INSTITUTE for REPLICATION

No. 43
I4R DISCUSSION PAPER SERIES

Replication Report on Altmann et al. (2022)

Sebastian Bachler

Andrea Erhart

Armando Holzknecht

June 2023

INSTITUTE for REPLICATION

I4R DISCUSSION PAPER SERIES

I4R DP No. 43

Replication Report on Altmann et al. (2022)

Sebastian Bachler[1], Andrea Erhart[2], Armando Holzknecht[1]

[1]Department of Banking and Finance, University of Innsbruck/Austria
[2]Department of Economics, University of Innsbruck/Austria

JUNE 2023

E-Mail: joerg.peters@rwi-essen.de
RWI – Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
45128 Essen/Germany
www.i4replication.org

ISSN: 2752-1931

Replication report on Altmann et al. (2022)*

Sebastian Bachler Andrea Erhart

Armando Holzknecht

Abstract

In the paper of, Altmann et al. (2022) the authors investigate whether positive effects which are due to behavioral policy interventions in policy-targeted domains come along with negative effects in policy non-targeted domains. Using lab and online experiments where subjects have to solve one policy-focused decision task and one non-focused background task, the authors show that increasing incentives or steering attention to the former led to higher attention spans, lower default adherence rates, and a higher choice quality in the decision task. However, because of steering participants focus to the decision task, lower choice quality and lower attention spans in the background task emerged as a consequence, which was particularly pronounced among individuals with lower cognitive capabilities and complex decision tasks. Essentially, the authors also describe that the negative effects in the background tasks offset the positive effects in the decision task, ultimately yielding a net-zero effect overall. Therefore, the authors emphasize policymakers to also consider the potential negative cognitive spillovers in order to not overestimate the benefits of behavioral policy interventions. All the results the authors in the main text report are significant on 5% and 1% significance levels. All findings presented in the main text of the paper can be replicated using the original Stata code and verified thoroughly using R. Additionally, we performed two robustness tests to ensure the reliability of the paper's main results, and they remained consistent. Hence, the reported findings in the paper appear to be robust.

*Authors: Bachler: University of Innsbruck, Department of Banking and Finance. E-Mail: sebastian.bachler@uibk.ac.at. Erhart: University of Innsbruck, Department of Economics. E-Mail: andrea.erhart@uibk.ac.at. Holzknecht: University of Innsbruck, Department of Banking and Finance. E-mail: armando.holzknecht@uibk.ac.at. Corresponding author: Armando Holzknecht
Financial support from the Network of Banking, Accounting, Auditing, Finance & IT (BAFIT) from the University of Innsbruck as well as from the Austrian Science Fund (SFB F63) are gratefully acknowledged.

1 Introduction

In this report, we first attempt to reproduce the results of the main text using exclusively materials from the replication package provided by Altmann et al. (2022, https://zenodo.org/record/5652808). Second, in order to validate the results, we try to replicate the findings in `R` using only the data from the aforementioned replication package. Finally, we check for robustness by conducting two different analysis specifications to the results of the analyses highlighted in the abstract. All results are shown in Section 5. The `R` script for our reproduction and robustness checks can be found separately on OSF (https://osf.io/kugbs/).

2 Reproduction

Regarding reproducibility, we find that all results described in the main text of the paper can be reproduced with the original Stata code and also validated to a full extent (100% out of all results) using `R` and the methods described in the paper, with only some minor deviations of test statistics and estimates at the 3rd decimal place. Beyond that, we had no substantial issues running the original code. Two minor issues we faced in the reproduction was that first, the declaration of control variables in the table notes is not sufficient and second, the derivation of payoffs and more precisely how the data had to be subsetted was not obvious for exact reproduction. Without checking the Stata code for this information, we would not have been able to fully reproduce the results of the paper.

3 Robustness

As a first robustness check, considering the binary nature of some of the dependent variables, we computed logit instead of linear probability models. The results of the logit models do not lead to any different conclusions and, therefore, support the findings from the paper. As a second check for robustness, we altered the battery of control variables in order to see if this influences any of the results. Especially,

the inclusion of interaction terms in the original battery of control variables seemed arbitrary to us. The alternative battery of control variables can be found in Table 1. However, we again find no qualitative differences in the results and can, therefore, confirm the authors' findings. Regarding continuous dependent variables, we did not deviate from using OLS regressions and only changed the selection of control variables. Either way, the core results of the paper remained unchanged and the reported results from the paper seem robust.

