 (without illustrating the last column).

The lottery was directly incentivized. The enumerator was equipped with a small sack in which we kept 10 chips, numbered from one to ten. The numbers correspond to the rows in the HL-task. Once the participant has made their choices for all 10 rounds, they would draw a number (indicating the respective row) from the sack which then was played for the US Dollar amount stated in Table 1. For example, if the participant drew a 9, then the enumerator would check for the lottery chosen in row 9 and play that respective lottery with they. The enumerator showed the sack and the numbers to the participant before playing the lottery and demonstrated a possible payout.

Also, to make sure the task is well understood by the enumerators and in the field, we undertook a number of measures: 1) We carefully selected the enumerators out of a pool of students together with the Royal University of Cambodia; 2) We undertook an extensive training with the enumerators, making sure everyone understood the task; 3) We undertook a comprehensive pilot under the supervision of one of the authors with daily feedback rounds individually with every enumerator, learning and adjusting the instructions accordingly; and 4) One author accompanied the enumerator team throughout the entire data collection, again including daily feedback rounds on possible issues arising.

2.2 Cognitive skills: Raven's Progressive Matrices

Following the literature, we decided to include a cognitive test in order to later assess our results from the HL-task. We use the universally excepted RPM-test (Raven, 1938) to capture cognitive skills. The RPM-test measures fluid intelligence, which is independent from the previously learned. It is non-verbal and can be used for all age groups (Dean et al., 2017; Raven & Rust, 2008; Raven, 1938). The RPM-test involves a sequence of shapes with one shape missing. Participants must choose which of several alternatives best fits in the missing space (see Figure A.3 for an exemplification). Like Mani et al. (2013) and as proposed by Dean et al. (2017), we also used a compressed version of the original RPM-test (Raven & Rust, 2008). In its raw version, the test comprises of 60 puzzles. After conducting a pilot with smallholder farmers in the region, we decided on twelve puzzles with ascending difficulty to print for the field study. Within this study, we refer to the results of the RPM-test as *Cognitive skills*. There was no time limit and no direct incentive scheme for a paper-based test in order to not put additional pressure on the test taker.

3 Study area and study group

We conducted a field study with 247 smallholder farmers from rural Cambodia. The study builds on a survey that was implemented throughout sixteen villages of the Ratanakiri Province, remotely situated in northeastern Cambodia. The data collection took place from August to October 2018, including an intensive pilot (N=40) after a two-week training session with the enumerators recruited from the Royal University of Phnom Phen. With a gross national income per capita of 1,075 USD, Cambodia is clustered as a least developed country (United Nations, 2018) with the province being classified as one of the poorest areas in the country, which is the reason for conducting our data collection and experiments in that exact region. Of the 150,000 citizens, 88 percent live in rural areas and depend predominantly on smallholder agriculture (Ritzema et al., 2019; ADB, 2014).

Every participant needed to have at least basic skills in speaking and understanding the national language and needed to be a smallholder farmer household head. Since there are no household lists for the region available to researchers, we rely on the expert knowledge of the extension workers from the regional government. We undertook a random selection of participants out of the pool of villages/ individuals identified by experts.[2] Local enumerators privately guided the participants throughout the research session, which comprised of an experimental part and a questionnaire. One author accompanied the enumerator team during the entire data collection process. In the questionnaire we ask the participants about all basic socio-economic characteristics, such as their age, years or education, gender, ethnicity and spiritual group. Further, we also ask them to assess their numeracy skills by asking "can you

[2]However, we need to point out that participants needed to fulfil certain basic requirements as pointed out (basic language skills and being a farm household head). Thus, while choosing randomly, there was a pre-selection of potential participants.

calculate" while the participant can answer with yes or no. We summarize this variable as *Can calculate*. After an approximately three-hour session, the participants received a payout equivalent of the minimum daily wage. As previously described, the payout consisted of two parts: A fixed amount and the additional win.

As seen in Table 2, the farmers in our sample are on average 39.58 years old. The majority has some basic literacy, as 57% can calculate and schooling took place for 2.87 years, on average. With a share of 59%, there are slightly more females than males in our sample. With respect to *Cognitive skills*, the farmers scored, on average, 4.06 out of 12 correct puzzles in the RPM-test.

