ONLINE APPENDIX

A Additional Figures and Tables

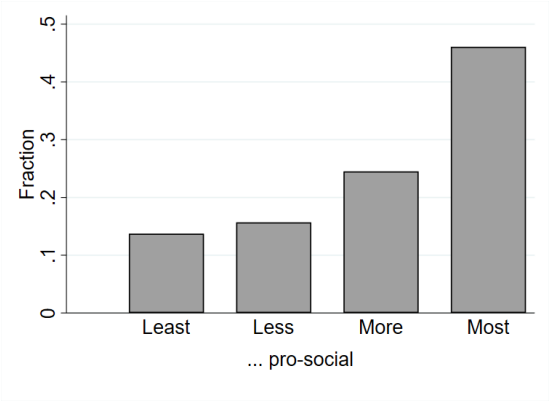
This appendix provides an additional figure and additional tables. Figure A1 provides a histogram of the pro-sociality index. Table A1 shows mean values for participants' well-being, separated both for the main aggregated WB categories as well as for the level of pro-sociality. As additional information it provides two-tailed p-values of the non-parametric Jonckheere's trend tests, an extension of the Wilcoxon rank sum test. Table A2 looks at all individual WB measures, dis-aggregating the four WB categories. It provides a more conservative test than before as it controls for multiple testing in relation to all 13 well-being measures (instead of four categories). Table A3a-d provide similar regressions to the ones of Table 3. In specification 1-4, only one of the main WB categories is included while specification 4 includes all four for comparison. Table A3a focuses on the pro-sociality index, while Table A3b-d look at the three decisions desperately (and we control for involved multiple testing). Table 3a shows that any potential effect of long-run HWB vanishes once other well-being measures are included. A more formal analysis reveals that any potential effect of long-run HWB is mediated to a large extent both by the change in short-run HWB (Goodman test: $p = 0.091$) and by EWB ($p = 0.026$). In contrast the coefficients of the change in short-run HWB and EWB stay relatively constant once other WB variables are included. We still find some evidence that EWB mediates some part of the effect of the change in short-run HWB ($p = 0.035$) and vice versa ($p = 0.068$). While there are some deviations (e.g. long-run HWB and trust), Table A3b-d broadly confirm the findings discussed before. Overall, these considerations provide further evidence for the importance of EWB. Table A3e provides similar regressions along the lines of Table 3 excluding (demographic) controls. Table A4 analyzes whether subjects scoring above the median of well-being measure are more likely to behave pro-socially and finds support. Finally, Table A5 and A6 provide analyses with respect to punishment

Table A1
Well-being categories according to how pro-social subjects behave - mean values

Well-being measures	DG and SPD					
	Least pro-social	**Less pro-social**	**More pro-social**	**Most pro-social**	**Jonckheere's Trend Test** adj. *p*-value	% SD
Hedonic Well-Being (HWB)						
Long-run HWB	-0.63	-0.45	-0.54	0.63	0.074^{+}	42%
Short-run HWB	0.15	-0.21	-0.24	0.15	0.495	0%
Change in hort-run HWB	-0.74	-0.39	0.03	0.33	0.074^{+}	61%
Eudaimonic Well-Being (EWB)	-1.32	-0.81	-0.31	0.83	0.002**	88%

Notes: $^{+}$/*/** indicates significance at the 10%/5%/1% level according to two-tailed Jonckheere's trend tests of the alternative hypothesis that there is a positive or negative trend between groups against the null hypothesis that there is no such trend. Those behaving pro-socially in all three (two/one/none) cooperation game decisions (giver, trustor, cooperator) are classified as most (more/less/least) pro-social. *p*-values are adjusted (Hochberg 1988) to account for the multiple testing implied by using four WB categories. % SD indicates how much *most pro-social* subjects are better (worse) off than *least pro-social* subjects, n=102.

Figure A1: Histogram – Pro-Sociality Index



Notes: Those behaving pro-socially in all three (two/one/none) cooperation game decisions (giver, trustor, cooperator) are classified as most (more/less/least) pro-social.