4 Conclusion

We conclude that all the main results reported in the paper are reproducible with both Stata (original code) and `R` (own code), while also being robust to robustness checks we introduced. Generally, we have the impression that the results from the paper would also hold for multiple other robustness checks, which ultimately could be checked with a multiverse analysis approach following Simonsohn et al. (2020). We believe this would go beyond the scope of this project, however, do not want to leave this unmentioned, as it circumvents the problem of finding significant results solely based on analytical heterogeneity.

Altmann et al. (2022)	Alternative
age, gender, study subject (harmonized categories), high school GPA (category), last math grade (category), self-perception of multitasking easy (binary), self-reported multitasking ability compared to other, wave fixed effect, ability in decision task, ability in background task, wave fixed effect × ability load, wave fixed effect × ability task	age, gender, subject (all categories), high school GPA (continuous), last math grade (continuous), self-reported multitasking ability compared to other, ability in decision task, ability in background task, wave fixed effect, decision task hard difficulty (binary)

Table 1: Control batteries

5 Results

Table 2: In-text statistics

Statistical Test	**Paper**	**Reproduction**
Kolmogrov-Smirnov test: Attention across Baseline/relative incentives treatments (p. 2306)	$p<0.001$ for all pairwise treatment comparisons	$p<0.001$ for all pairwise treatment comparisons
Correlation: Attention and relative incentives (p. 2307)	$\rho=0.0463$, $p<0.001$	$\rho=0.0463$, $p<0.001$
Correlation: Propensity devote and relative incentives (p. 2307)	$\rho=0.0425$, $p<0.001$	$\rho=0.0425$, $p<0.001$
Correlation: Performance and relative incentives (p. 2308)	$\rho=0.0496$, $p<0.001$	$\rho=0.0496$, $p<0.001$
Correlation: Sticking to default and relative incentives (p. 2308)	$\rho=-0.0439$, $p<0.001$	$\rho=-0.0439$, $p<0.001$
Correlation: Sticking to default and relative incentives (p. 2308)	$\rho=-0.0439$, $p<0.001$	$\rho=-0.0439$, $p<0.001$
Correlation: Sticking to default and relative incentives when default correct (p. 2309)	$\rho=-0.071$, $p=0.219$	$\rho=-0.071$, $p=0.219$
Correlation: Sticking to default and relative incentives when default incorrect (p. 2309)	$\rho=-0.528$, $p<0.01$	$\rho=-0.528$, $p<0.001$

Notes: The table reports the statistical tests mentioned in-line in the main text of Altmann et al. (2022). The first column shows which statistical test was conducted including a page reference. The second column includes the estimates reported in paper, while the third column reports the estimates from our reproduction. Correlation estimates reported are Spearman's rank correlation coefficients. Besides the statistical tests reported in this table, the authors also include p-values of t-tests for coefficient estimates of multiple regressions of the main text. As we also reproduce these regressions, and hence also the associated t-tests, we do not include the reproduction of these in this table. For this we refer to the other tables in the Appendix.

Table 3: Attention and choice quality in Baseline

Paper

	Avg. attention		Attention = 0		Choice quality	
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline-20	1.798	1.251	−0.059	−0.045	0.031	0.046
	(1.810)	(1.984)	(0.066)	(0.075)	(0.040)	(0.044)
Baseline-40	4.223*	2.964	−0.190***	−0.167**	0.106***	0.109***
	(2.227)	(2.429)	(0.060)	(0.072)	(0.038)	(0.042)
Baseline-Ample	11.627***	12.295***	−0.282***	−0.291***	0.252***	0.233***
	(1.174)	(1.480)	(0.035)	(0.043)	(0.023)	(0.025)
Constant	14.137***	9.702	0.307***	0.319	0.583***	0.335***
	(0.961)	(6.135)	(0.032)	(0.201)	(0.018)	(0.109)
Controls	No	Yes	No	Yes	No	Yes
N	4060	4060	5960	5960	5960	5960
No. Subjects	203	203	298	298	298	298
$R2$	0.171	0.253	0.100	0.181	0.058	0.090
Baseline-20=Baseline-40	0.338	0.475	0.090	0.100	0.129	0.158
Baseline-20=Baseline-A.	0.000	0.000	0.000	0.005	0.000	0.000
Baseline-40=Baseline-A.	0.001	0.001	0.086	0.143	0.000	0.012