TABLE 2: Descriptive statistics (N=247)

Statistic	Unit	Mean	St. Dev.	Min	Max
Age	years	39.58	14.51	13.00	76.00
Can calculate	dummy[a]	0.57		0.00	1.00
Cognitive skills	ordinal[b]	4.06	2.67	0.00	12.00
Education	years	2.87	3.18	0.00	12.00
Ethnicity					
Jarai	dummy[a]	0.48		0.00	1.00
Khmer	dummy[a]	0.26		0.00	1.00
Other	dummy[a]	0.26		0.00	1.00
Gender	dummy[c]	0.59		0.00	1.00
Spiritual Group					
Buddhist	dummy[a]	0.54		0.00	1.00
Other	dummy[a]	0.36		0.00	1.00
None	dummy[a]	0.10		0.00	1.00
Answering patterns in the HL-task					
Consistent	dummy	0.30			
Single backswitch (ABA)	dummy	0.02			
Single backswitch (BA)	dummy	0.01			
Monotonous pattern (only A)	dummy	0.07			
Multiple switching	dummy	0.60			

[a] Self-assessment question, Yes=1, No=0; [b]Measured as the number of correct answers given in the RPM-test ranging from 0-12;[c] 1= Female, 0= Male;

As a starting point for the inconsistency variable, we report the choice pattern in the HL-task in detail in Table 2. First, in a very strict examination, 30% of our sample answer in

a consistent pattern. 2% of the sample switches back once, after moving from A to B (*Single switchback (ABA))*. Furthermore, 1% of our sample starts off with B and moves back once to A (*Single switchback (BA))* Next, *Monotonous pattern* reports those individuals who answer in a monotonous pattern (only A). In our sample this refers to 7 %. Monotonous answers in A might be of special interest in the specific context of our study, as this choice pattern can mirror another dimension of inconsistency. Here, participants either did not grasp the task, as hypothesised by (Liu, 2013), or did not set out to engage in the task, i.e. inattentiveness, as suggested by Amador-Hidalgo et al. (2021) when choosing A in the last row. Finally, 60% of our sample shows multiple switching behavior, i.e. switches more than twice between Lottery A and Lottery B.
For our analysis, we will use two inconsistency variables. First, a moderate indicator (*Inconsistency HL-task I*), which does not consider monotonous choice pattern to be inconsistent. In our sample, this accounts for 63 % of our sample (i.e. in this case 37% of the sample is clustered as consistent). Second, a strict indicator (*Inconsistency HL-task II*), where we add those who chose the dominated option in row 10, i.e. indicating the lack of attention rather than a calculation error (Amador-Hidalgo et al., 2021). In other words, *Inconsistency HL-task II* captures all participants that violated first-order stochastic dominance in the lottery. This applies to 70% of our sample. Moreover, to further examine multiple switching behavior, Figure 1 illustrates the relative share of individuals within our sample switching in a given row of the HL-task. We can report that below 10 % of our sample switches six, seven, eight, or nine times, respectively, with nine switches occurring only in rare occasions. However, 17 % of our sample switches three times and 15% switches five times. Thus, while switching back and forth throughout the entire task is not the typical choice pattern for the multiple switchers, they do appear to bounce back and forth from Lottery A to Lottery B quite frequently.

4 Results and Discussion

One main hypothesis for inconsistent answering behavior of low-numeracy subjects in the HL-task is due to its complexity, as participants have different levels of cognitive skills (He et al., 2018; Anderson & Mellor, 2009; Dave et al., 2010). We test this hypothesis with the data collected in the RPM-test. The results are depicted in Table 3.[3] In model (1) we consider only cognitive skills as a independent variable for explaining inconsistencies. Cognitive skills reveal to be a statistically significant determinant at the 1%-level of inconsistent choice behavior in the HL-task. If an individual scores more puzzles in the RPM-test, they decrease their likelihood of inconsistency in the HL-task by approximately 5 percentage points.

[3]We present the correlation coefficients of all variables included in Table A.1, showing that we only included variables below a correlation coefficient of 0.6.