Table A2
Well-being according to how pro-social subjects behave - mean values

Well-being measures	DG and SPD					
	Least pro-social	**Less pro-social**	**More pro-social**	**Most pro-social**	**Jonckheere's Trend Test** (adj.) *p*-value	% SD
Hedonic Well-Being (HWB)						
Long-run HWB						
Overall Happiness (OH)	6.14	6.19	6.28	6.51	0.356	32%
Positive Affect Schedule (PAS)	31.79	33.25	32.92	34.87	0.148	52%
Negative Affect Schedule (NAS)	21.57	22.63	24.12	20.72	0.609	(11%)
Highest/ lowest happiness						
Highest Happiness (HH)	7.14	7.38	7.40	7.62	0.388	42%
Lowest Happiness (LH)	4.00	3.38	3.24	3.55	0.825	(31%)
Life satisfaction						
Satisfaction With Life (SWL)	24.79	23.13	24.04	24.53	0.665	(5%)
Short-run HWB						
Now Happiness (NH)	5.50	5.44	5.44	5.62	0.641	9%
Mood Index (MI)	43.07	39.56	39.20	42.06	0.763	(9%)
Change in short-run HWB						
Now Happiness Change (NHC)	-0.21	-0.44	0.08	0.26	0.318	36%
Mood Index Change (MIC)	-7.29	-2.25	-1.84	-0.23	0.264	72%
Eudaimonic Well-Being (EWB)						
PWB Index (PWBI)	28.14	26.69	27.36	29.19	0.148	27%
Self-Actualization Index (SAI)	40.57	41.69	42.24	44.68	0.031*	77%
Social Well-Being (SoWB)	52.93	58.88	60.96	63.00	0.014*	112%
Material well-being						
Monthly expenditures (MWB 1)	652.50	554.06	603.20	712.55	0.200	19%
Parents' income (MWB 2)	2.57	3.44	3.00	3.36	0.370	46%
Cognitive ability						
Cognitive Reflection Test (CRT)	1.21	1.50	1.68	1.70	0.190	44%
N	14	16	25	47		

Notes: $^{+}$/*/** indicates significance at the 10%/5%/1% level according to two-tailed Jonckheere's trend tests of the alternative hypothesis that there is a positive or negative trend between groups against the null hypothesis that there is no such trend. Those behaving pro-socially in all three (two/one/none) cooperation game decisions (giver, trustor, cooperator) are classified as most (more/less/least) pro-social. Due to the large number of 13 different well-being measures, the more lenient procedure of Benjamini and Hochberg (1995) is used to adjust *p*-values for the multiple. % SD indicates how much *most pro-social* subjects are better (worse) off than *least pro-social* subjects, n=102.

Table A3a
Ordered logit regressions for dictator game and seq. prisoner's dilemma with demographic controls

	(1) Pro-socialty Index	(2) Pro-socialty I.	(3) Pro-socialty I.	(4) Pro-socialty I.	(5) Pro-socialty I.
Long-run HWB	0.083				-0.035
	(0.108)				(0.116)
Short-run HWB		-0.136			0.019
		(0.153)			(0.173)
Change in short-run HWB			0.412**		0.362**
			(0.123)		(0.137)
EWB				0.403**	0.356**
				(0.119)	(0.134)
Expenditures	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parents' Income	0.064	0.048	0.091	0.163	0.189
	(0.165)	(0.166)	(0.166)	(0.177)	(0.183)
Cognitive Ability (CRT)	0.276	0.337	0.346	0.069	0.139
	(0.407)	(0.393)	(0.411)	(0.427)	(0.448)
Additional controls (& Cons.)	✓	✓	✓	✓	✓
N	102	102	102	102	102
Pseudo R^2	0.060	0.061	0.102	0.105	0.133

Notes: Robust standard errors in parentheses; $^{+}$/*/** indicates significance at the 10%/5%/1% level. WB measures are aggregated. Additional control variables as in Table 3.