Reproduction

	Avg. attention		Attention = 0		Choice quality	
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline-20	1.798	1.251	−0.059	−0.045	0.031	0.046
	(1.810)	(1.984)	(0.066)	(0.075)	(0.040)	(0.044)
Baseline-40	4.223*	2.964	−0.190***	−0.167**	0.106***	0.109***
	(2.227)	(2.429)	(0.060)	(0.072)	(0.038)	(0.042)
Baseline-Ample	11.627***	12.295***	−0.282***	−0.291***	0.252***	0.233***
	(1.174)	(1.480)	(0.035)	(0.043)	(0.023)	(0.025)
Constant	14.137***	9.702	0.307***	0.319	0.583***	0.335***
	(0.961)	(6.135)	(0.032)	(0.201)	(0.018)	(0.109)
Controls	No	Yes	No	Yes	No	Yes
N	4060	4060	5960	5960	5960	5960
No. Subjects	203	203	298	298	298	298
$R2$	0.171	0.253	0.100	0.181	0.058	0.090
Baseline-20=Baseline-40	0.337	0.475	0.089	0.099	0.128	0.157
Baseline-20=Baseline-A.	0.000	0.000	0.000	0.005	0.000	0.000
Baseline-40=Baseline-A.	0.000	0.001	0.085	0.142	0.000	0.011

Notes: The table refers to Table B.1 of Altmann et al. (2022, p. 2325)reports results of OLS regressions of attention devoted to the decision task and choice quality in this task for the treatments in the Baseline environment. On top the results in the paper are shown and below are the results of our reproduction using `R`. The dependent variable in Columns (1)–(2) is the number of seconds that a subject enters the decision task in a given round of the experiment. The dependent variable in Columns (3)–(4) is an indicator equal to 1 if a subject does not enter the decision task in a given round of the experiment and 0 otherwise. The dependent variable in Columns (5)–(6) is an indicator equal to 1 if a subject correctly solves the decision task in a given round of the experiment and 0 otherwise. Specifications with controls include the set of control variables specified in Table 2 of Altmann et al. (2022, p. 2310). The lower part of the table reports p-values from post-estimation tests of the equality of selected coefficients (Wald tests). Robust standard errors, clustered at the subject level, are reported in parentheses. ***, **, * indicates significance at the 1%, 5%, and 10% level, respectively.

Table 4: Behaviour in the targeted choice domain

Paper

	Panel (a): Choice quality				Panel (b): Default adherence			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intervention	0.107***	0.110***			−0.217***	−0.219***		
	(0.018)	(0.017)			(0.022)	(0.022)		
Directed Attention			0.105***	0.110***			−0.137***	−0.139***
			(0.020)	(0.019)			(0.026)	(0.025)
Forced Choice			0.109***	0.111***			−0.300***	−0.302***
			(0.021)	(0.019)			(0.021)	(0.021)
Constant	0.599***	0.262***	0.599***	0.262***	0.562***	0.689***	0.562***	0.678***
	(0.015)	(0.068)	(0.015)	(0.068)	(0.020)	(0.088)	(0.020)	(0.085)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
N	12,160	12,160	12,160	12,160	12,160	12,160	12,160	12,160
No. subjects	608	608	608	608	608	608	608	608
R^2	0.012	0.041	0.012	0.041	0.043	0.053	0.061	0.070
Directed=Forced			0.848	0.949			0.000	0.000

Reproduction

	Panel (a): Choice quality				Panel (b): Default adherence			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intervention	0.107***	0.110***			−0.217***	−0.219***		
	(0.018)	(0.017)			(0.022)	(0.022)		
Directed Attention			0.599***	0.262***			0.562***	0.678***
			(0.015)	(0.068)			(0.020)	(0.085)
Forced Choice			0.105***	0.110***			−0.137***	−0.139***
			(0.020)	(0.019)			(0.026)	(0.025)
Constant	0.599***	0.262***	0.109***	0.111***	0.563***	0.689***	−0.300***	−0.302***
	(0.015)	(0.068)	(0.021)	(0.019)	(0.020)	(0.088)	(0.021)	(0.021)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
N	12,160	12,160	12,160	12,160	12,160	12,160	12,160	12,160
No. subjects	608	608	608	608	608	608	608	608
R^2	0.012	0.041	0.012	0.041	0.043	0.053	0.061	0.070
Directed=Forced			0.848	0.949			0.000	0.000