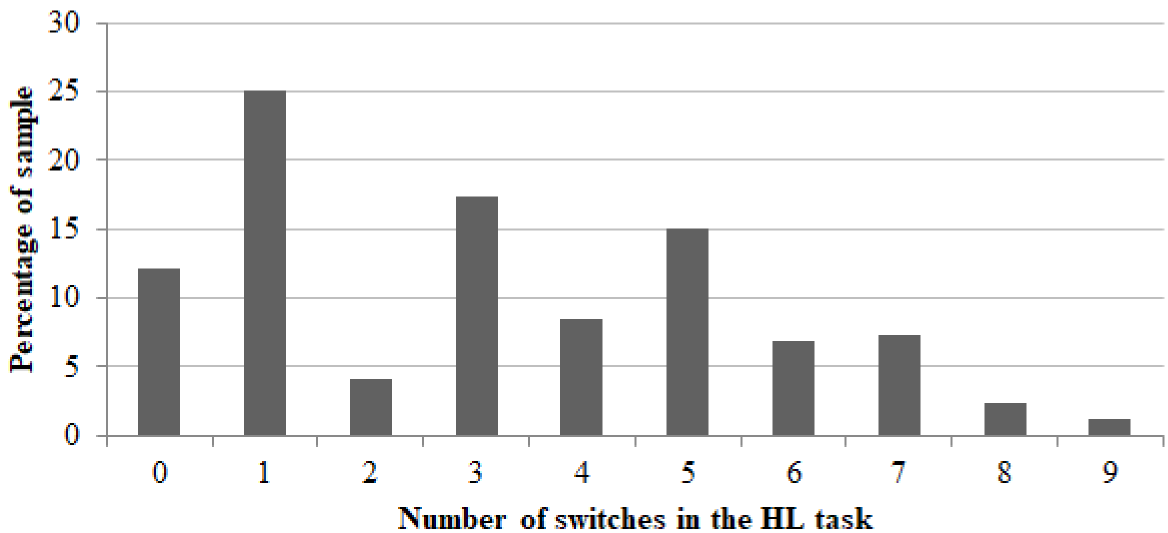

FIGURE 1: Switching behavior within the HL-task (N=247)

Note: This figure illustrates the number of switches within an individual HL task. For example, the first bar illustrates that 12.15% of participants in our sample did not switch at all, i.e. 0 switches.

TABLE 3: Results of the Logit Regression on inconsistencies (N=247)

	(1)[a]		(2)[a]		(3)[b]		(4)[c]	
	Coefficient	St. Err.	Coefficient	St. Err.	Coefficient	St. Err.	Coefficient	St. Err.
Age			0.02*	0.01	0.03**	0.01	-0.05	0.03
Can calculate			-0.12	0.36	-0.14	0.37	-0.19	0.75
Cognitive skills	-0.24***	0.06	-0.40***	0.11	-0.48***	0.12	-0.26	0.10
Cognitive skills*Gender			0.27**	0.13	0.38***	0.14	-0.02	0.26
Education			-0.01	0.05	-0.01	0.06	0.04	0.10
Gender (female=1)			-0.95	0.67	-1.34*	0.71	-0.03	1.30
Cons	1.92***	0.29	1.81	0.75	1.66	0.79	1.26	1.63
LR chi^2	20.59***		29.52***		33.11***		7.06	
N	247		247		222		89	

Note: $***p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$; [A]Logit regression on Inconsistency (*Inconsistency HL-task **II***); [B]Logit regression on Inconsistency (*Inconsistency HL-task **I***) [c]Logit regression on inattentive behavior, i.e. monotonous choice pattern of A. The subsample consists of those who showed monotonus behavior (dependent variable =1) and those who showed consistent behavior (dependent variable = 0).

In model (2) we add sociodemographic control variables to our analysis. Cognitive skills remain a statically significant determinant of inconsistency levels at the 1%-level. Age is also a statistically significant determinant of inconsistencies at the 5%-level for the HL-task. Thus, older participants are significantly more likely to behave inconsistently. For further robustness, we add model (3). In this model, we regress on *Inconsistency HL-task I*, i.e. only on those participants who showed multiple switching behavior. Generally, statistical significance, direction, and magnitude of the coefficients in model (3) are mostly in line with previous findings: Amador-Hidalgo et al. (2021), who investigate a large sample of Spanish Business Economics students and report a negative relationship between cognitive ability and inconsistent choice patterns in risk-related tasks (including the HL-task). Furthermore, Andersson et al. (2016) rely on a subject pool from the general Danish population and report similar findings, namely a positive relationship between consistent choice patterns in a multiple price list elicitation method and cognitive ability scores.

To shed light on different choice patterns, we add a further model. Based on Figure 1, we believe that those individuals who show multiple switching choose option A by chance in row 10. However, as we want to capture inattentive players separately, we strongly believe that these are represented in monotonous choices of A (thus also choosing A in row 10). Therefore, to study those individuals who are less engaged in the task rather than having potential difficulties in calculating separately, we add model (4) to Table 3. In this Logit regression, we use monotonous choice pattern as the dependent variable, i.e. monotonous choice pattern =1, consistent choice pattern =0. Cognitive skills do not remain a statistically significant driver when only considering monotonous choice patterns. This finding is in line with Amador-Hidalgo et al. (2021), who find no effect of cognitive skills on inattentive behavior among their comparably highly educated sample of students.