Table A3b
Logit regressions for dictator game with demographic controls

	(1) Giver	(2) Giver	(3) Giver	(4) Giver	(5) Giver
Long-run HWB	0.014				-0.089
	(0.116)				(0.124)
Short-run HWB		0.053			0.107
		(0.151)			(0.176)
Change in short-run HWB			0.157		0.217
			(0.138)		(0.176)
EWB				0.396**	0.422*
				(0.135)	(0.153)
Expenditures	0.002	0.002	0.002^{+}	0.002	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parents' Income	0.120	0.120	0.138	0.239	0.260
	(0.201)	(0.198)	(0.205)	(0.225)	(0.224)
Cognitive Ability (CRT)	0.755	0.775	0.795	0.612	0.756
	(0.545)	(0.536)	(0.531)	(0.595)	(0.612)
Additional controls (& Cons.)	✓	✓	✓	✓	✓
N	102	102	102	102	102
Pseudo R^2	0.145	0.146	0.155	0.202	0.217

Notes: Robust standard errors in parentheses; $^{+}$/*/** indicates significance at the 10%/5%/1% level. WB measures are aggregated. Additional control variables as in Table 3.

Table A3c
Logit regressions for seq. prisoner's dilemma with demographic controls

	(1) Trustor	(2) Trustor	(3) Trustor	(4) Trustor	(5) Trustor
Long-run HWB	0.304				0.347*
	(0.157)				(0.150)
Short-run HWB		-0.288			-0.478*
		(0.206)			(0.204)
Change in short-run HWB			0.188		-0.123
			(0.140)		(0.203)
EWB				0.410*	0.473*
				(0.171)	(0.204)
Expenditures	0.001	0.001	0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parents' Income	-0.045	-0.081	-0.077	0.064	0.060
	(0.219)	(0.236)	(0.231)	(0.226)	(0.226)
Cognitive Ability (CRT)	-0.675	-0.365	-0.323	-0.734	-1.195^{+}
	(0.685)	(0.622)	(0.608)	(0.647)	(0.712)
Additional controls (& Cons.)	✓	✓	✓	✓	✓
N	102	102	102	102	102
Pseudo R^2	0.325	0.311	0.299	0.336	0.400

Notes: Robust standard errors in parentheses; $^{+}$/*/** indicates significance at the 10%/5%/1% level. WB measures are aggregated. Additional control variables as in Table 3.

Table A3d
Logit regressions for seq. prisoner's dilemma with demographic controls

	(1) Cooperator	(2) Cooperator	(3) Cooperator	(4) Cooperator	(5) Cooperator
Long-run HWB	0.029				-0.189
	(0.130)				(0.145)
Short-run HWB		0.111			0.339
		(0.165)			(0.213)
Change in short-run HWB			0.573*		0.600*
			(0.235)		(0.249)
EWB				0.662**	0.611**
				(0.175)	(0.185)
Expenditures	-0.000	-0.000	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Parents' Income	-0.059	-0.057	-0.092	0.117	0.085
	(0.187)	(0.187)	(0.182)	(0.212)	(0.202)
Cognitive Ability (CRT)	0.399	0.411	0.454	0.081	0.303
	(0.547)	(0.508)	(0.656)	(0.644)	(0.796)
Additional controls (& Cons.)	✓	✓	✓	✓	✓
N	102	102	102	102	102
Pseudo R^2	0.120	0.124	0.194	0.248	0.310

Notes: Robust standard errors in parentheses; +/*/** indicates significance at the 10%/5%/1% level. WB measures are aggregated. Additional control variables as in Table 3.

Table A3e
(Ordered) Logit regressions for dictator game and seq. prisoner's dilemma

	(1) **Pro-socialty Index**	(2) Giver Dummy	(3) Trustor Dummy	(4) Coop. Dummy
Long-run HWB	0.042	0.008	0.197	-0.090
	(0.106)	(0.112)	(0.099)	(0.131)
Short-run HWB	0.003	0.043	-0.270	0.131
	(0.152)	(0.142)	(0.170)	(0.193)
Change in short-run HWB	0.240*	0.094	0.064	0.389
	(0.122)	(0.141)	(0.150)	(0.184)
EWB	0.322**	0.305*	0.301*	0.410*
	(0.113)	(0.115)	(0.132)	(0.167)
N	102	102	102	102
Pseudo R^2	0.090	0.097	0.139	0.182

Notes: Ordered logit regression in (1). Logit regressions in (2)-(4). Robust standard errors in parentheses; +/*/** indicates significance at the 10%/5%/1% level. In specifications (2)-(4), *p*-values are adjusted (Hochberg 1988) to account for the multiple testing implied by looking at decisions separately.