Notes: The table refers to Table 2 of Altmann et al. (2022, p. 2310) and reports results of OLS regressions of treatment differences in choice quality and default adherence in the decision task. On top the results in the paper are shown and below are the results of our reproduction using R. The dependent variable in Columns (1)–(4) is an indicator equal to 1 if a subject correctly solves the decision task in a given round of the experiment and 0 otherwise. The dependent variable in Columns (5)–(8) is an indicator equal to 1 if a subject's choice in a given round of the experiment coincides with the (choice alternative corresponding to the) default option and 0 otherwise. Control variables in Columns (2), (4), (6), and (8) include subjects' age, gender, field of study, high school GPA and last high school math grade, ability in the background and decision task, perceived multitasking ability, and indicator variables for subjects' assigned payoff scheme (cp. Table 1 of Altmann et al. (2022, p. 2300)) and experiment wave. The lower part of the table reports p-values from post-estimation tests of the equality of selected coefficients (Wald tests). Robust standard errors, clustered at the subject level, are reported in parentheses. ***, **, * indicates significance at the 1%, 5%, and 10% level, respectively

Table 5: Cognitive spillovers

Paper

	Lab experiment				Online experiment	
	(1)	(2)	(3)	(4)	(5)	(6)
Intervention	−0.028*	−0.038***			−0.114***	−0.113***
	(0.017)	(0.014)			(0.019)	(0.017)
Directed Attention			−0.023	−0.032*		
			(0.020)	(0.016)		
Forced Choice			−0.034*	−0.044***		
			(0.020)	(0.016)		
Constant	0.760***	0.519***	0.760***	0.519***	0.733***	0.646***
	(0.014)	(0.064)	(0.014)	(0.064)	(0.012)	(0.060)
Controls	No	Yes	No	Yes	No	Yes
N	12,160	12,160	12,160	12,160	5,472	5,472
No. Subjects	608	608	608	608	608	608
R^2	0.001	0.074	0.001	0.075	0.031	0.116
Directed=Forced		0.606	0.453			

Reproduction

	Lab experiment				Online experiment	
	(1)	(2)	(3)	(4)	(5)	(6)
Intervention	−0.028*	−0.038***			−0.114***	−0.113***
	(0.017)	(0.014)			(0.019)	(0.017)
Directed Attention			−0.023	−0.032*		
			(0.020)	(0.016)		
Forced Choice			−0.034*	−0.044***		
			(0.020)	(0.016)		
Constant	0.760***	0.519***	0.760***	0.519***	0.733***	0.646***
	(0.014)	(0.064)	(0.014)	(0.064)	(0.012)	(0.060)
Controls	No	Yes	No	Yes	No	Yes
N	12,160	12,160	12,160	12,160	5,472	5,472
No. Subjects	608	608	608	608	608	608
R^2	0.001	0.074	0.001	0.075	0.031	0.116
Directed=Forced		0.606	0.453			

Notes: The table refers to Table 3 of Altmann et al. (2022, p. 2312) and reports results of OLS regressions of treatment differences in choice quality in the background task. On top the results in the paper are shown and below are the results of our reproduction using `R`. The dependent variable in Columns (1)–(4) is an indicator equal to 1 if a subject correctly solves the background task in a given round of the laboratory experiment and 0 otherwise. The dependent variable in Columns (5)–(6) is the fraction of pairs that a subject uncovers in the Memory game in a given round of the online experiment. Columns (2) and (4) include the set of control variables specified in Table 2 of Altmann et al. (2022, p. 2310). Control variables in Column (6) include subjects' age, gender, and screen size. The lower part of the table reports p-values from post-estimation tests of the equality of selected coefficients (Wald tests). Robust standard errors, clustered at the subject level, are reported in parentheses. ***, **, * indicates significance at the 1%, 5%, and 10% level, respectively.