Furthermore, we go beyond mean effects by adding interactions with RPM-test results to our model. While almost all interactions lack statistical significance, one interaction, namely the one between Cognitive skills and Gender *is* statistically significant. Therefore, we add this interaction to our set of independent variables. The statistically significant coefficient of the interaction variable indicates that the decrease of inconsistent choices through higher cognitive skills is significantly reduced for female participants, i.e. the drop of inconsistent choices with higher cognitive skills is smaller for females. More specifically, the effect of cognitive skills on inconsistent choice behavior is not significant (p=0.181 for model (3)) for females. Thus, male participants are the dominant driver of the correlation between cognitive skills and inconsistent choices.

To demonstrate these differences between men and women within our sample, we illustrate inconsistent choice behavior over cognitive skills for both genders separately. Figure 2 presents the inconsistency in percent (y-axes) for each respective puzzle of the RPM-test (x-axes) by gender. Female participants (red line) show no clear pattern of inconsistent behavior over the RPM-test. For example, inconsistencies decrease for puzzle three and puzzle seven and peak in puzzle six, nine and twelve. However, for male

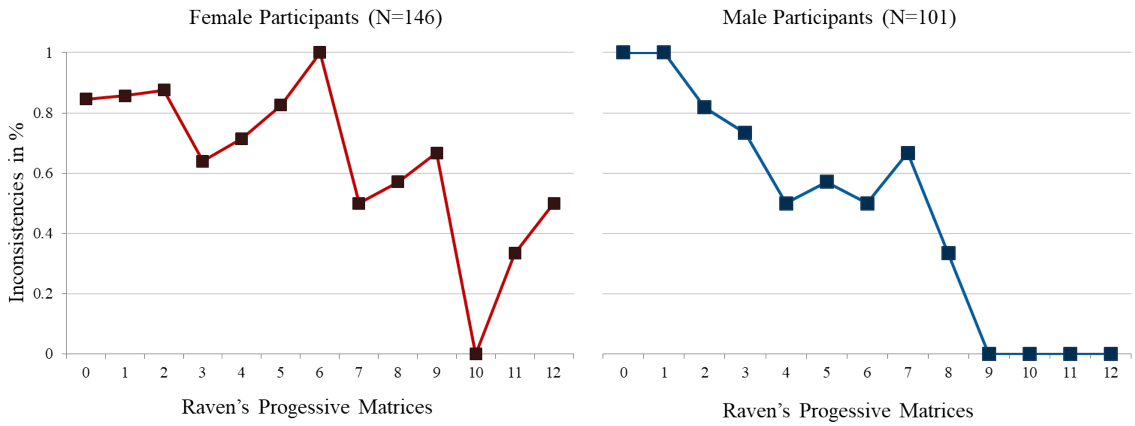

FIGURE 2: Results by Gender

participants we see a rather continual decline on inconsistencies when moving from zero correct puzzles to 12 correct puzzles in the RPM-test. Taken together, our results reveal that cognitive skills *are not* a reliable indicator for inconsistent choice behavior of women, however, cognitive skills *are* a reliable indicator for inconsistent choice behavior of men.

Our inconsistency levels are relatively high compared to other studies[4], yet still in line with the general pattern of findings. Brick et al. (2012) come up with an inconsistency rate of 41% in South Africa, Doerr et al. (2011) find 39% in Ethiopia, Jacobson & Petrie (2009) find 50% in Rwanda, Galarza (2009) finds that 50% of the farmers in Peru switch at least twice and Charness & Viceisza (2011) report inconsistencies of 75 % among farmers from rural Senegal. Even though Anderson & Mellor (2009) suggest that inconsistencies might indicate an indifference between the alternatives, it is more likely - considering literature and our findings - that it indicates the lack of understanding of the task. While there is no doubt that low educated individuals can engage in probability estimations (see for example Fontanari et al. (2014)), the special circumstances of a sudden experiment as well as the very nature of a complex lottery might cause lacking accuracy when played the field. Brick et al. (2012) suggest that MPL methods “although widely used in both student samples and in the field, might not be the most appropriate elicitation tool within a developing country context” (p.141). This opinion is shared by many other researchers as well (Charness & Viceisza, 2016; Holden, 2014; Nielsen et al., 2013; Dave et al., 2010).