Table A4
Proportion tests of those who score high and low on different WB scales

	long-run HWB	**Short-run HWB**	**Change in short-run HWB**	**EWB**
Pro-sociality Index = 3				
High	56.8%	50.9%	57.1%	58.8%
Low	35.2%	41.1%	35.8%	33.3%
Adj. *p*-value	0.062	0.320	0.062	0.039

Notes: For each aggregate WB measure, subjects are split into those who score at or above the median in terms of that measure (High) and those who are below the median (Low). For every variable the High/Low row shows the percentage of *most pro-social* subjects. Those behaving pro-socially in all three cooperation game decisions (giver, trustor, cooperator) are classified as most pro-social. *p*-values are provided by two-tail z-tests of the hypothesis that the percentage of the High group exceeds (or undercuts) the percentage of the Low group. *p*-values are adjusted (Hochberg 1988) to account for the multiple testing implied by using four WB categories. n=102.

Table A5
Logit regressions for SP- and TP-punishment games

	(1) SP-Punisher Dummy	(2) TP-Punisher Dummy *(low cost)*	(3) TP-Punisher Dummy *(high cost)*
Long-run HWB	-0.067	0.038	-0.088
	(0.114)	(0.127)	(0.127)
Short-run HWB	-0.057	-0.167	-0.140
	(0.124)	(0.139)	(0.175)
EWB	-0.041	0.166	0.316*
	(0.122)	(0.150)	(0.160)
Expenditures	0.000	-0.001	-0.000
	(0.001)	(0.001)	(0.001)
Parents' Income	-0.238	0.038	0.074
	(0.184)	(0.165)	(0.196)
Cognitive Ability (CRT)	0.121	-0.725	-0.370
	(0.472)	(0.544)	(0.611)
Additional controls (& Cons.)	✓	✓	✓
N	102	102	102
Pseudo R^2	0.073	0.126	0.122

Notes: Robust standard errors in parentheses; +/*/** indicate significance at the 10%/5%/1% level. WB measures are aggregated. Additional control variables as in Table 3.

Table A6
Well-being for the second-party [SP] (Mini-UG), low-cost third-party [TP] (SPD-P) and high-cost third-party (DG-P) punishment - mean values

	Well-being measure	**Nonpun.**	**Punisher**	**adjusted *p*-value**	**% SD**
	Hedonic Well-Being (HWB)				
	Long-run HWB	0.43	-0.39	0.324	(27%)
(a)	Short-run HWB	0.18	-0.16	0.636	(18%)
SP-Pun.					
(Mini-UG)	*Eudaimonic Well-Being (EWB)*	0.33	-0.39	0.324	(25%)
	N	48	54		
	Hedonic Well-Being (HWB)				
(b)	Long-run HWB	-0.01	0.01	0.883	1%
high-cost	Short-run HWB	0.09	-0.08	0.836	(8%)
TP-Pun.					
(SPD-P)	*Eudaimonic Well-Being (EWB)*	-0.20	0.19	0.836	15%
	N	50	52		
	Hedonic Well-Being (HWB)				
(c)	Long-run HWB	-0.09	0.29	0.682	13%
low-cost	Short-run HWB	0.03	-0.12	0.682	(7%)
TP-Pun.					
(DG-P)	*Eudaimonic Well-Being (EWB)*	-0.30	0.94	0.024*	50%
	N	77	25		

Notes: +/*/** indicate significance at the 10%/5%/1% level according to two-tailed non-parametric rank-sum tests. *p*-values are adjusted (Hochberg 1988) to account for the multiple testing implied by using four WB categories. % *SD* indicates how big the difference between those who punish and those who do not is in percent of the standard deviation. $n = 102$

B Appendix: Additional Material, Analyses and Robustness

This appendix provides four pieces of additional material. First, a detailed description of our experimental games as well as basic results about subjects' behavior in these games and their answers in the WB questionnaire and are presented. Due to space constraints, these results could not have been outlined in the main text. Second, robustness checks of the regression analysis of the main text are outlined. Third, completely unaggregated data is presented for completeness. Finally, a more elaborated analysis of the effects of pro-sociality on mood (pro-sociality → mood) is provided.