Table 6: Effect of interventions on overall payoffs

Paper

	(Payoff Share)			
	(1)	(2)	(3)	(4)
Intervention	0.006	−0.000		
	(0.013)	(0.011)		
Directed Attention			0.005	0.001
			(0.016)	(0.012)
Forced Choice			0.006	−0.001
			(0.015)	(0.012)
Constant	0.725***	0.475***	0.725***	0.475***
	(0.011)	(0.050)	(0.011)	(0.050)
Controls	No	Yes	No	Yes
N	12,160	12,160	12,160	12,160
No. Subjects	608	608	608	608
R^2	0.000	0.074	0.000	0.074
Directed=Forced			0.936	0.928

Reproduction

	(Payoff Share)			
	(1)	(2)	(3)	(4)
Intervention	0.006	−0.000		
	(0.013)	(0.011)		
Directed Attention			0.005	0.001
			(0.016)	(0.012)
Forced Choice			0.006	−0.001
			(0.015)	(0.012)
Constant	0.725***	0.475***	0.725***	0.475***
	(0.011)	(0.050)	(0.011)	(0.050)
Controls	No	Yes	No	Yes
N	12,160	12,160	12,160	12,160
No. Subjects	608	608	608	608
R^2	0.000	0.074	0.000	0.074
Directed=Forced			0.936	0.928

Notes: The table refers to Table 4 of Altmann et al. (2022, p. 2317) and reports results of OLS regressions of treatment differences in subjects' overall payoffs. On top the results in the paper are shown and below are the results of our reproduction using R. The dependent variable is a subject's realized share of the maximally attainable overall payoff in the subject's treatment (calculated on the subject-round level). Specifications with controls include the set of control variables specified in Table 2 Altmann et al. (2022, p. 2310). The lower part of the table reports p-values from post-estimation tests of the equality of selected coefficients (Wald tests). Robust standard errors, clustered at the subject level, are reported in parentheses. ***, **, * indicates significance at the 1%, 5%, and 10% level, respectively.

Table 7: Hypothetical payoffs across incentive conditions

Paper

	(Hypothetical) payoffs in treatment		
	Baseline-10	Baseline-20	Baseline-40
Subjects in Baseline-10	0.736	0.711	0.679
	(0.161)	(0.146)	(0.140)
Subjects in Baseline-20	0.730	0.711	0.687
	(0.129)	(0.122)	(0.132)
Subjects in Baseline-40	0.656	0.661	0.668
	(0.206)	(0.176)	(0.146)
Difference to highest hypothetical payoff	0.006	0.000	−0.019
p-values	0.832	0.988	0.629

Reproduction

	(Hypothetical) payoffs in treatment		
	Baseline-10	Baseline-20	Baseline-40
Subjects in Baseline-10	0.736	0.710	0.679
	(0.161)	(0.146)	(0.141)
Subjects in Baseline-20	0.730	0.711	0.687
	(0.129)	(0.122)	(0.132)
Subjects in Baseline-40	0.656	0.661	0.668
	(0.206)	(0.176)	(0.145)
Difference to highest hypothetical payoff	0.006	0.000	−0.019
p-values	0.811	0.987	0.644

Notes: The table refers to Table 5 of Altmann et al. (2022, p. 2319) and reports empirically observed and hypothetical shares of overall payoffs in the experiment. On top the results in the paper are shown and below are the results of our reproduction using R. Within a given row, the behaviour of subjects in the respective treatment is held fixed. Values on the diagonal depict the empirically observed share of the maximally attainable overall payoff in a given treatment. Off-diagonal values depict the corresponding payoff share subjects would have earned in the treatments denoted at the top of each column. At the bottom of the table, we report the difference between the empirically observed payoff share and the highest hypothetical value in each column, and p-values from t-tests of equality of the empirically observed and highest hypothetical payoff share.