5 Concluding Remarks

We have analyzed inconsistency levels within the HL-task among 247 low-literacy small-holder farmers. For this purpose, we collected data on sociodemographic and conducted the HL-task in rural Cambodia. We illustrated the HL-task using colored cards and the lottery

[4]However, even though our inconsistency rate is rather high, we are not surprised by it due to the low scores in the RPM-test.

table. In synthesis, our field study reveals that cognitive skills statistically significantly explain the inconsistencies in the HL-task at the 1% level. Our results therefore strongly suggest that - as hypothesized in literature - inconsistencies are a signal for the task at hand being too complex for the participant to fully engage in it, at least among low-literacy subjects. Furthermore, our analysis suggests cognitive skills are a relatively more reliable indicator for inconsistent choice behavior of men.

What do we draw from this? When attempting to measure risk attitudes in the field, researchers need to face the trade-off between complex and simple methods. A complex task - such as the HL-task - might reveal finer information, however the participants might not understand it and therefore the information is less meaningful. Our study suggests that the reason for inconsistencies is the lack of understanding, even though we can only confirm that for men. Therefore, if applied in the field, researchers should consider incorporating the RPM-test into their pre-test endeavors to understand potential suitability of the HL-task for the respective sample. This might be one way to use a complex lottery - thus reaping the benefits of rich information - while giving the participant the chance to actually cognitively engage in the lottery. Furthermore, based on recent suggestions form Branas-Garza et al. (2021) on the lacking impact of payment schemes for the performance in field experiments, it would be an interesting endeavor to conduct the HL-task in a similar study but with an indirect incentive scheme. Finally, it would be interesting for future research to investigate why the relationship between cognitive skills and inconsistencies is stronger for men than for women.

References

ADB, A. D. B. (2014). *Cambodia; country Poverty Analysis 2014*. O Books Sun & Moon.

Amador-Hidalgo, L., Brañas-Garza, P., Espín, A. M., García-Muñoz, T., & Hernández-Román, A. (2021). Cognitive abilities and risk-taking: Errors, not preferences. *European Economic Review*, *134*, 103694.

Anderson, L. R. & Mellor, J. M. (2009). Are risk preferences stable? Comparing an experimental measure with a validated survey-based measure. *Journal of Risk and Uncertainty*, *39*(2), 137–160.

Andersson, O., Holm, H. J., Tyran, J.-R., & Wengström, E. (2016). Risk aversion relates to cognitive ability: Preferences or noise? *Journal of the European Economic Association*, *14*(5), 1129–1154.

Branas-Garza, P., Estepa Mohedano, L., Jorrat, D., Orozco, V., & Rascon-Ramirez, E. (2021). To pay or not to pay: Measuring risk preferences in lab and field. *Judgment and Decision Making*, *16*(5), 1290–1313.

Brañas-Garza, P., Guillen, P., & del Paso, R. L. (2008). Math skills and risk attitudes. *Economics Letters*, *99*(2), 332–336.

Branas-Garza, P. & Rustichini, A. (2011). Organizing effects of testosterone and economic behavior: not just risk taking. *PloS one*, *6*(12), e29842.

Brick, K., Visser, M., & Burns, J. (2012). Risk aversion: Experimental evidence from South African fishing communities. *American Journal of Agricultural Economics*, *94*(1), 133–152.

Charness, G., Eckel, C., Gneezy, U., & Kajackaite, A. (2018). Complexity in risk elicitation may affect the conclusions: A demonstration using gender differences. *Journal of Risk and Uncertainty*, *56*(1), 1–17.

Charness, G., Gneezy, U., & Imas, A. (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization*, *87*, 43–51.

Charness, G. & Viceisza, A. (2011). Comprehension and risk elicitation in the field: Evidence from rural Senegal. IFPRI discussion papers 1135. *International Food Policy Research Institute (IFPRI)*.

Charness, G. & Viceisza, A. (2016). Three Risk-elicitation Methods in the Field: Evidence from Rural Senegal. *Review of Behavioral Economics*, *3*(2), 145–171, https://doi.org/10.1561/105.00000046 http://dx.doi.org/10.1561/105.00000046.

Crosetto, P. & Filippin, A. (2013). The "bomb" risk elicitation task. *Journal of Risk and Uncertainty*, *47*(1), 31–65.

Dave, C., Eckel, C. C., Johnson, C. A., & Rojas, C. (2010). Eliciting risk preferences: When is simple better? *Journal of Risk and Uncertainty*, *41*(3), 219–243.

Dean, E. B., Schilbach, F., & Schofield, H. (2017). Poverty and cognitive function. In *The economics of poverty traps* (pp. 57–118). University of Chicago Press.

Doerr, U., Toman, O. M., & Schmidt, U. (2011). Overconfidence and risk management of Ethiopian farmers. *University of Kiel (Working paper)*.