B.1 Games, Behavior and WB Data

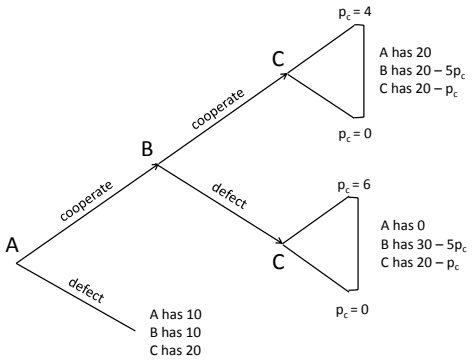
As indicated in the main text, participants played a total of five games. The sequential prisoner's dilemma (SPD) and the dictator game (DG) games measure pro-social behavior such as giving, trust, and cooperation. On the other hand, the mini-ultimatum game (Mini-UG), the SPD with punishment (SPD-P), and the DG with punishment (DG-P) games measure second-party and third-party punishment.

To measure giving and replicate the results of KE, I implemented a dictator game (DG). The dictator has an endowment of 20 € and can send 0 €, 2 €, ... , 20 € to the recipient who has an initial endowment of 0 € and who has to accept any choice the dictator makes. A dictator giving any positive amount will classified as a 'giver'.

To measure trust and cooperation, I used the *seq. prisoner's dilemma* (SPD) [or bilateral trust game], in a version proposed by Anderson et al. (2013). In this game, both players are initially endowed with 10 €. The first mover can either send 0 € or 10 € to the second mover, who then decides how much to return from the options of 0 €, 2 €, 4 €, 6 €, 8 €, or 10 €. The experimenter doubles the amount sent by either player. If both players send 10 €, they each end up with 20 €, which is the social optimum. However, the subgame perfect equilibrium with selfish preferences is for both players to send nothing, leading to a final payoff of 10 €. I classify a player who sends their endowment as a person who trusts, and follow Anderson et al. (2013) in classifying (conditional) cooperators.[1] After the game, I elicited unincentivized beliefs about what other people typically return on average, which I will use as a control in the regression analyses.

[1]There are three "pure" types: Second movers who always return 0 € independent of the first mover's choice, are classified as pure *free-riders*. Second movers who choose the most cooperative action available (always return 10 €) are classified as pure *unconditional cooperators*. Second movers who exactly return what has been sent to them by the first mover (0 € and 10 €) are classified as pure *conditional cooperators*. For subjects not behaving as one of the pure types, the Euclidean distance between his or her decision and the decision of each of the pure types is calculated. The subject is then assigned to the least distant type category. Since only observe 5% of cooperating subjects are classified as unconditional but 72% as conditional cooperators, both types are merged.

Figure B1: Seq. Prisoner's Dilemma with TP-Punishment



The *mini-ultimatum game* (Mini-UG) measures second-party (SP) punishment, is adapted from Falk et al. (2003), and has a two-stage structure. In the first stage, the proposer makes a proposal on how to divide a pie of 20 €. The two options are an unequal split where the proposer gets 18 € and the responder only receives 2 €, and an equal split.In the second stage, the responder can either accept or reject the proposer's choice. If the responder accepts, the chosen proposal is implemented. If the responder rejects the proposal, both players receive 0 €. Since even an egoistic proposer chooses the equal split when they believe that the unequal split will be rejected, I will focus on the punishment decision: a subject rejecting the equal split will be classified as a person punishing unfair behavior in a second-party (SP) punishment setting.[2]