Table 8: Attention and choice quality in Baseline
Robustness

	Attention			Attention = 0		
	(1)	(1a)	(1b)	(2)	(2a)	(2b)
Baseline-20	1.251	1.251	0.530	−0.045	−0.048	−0.003
	(1.984)	(1.988)	(2.149)	(0.075)	(0.072)	(0.081)
Baseline-40	2.964	2.964	2.411	−0.167**	−0.168**	−0.140*
	(2.429)	(2.435)	(2.571)	(0.072)	(0.067)	(0.084)
Baseline-Ample	12.295***	12.295***	11.894***	−0.291***	−0.285***	−0.279***
	(1.480)	(1.483)	(1.491)	(0.043)	(0.033)	(0.034)
Constant	9.702	11.152*	6.607	0.319		
	(6.135)	(6.197)	(5.925)	(0.201)		
Controls	Original	Original	Alternative	Original	Original	Alternative
Round FE	No	Yes	No	No	No	No
N	4060	4060	3960	5960	5960	5820
No. Subjects	203	203	198	298	298	291
R^2	0.253	0.262	0.22	0.181	0.228	0.201
Baseline-20=Baseline-40	0.475	0.476	0.475	0.100	0.125	0.181
Baseline-20=Baseline-A.	0.000	0.000	0.000	0.005	0.001	0.000
Baseline-40=Baseline-A.	0.001	0.001	0.001	0.143	0.029	0.017

Notes: Columns (1) and (2) refer to Columns (2) and (4) in Table B.1 of Altmann et al. (2022, p.2352), reporting results of OLS regressions of attention devoted to the decision task for the treatments in the Baseline environment. Alongside results of robustness checks with two different specifications (Columns a and b, respectively) are reported. The dependent variable in Columns (1)–(1b) is the number of seconds that a subject enters the decision task in a given round of the experiment. The dependent variable in Columns (2)–(2b) is an indicator equal to 1 if a subject does not enter the decision tasks in a given round and 0 otherwise. Column (1a) additionally controls for round fixed effects. Column (1b) includes an alternative battery of control variables (see Table 1). Column (2a) shows the result of a logit regression with the same explanatory variables as Column (2). Column (2) reports also the results of a logit regression with the same alternative battery of control variables as Column (1b). For logit regression average marginal effects are shown. The lower part of the table reports p-values from post-estimation tests of the equality of selected coefficients (Wald tests). Robust standard errors, clustered at the subject level, are reported in parentheses. ***, **, * indicates significance at the 1%, 5%, and 10% level, respectively. R^2 for logit regressions refers to McFadden's R^2.

Table 9: Behaviour in the targeted choice domain
Robustness

	Default adherence			Choice quality		
	(1)	(1a)	(1b)	(2)	(2a)	(2b)
Intervention	−0.219***	−0.219***	−0.224***	0.110***	0.111***	0.112***
	(0.022)	(0.022)	(0.022)	(0.017)	(0.016)	(0.017)
Constant	0.689***			0.262***		
	(0.088)			(0.068)		
Controls	Original	Original	Alternative	Original	Original	Alternative
N	12,160	12,160	11,920	12,160	12,160	11,920
No. Subjects	608	608	596	608	608	596
R^2	0.053	0.039	0.044	0.041	0.033	0.042

Notes: Columns (1) and (2) refer to Columns (6) and (2) in Table 2 of Altmann et al. (2022, p.2310), both reporting results of OLS regressions of intervention treatment differences in default adherence rate and choice quality in the decision task. Alongside results of robustness checks with two different specifications (Columns a and b, respectively) are reported. The first alternative specification (Columns a) show results of a logit regression with the same explanatory variables as in the original specification. The second specification (Columns b) also shows results of a logit regression, but with an alternative battery of control variables (see Table 1). The dependent variable in Columns (1)–(1b) is an indicator to 1 if a subject's choice in a given round of the experiment coincides with the (choice alternative corresponding to the) default option and 0 otherwise. The dependent variable of Columns (2)–(2b) is an indicator equal to 1 if a subject correctly solves the decision task in a given round of the experiment and 0 otherwise. For logit regressions average marginal effects are shown. Robust standard errors, clustered at the subject level, are reported in parentheses. ***, **, * indicates significance at the 1%, 5%, and 10% level, respectively. R^2 for logit regressions refers to McFadden's R^2.