Estepa-Mohedano, L. & Espinosa, M. P. (2021). Comparing risk elicitation in lotteries with visual or contextual framing aids. *Munich Personal RePEc Archive*, (108440).

Filippin, A. & Crosetto, P. (2016). A reconsideration of gender differences in risk attitudes. *Management Science*, *62*(11), 3138–3160.

Fontanari, L., Gonzalez, M., Vallortigara, G., & Girotto, V. (2014). Probabilistic cognition in two indigenous Mayan groups. *Proceedings of the National Academy of Sciences*, *111*(48), 17075–17080.

Galarza, F. B. (2009). *Choices under Risk in Rural Peru*. Staff Papers 92247, University of Wisconsin-Madison, Department of Agricultural and Applied Economics https://ideas.repec.org/p/ags/wisagr/92247.html.

He, P., Veronesi, M., & Engel, S. (2018). Consistency of risk preference measures: An artefactual field experiment from rural china. *The Journal of Development Studies*, *54*(11), 1955–1973.

Holden, S. T. (2014). Risky choices of poor people: Comparing risk preference elicitation approaches in field experiments.

Holt, C. A. & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, *92*(5), 1644–1655.

Ihli, H. J., Gassner, A., & Musshoff, O. (2018). Experimental insights on the investment behavior of small-scale coffee farmers in central Uganda under risk and uncertainty. *Journal of Behavioral and Experimental Economics*, *75*, 31–44.

Jacobson, S. & Petrie, R. (2009). Learning from mistakes: What do inconsistent choices over risk tell us? *Journal of risk and uncertainty*, *38*(2), 143–158.

Lejuez, C. W., et al. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, *8*(2), 75.

Liu, E. M. (2013). Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China. *Review of Economics and Statistics*, *95*(4), 1386–1403.

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *science*, *341*(6149), 976–980.

Mosley, P. & Verschoor, A. (2005). Risk attitudes and the 'vicious circle of poverty'. *The European journal of development research*, *17*(1), 59–88.

Nielsen, T., Keil, A., & Zeller, M. (2013). Assessing farmers' risk preferences and their determinants in a marginal upland area of Vietnam: a comparison of multiple elicitation techniques. *Agricultural Economics*, *44*(3), 255–273.

Raven, J. C. (1938). *Guide to Using Progressive Matrices (1938):(general Arrangement of the Scale-the Individual Test, the Self Administered Or Group Test-records and Marking-key.)*. HK Lewis.

Raven, J. C. & Rust, J. (2008). *Standard Progressive Matrices-Plus Version and Mill Hill Vocabulary Scale: Manual;[SPM+/MHV]*. Pearson.

Ritzema, R. S., et al. (2019). Household-level drivers of dietary diversity in transitioning agricultural systems: evidence from the Greater Mekong Subregion. *Agricultural Systems*, *176*, 102657.

Rommel, J., Hermann, D., Müller, M., & Mußhoff, O. (2019). Contextual framing and monetary incentives in field experiments on risk preferences: Evidence from German farmers. *Journal of Agricultural Economics*, *70*(2), 408–425.

United Nations (2018). The Least Developed Country Category: 2018 Country Snapshots.

Verschoor, A., D'Exelle, B., & Perez-Viana, B. (2016). Lab and life: Does risky choice behaviour observed in experiments reflect that in the real world? *Journal of Economic Behavior & Organization*, *128*, 134–148.

Vollmer, E., Hermann, D., & Mußhoff, O. (2017). Is the risk attitude measured with the Holt and Laury task reflected in farmers' production risk? *European Review of Agricultural Economics*, *44*(3), 399–424.

World Bank (2016). Poverty: Overview. *World Bank*.

World Bank (2018). Poverty and shared prosperity 2018: Piecing together the poverty puzzle.

World Bank (2020). World Development Indicators database. data retrieved online, https://datacatalog.worldbank.org/dataset/world-development-indicators.

Appendix

Appendix A: HL-task instruction example

In the following, we present the example from our training session, adjusted throughout the pilot. Moderation and suggested actions are in italic:

Dear Joe, We will now play a lottery. The lottery is simple. We will play 10 rounds and in every round you can choose between 2 Options. Option A and Option B. I will show you what the two Options hold for you using these cards. If you have any questions, please interrupt me anytime. At the end of the 10 rounds we will draw a number between 1-10 and for the respective round you will get the actual payout of the gamble! So we play for real money that you can win on top of your salary for this session!!! Ok so let's start!