The *SPD with punishment* (SPD-P) is an adaptation of a game by Hoff et al. (2011) and measures (low-cost) third-party punishment. The game proceeds as follows: the first mover, A, decides whether to send or keep their money. Next, the second mover, B, faces the same binary choice. Finally, the third mover, C, can choose to punish B by investing 1 € to reduce B's payoff by 5 €. The punishment is administered by an uninvolved third party, and it is relatively cheap. Therefore, punishing can improve relative payoffs in this setting. My focus is on the punishment decision of C. A subject who chooses to punish defectors will be classified as an individual punishing unfair behavior in a third-party (TP) punishment setting with low costs. Figure 1 provides a detailed game tree that shows the precise payoffs.

The *DG with punishment* (DG-P) is a variation of the dictator game that introduces a (high-cost) third-party (TP) punishment opportunity. In this game, the first, second, and third players receive an endowment of 15 €, 5 €, and 20 €, respectively. The first player can choose to send 0 €, 2.5 €, or 5 € to the second player. The second player has no decision to make, but the third player has the opportunity to invest 1 €, 2 €, 3 €, 4 €, or 5 € to punish the first player for unequal sharing. However, punishing is relatively costly, as investing 1 € in punishment only reduces the first player's payoff by 1 €. This

[2]After the Mini-UG, subjects played a joy-of-destruction game (JOY-G, Abbink and Herrmann 2011), measuring spite. I exclude this game from my main analysis since I almost observe no spite.

means that the punisher cannot improve his relative payoffs. I will classify a subject who punishes first players who do not share equally as an individual punishing unfair behavior in a TP punishment setting with high costs.

Besides the JOY-G (5% spite instead 26% in Abbink and Herrmann 2011), all other games replicate the conventional results of the literature: In the DG, 61% of subjects gave whereas 39% were *Nongivers* (KE's result: 40% *Nongivers*). In the SPD, 64% of the subjects trusted and 77% of the subjects are classified as (conditional) cooperators, and 23% as free-riders (Anderson et al. 2013 – *self-selected students* sample: 55% trust, 63% cooperators, 35% free-riders). In the Mini-UG, 53% of subjects rejected the unfair offer (Falk et al. 2013: 44% rejection). In the SPD-P, about 50% punished free-riders (whereas only 3% punished cooperators). In the DG-P, about 25% punished dictators who did not share equally. Results are roughly in line with Fehr and Fischbacher (2014).

Table B1 shows that WB measures are line with results from the literature. It shows the Spearman-rank correlation coefficients (and p-values) for HWB and EWB measures. All correlations have the expected sign and out of 55 correlations, 47 are significant at the 5% level. Moreover, Table B2 summarizes the Spearman-rank correlation of HWB and EWB measures with the Marlowe-Crowne social desirability scale, the two material WB measures, and cognitive ability. Compared to KE, the only real difference is that correlations between WB and the Marlowe-Crowne scale are slightly higher (at least for EWB), indicating that social desirability plays a more important role in a single-blind than in a double-blind setting (for which I control in the regressions). Table B3 provides clear evidence that long-run HWB and EWB are related in line with the idea that HWB emerges as a favorable by-product of EWB.

Finally, for EWB measures (and especially for the *Social Well-Being*), one may be concerned about particular items that ask something that is directly related to pro-social behavior. Overall, excluding these problematic questions from EWB seems not to have a systematic effect and results remain reasonable robust to conclude that my central findings do not depend on correlations between single questions and behavior in games. More precisely: *Social Well-Being* consists of five dimensions, for which two are potentially problematic: *social contribution* and *social acceptance*. Excluding the *social contribution* dimension (e.g. Question 4: "I have something valuable to give to the world" – App. D) from SoWB as wells as question 13 ("I do not feel responsible to help anybody.") from the Self-Actualization Index does not qualitatively change my main results – as presented in Figure 1 (related Table A1) and Table 3 – for pro-sociality. It slightly worsens results for punishing in the DG-P. Finally, for the trust decision, one may be concerned about questions hinting to what extent people consider others as trustworthy. Excluding the *social acceptance* dimension (e.g. Question 14: "I believe that other people are kind."), question 3 ("I believe that people are essentially good and can be trusted") from the Self-Actualization Index, and question 16 ("I have not experienced many warm and trusting