Table 10: Cognitive spillovers
Robustness

	Lab experiment		
	(1)	(1a)	(1b)
Intervention	−0.038***	−0.037***	−0.037**
	(0.014)	(0.014)	(0.015)
Constant	0.519***		
	(0.064)		
Controls	Original	Original	Alternatve
N	12,160	12,160	11,920
No. Subjects	608	608	596
R^2	0.074	0.065	0.076

Notes: Column (1) refers to Column (2) in Table 3 Altmann et al. (2022, p. 2312), reporting results of a OLS regression of treatment differences in choice quality in the background task. Alongside results of robustness checks with two different specifications (Columns 1a and 1b) are reported. The dependent variable in in all columns is an indicator equal to 1 if a subject correctly solves the background task in a given round of the laboratory experiment and 0 otherwise. Column (1a) reports the results of a logit regression with the same explanatory variables as Column (1). Column (1b) reports the results of a logit regression with an alternative batteroy of control variables (see Table 1). For logit regressions average marginal effects are shown. Robust standard errors, clustered at the subject level, are reported in parentheses. ***, **, * indicates significance at the 1%, 5%, and 10% level, respectively. R^2 for logit regressions refers to McFadden's R^2.

Table 11: Effect of interventions on overall payoffs
Robustness

	(Payoff Share)	
	(1)	(1a)
Intervention	−0.000	0.000
	(0.011)	(0.011)
Constant	0.475***	0.514***
	(0.050)	(0.053)
Controls	Original	Alternative
N	12,160	11,920
No. Subjects	608	596
R^2	0.074	0.092

Notes: Column (1) refers to Column (2) in Table 4 of Altmann et al. (2022, p.2317) and reports results of OLS regressions of treatment differences in subjects' overall payoffs. Column (1a) reports results of an OLS regression with an alternative battery of control variables (see Table 1). The dependent variable is a subject's realized share of the maximally attainable overall payoff in the subject's treatment (calculated on the subject-round level). Robust standard errors, clustered at the subject level, are reported in parentheses. ***, **, * indicates significance at the 1%, 5%, and 10% level, respectively.

Table 12: Heterogeneity of Cognitive Spillovers
Robustness

	Raven Score			Task Difficulty		
	(1)	(1a)	(1b)	(2)	(2a)	(2b)
Intervention	−0.064***	−0.334***	−0.325***	−0.018	−0.103	−0.099
	(0.022)	(0.114)	(0.121)	(0.015)	(0.096)	(0.102)
Intervention × High Raven	0.071**	0.382**	0.367**			
	(0.032)	(0.174)	(0.181)			
Intervention × High Difficulty				−0.037**	−0.181**	−0.191**
				(0.015)	(0.090)	(0.090)
High Raven	−0.024	−0.127	−0.071			
	(0.025)	(0.142)	(0.145)			
High Difficulty				−0.073***	−0.441***	−0.435***
				(0.012)	(0.073)	(0.073)
Constant	0.505***	−0.049	0.268	0.560***	−0.297	0.221
	(0.063)	(0.340)	(0.363)	(0.064)	(0.350)	(0.358)
Controls	Original	Original	Alternative	Original	Original	Alternative
N	9,340	9,340	9,200	12,160	12,160	11,920
No. subjects	467	467	460	608	608	596
R^2	0.083	0.071	0.081	0.087	0.077	0.077

Note: Columns (1) and (2) refer to Column (2) in Table O.4 and Column (2) in Table O.5 in the Online Appendix of Altmann et al. (2022, p.5-6) respectively, reporting results of OLS regression of treatment differences in choice quality in the background task. Alongside results robustness checks with two different specifications (Columns a and b, respectively) are reported. The dependent variable in Columns (1)–(2b) is an indicator equal to 1 if a subject correctly solves the background task in a given round of the experiment and 0 otherwise. 'High Raven' is an indicator equal to 1 if a subject's Raven score lies above the median. 'High Difficulty' is an indicator equal to 1 if the tasks difficulty of the decision task in a given round of the experiment lies above the median difficulty. The first specification (Columns a) report results of a logit regression with the same explanatory variables as the original specification (Columns 1 and 2). The second specification (Columns b) report results of a logit regression with a different battery of control variables (see Table 1). Results of logit regression are reported in terms of coefficients. Robust standard errors, clustered at the subject level, are reported in parentheses. ***, **, * indicates significance at the 1%, 5%, and 10% level, respectively. R^2 for logit regressions refers to McFadden's R^2.

References

Altmann, S., Grunewald, A. and Radbruch, J.: 2022, Interventions and cognitive spillovers, *The Review of Economic Studies* **89**(5), 2293–2328.

Simonsohn, U., Simmons, J. P. and Nelson, L. D.: 2020, Specification curve analysis, *Nature Human Behaviour* **4**(11), 1208–1214.