	A		B	
	$2	$1.60	$3.85	$0.1
1	1	9	1	9
2	2	8	2	8
3	3	7	3	7
4	4	6	4	6
5	5	5	5	5
6	6	4	6	4
7	7	3	7	3
8	8	2	8	2
9	9	1	9	1
10	10	0	10	0

FIGURE A.1: Explanation of the HL-task

Please show him/her the HL table and explain the lotteries as fields. As you can see, there are two options. Option A and Option B. These are the two options you can choose between. Here you can see the constant payouts. In Option A this will always be $2 or $1.60 and in Option B this will always be $3.85 or $0.10. This will never change. The only thing that changes are the likelihoods of winning one of the two amounts. *Now set the table* *aside for a moment and take out your cards. To illustrate the probabilities to you, I will use colored cards. Lay down 10 cards.*

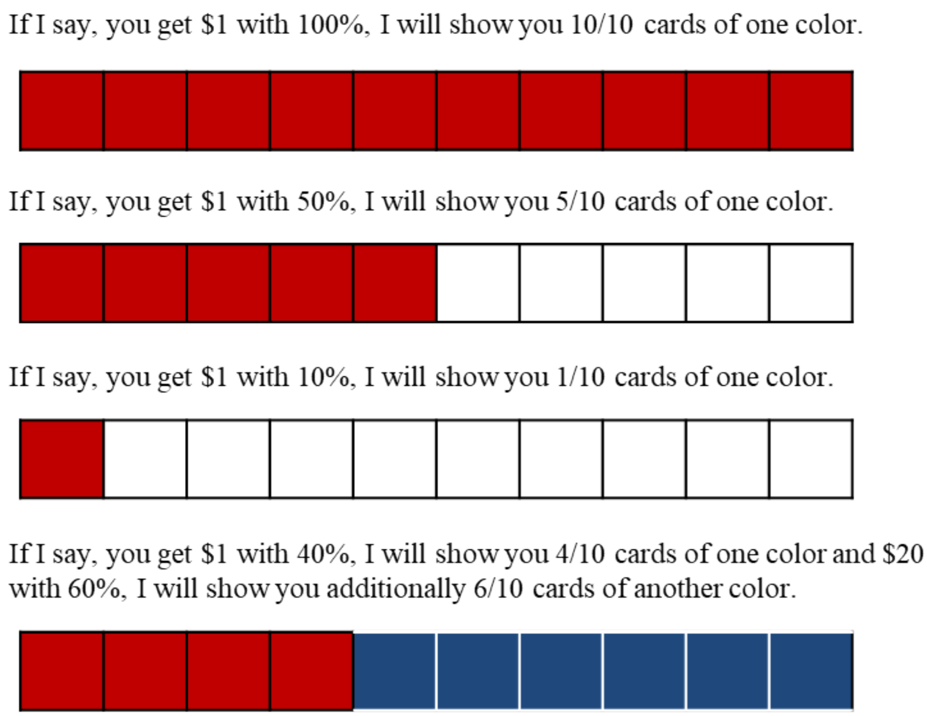

FIGURE A.2: Explanation of probabilities and the use of colored cards during the study

As you can see: More cards of one color mean higher chances to win that respective amount of money! Fewer cards mean lower chances! So, now let's get back to the game. I will start with round one and you tell me which option you like best. And always keep in mind, you are playing for real money! You can win the amounts shown, so choose wisely! Let's start with a test round. (Please let the participant point out any random row of the table and pick that row as an example. I use round one to illustrate the procedure: In round one, A (i.e. in field A), you have the chance to win $2 with 10% (lay down one red card) or $1.60 with 90% (lay down nine blue cards) and in B (i.e. in field B), you have the chance to win $3.85 with 10% (lay down one red card) or $0.10 with 90% chance (lay down 9 blue cards). Which gamble A or B do chose in this round? *The participant can ask questions. Please wait patiently until he/she chooses Lottery A or B. Then, please illustrate the lottery by then actually playing the given option with him/her. Also, please illustrate what would have happened if he/she had chosen the other lottery. Please ask the participant afterwards if there are any remaining questions, otherwise the game starts. Please start with round one note down the results in your EXPERIMENT RESULTS SHEET and continue with the next row. Do so until down with all rows.*

Great job! Now, as promised, you can draw a card out of this box. There is a number between one and ten and will determine which row we will play for real money. (*let him/her draw the number out of your box A*).