Table B1
Spearman correlation matrix for EWB and HWB measures (correlations with coefficients and p-values)

	OH	PAS	NAS	NH1	MI1	HH	LH	SWL	PWBI	SAI
Overall Happiness (OH)										
Positive Affect Schedule (PAS)	0.39									
	0.000									
Negative Affect Schedule (NAS)	-0.14	-0.21								
	0.172	**0.037**								
Now Happiness (NH)	0.35	0.25	-0.24							
	0.000	**0.012**	**0.016**							
Mood Index (MI)	0.38	0.38	-0.31	0.64						
	0.000	**0.000**	**0.002**	**0.000**						
Highest Happiness (HH)	0.40	0.40	-0.13	0.56	0.45					
	0.000	**0.000**	0.204	**0.000**	**0.000**					
Lowest Happiness (LH)	0.28	0.25	-0.45	0.34	0.51	0.32				
	0.005	0.140	**0.000**	**0.000**	**0.000**	**0.001**				
Satisfaction with Life (SWL)	0.51	0.29	-0.25	0.33	0.27	0.33	0.19			
	0.000	**0.004**	**0.010**	**0.001**	**0.005**	**0.001**	0.059			
Psych. Well-Being Index (PWBI)	0.45	0.46	-0.31	0.22	0.37	0.28	0.22	0.56		
	0.000	**0.000**	**0.002**	**0.025**	**0.000**	**0.004**	**0.029**	**0.000**		
Self-Actualization Index (SAI)	0.16	0.32	-0.43	0.10	0.11	0.11	0.24	0.32	0.44	
	0.114	**0.001**	**0.000**	0.327	0.293	0.264	**0.014**	**0.001**	**0.000**	
Social Well-Being (SoWB)	0.32	0.31	-0.30	0.15	0.23	0.22	0.23	0.25	0.46	0.48
	0.001	**0.002**	**0.002**	0.127	**0.021**	**0.023**	**0.022**	**0.010**	**0.000**	**0.000**

Notes: *p*-values below the 5% level are in bold type, n=102.

Table B2
Spearman correlation matrix for EWB and HWB with MC social desirability scale, material WB, cognitive ability (coefficients and p-values)

	OH	PAS	NAS	NH1	MI1	HH	LH	SWL	PWBI	SAI	SoWB
Marlowe-Crowne scale (MC)	0.13	0.33	-0.36	0.17	0.21	0.15	0.28	0.15	0.33	0.42	0.24
	0.187	**0.001**	**0.000**	0.090	**0.032**	0.142	**0.004**	0.142	**0.001**	**0.000**	**0.014**
Monthly expenditures	0.27	0.24	0.04	0.02	0.13	0.05	0.05	0.19	0.17	0.22	0.06
	0.006	**0.016**	0.711	0.846	0.190	0.587	0.643	0.062	0.079	**0.029**	0.572
Parents' income	0.17	0.15	-0.07	0.14	0.10	0.24	-0.10	0.38	0.20	-0.02	-0.02
	0.085	0.135	0.482	0.166	0.338	**0.015**	0.333	**0.000**	**0.049**	0.873	0.868
Cognitive Ability (CRT)	0.11	0.03	-0.01	-0.08	-0.04	0.00	-0.12	0.06	0.09	0.02	-0.04
	0.291	0.736	0.937	0.421	0.721	0.979	0.218	0.573	0.379	0.808	0.716

Notes: *p*-values below the 5% level are in bold type, n=102.