Let's say, the number is 1. Then you check you sheet to see what he/she has answered in round 1. Let's say he/she chose option A in round 1. Then you use your Box B and put the right share of two different colored cards in there: here, 1 blue and 9 red. *Then, with closed eyes the participant draws one card. The respective amount is written down on the EXPERIMENT RESULTS SHEET and will be added at the end.*

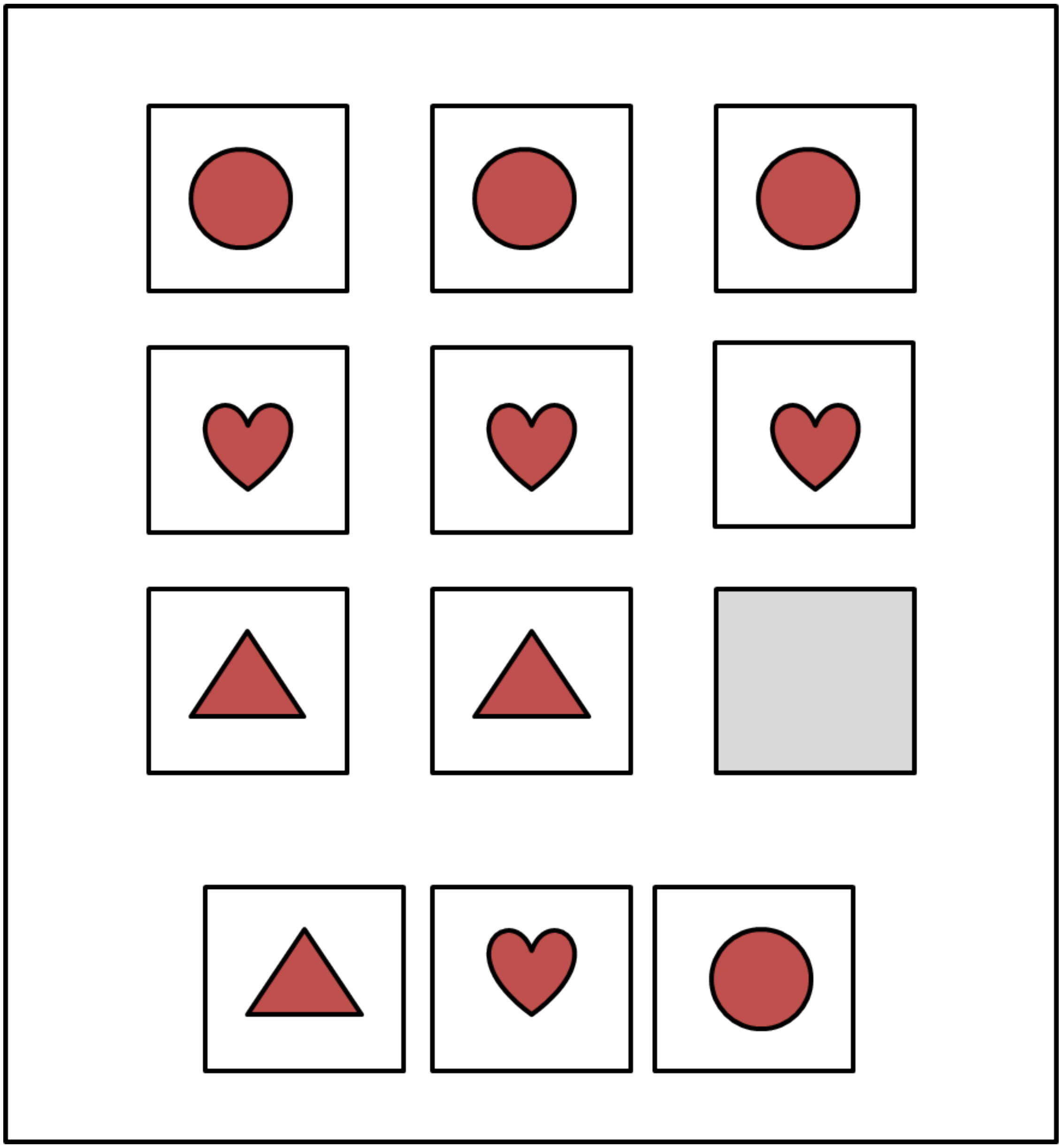

FIGURE A.3: Illustration of the core mechanism of the RPM-test, designed by author

Note: In our study, we exclusively rely on the original Raven's Matrices (Raven & Rust, 2008). This self-compiled figure is merely for demonstration purposes of the test's logic and not one of the puzzles used.

TABLE A.1: Spearman Correlation

	Age	Can calculate	Cognitive skills	Education	Gender
Age	1.00				
Can calculate	-0.15**	1.00			
Cognitive skills	-0.17***	0.32***	1.00		
Education	-0.27***	0.56***	0.31***	1.00	
Gender	-0.12	-0.22***	0.02	-0.01	1.00