Table B3
Results on EWB and HWB

		Mean scores OH	PAS	NAS	HH
EWB measure					
Index of PWB (PWBI)	High	6.69	35.78	20.66	7.69
	Low	5.91	31.00	23.70	7.16
	p-value	**0.00**	**0.00**	**0.02**	**0.01**
Self-Actualization Index (SAI)	High	6.47	34.92	19.64	7.53
	Low	6.22	32.41	24.49	7.39
	p-value	0.14	**0.02**	**0.00**	0.27
Social Well-Being (SoWB)	High	6.64	35.02	20.89	7.73
	Low	6.00	32.13	23.28	7.13
	p-value	**0.00**	**0.01**	0.06	**0.00**

Notes: Subjects are split into those who score at or above the median of different EWB measures and those who are below. For each group mean values of different HWB measures (OH, PAS, NAS, SWL) are reported. Additionally, p-values for the null hypothesis that the High EWB group is better off than the Low group are reported. All p-values below 5% are in bold type, n=102.

relationships with others") from the PWB Index does not lead to a qualitative change of my results concerning trusting.

B.2 Robustness checks – Regression Analysis

This section provides three additional robustness checks for the regression analysis. First, I will check whether results are robust to including personality characteristics. Second, an analysis that further controls for the multicollinearity problem is provided. Finally, an unaggregated analysis is presented.

As outlined before, subjects also answered a short Big Five Inventory. Because well-being measures and some Big Five measures have a fairly high correlation, they are not included in the regression analysis of the main text to avoid further amplifying the problem of multicollinearity. Table B4a-b show, however, that including the Big Five measure leads to fairly similar results.[3] Only the change in short-run HWB is – unlike in Table 3 – not significant any more for the pro-sociality index (specification 1). More importantly, EWB remains significant in all specifications and the change in short-run HWB remains at least significant in specification 4 when looking at the cooperation decision. Thus, results generally remain robust when controlling for personality measures.

As noted before, a crucial concern with any regression analysis in my context is multicollinearity. Individual WB measures are highly correlated with each other (see App. B.1). Aggregating WB measures already reduces this problem since it reduces the number of variables. Nonetheless, it does not completely resolve the problem since aggregate measures also remain correlated. To fully avoid this issue, KE regress one WB measure

[3]Notably, as in Table 4a, the coefficients' p-values are adjusted in Table B4a to take into account that we look at the three decisions separately.

Table B4a
(Ordered) Logit regressions for DG and SPD with demographic controls & Big Five Inventory

	(1) Pro-Sociality Index	(2) Giver Dummy	(3) Trustor Dummy	Coop. Dummy
Long-run HWB	**-0.010**	-0.164	0.419	-0.205
	(0.137)	(0.148)	(0.232)	(0.185)
Short-run HWB	**-0.030**	0.119	-0.690*	0.323
	(0.214)	(0.181)	(0.317)	(0.284)
Change in short-run HWB	**0.240**	0.066	-0.100	0.692*
	(0.173)	(0.190)	(0.301)	(0.256)
EWB	**0.417****	0.418*	0.474*	0.651*
	(0.161)	(0.177)	(0.227)	(0.241)
Expenditures	**0.000**	0.002	-0.000	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)
Parents' Income	**0.285**	0.320	0.222	0.284
	(0.184)	(0.213)	(0.204)	(0.170)
Cognitive Ability (CRT)	**0.165**	0.818	-1.682	0.734
	(0.486)	(0.642)	(0.890)	(0.725)
Extraversion	**0.088**	0.100	0.155	0.143
	(0.117)	(0.124)	(0.268)	(0.228)
Agreeableness	**0.136**	0.165	-0.041	0.041
	(0.133)	(0.151)	(0.297)	(0.196)
Conscientiousness	**-0.036**	-0.058	0.134	-0.053
	(0.113)	(0.122)	(0.174)	(0.146)
Neuroticism	**-0.117**	0.072	-0.086	-0.310^{+}
	(0.119)	(0.137)	(0.165)	(0.170)
Openness	**0.384****	0.344*	0.754*	0.641*
	(0.135)	(0.172)	(0.294)	(0.263)
Additional controls (& Cons.)	✓	✓	✓	✓
N	**102**	102	102	102
Pseud R^2	**0.185**	0.284	0.515	0.406

Notes: Ordered logit regression in (1). Logit regressions in (2)-(4). Robust standard errors in parentheses; $^{+}$/*/** indicates significance at the 10%/5%/1% level. In specifications (2)-(4), *p*-values are adjusted (Hochberg 1988) to account for the multiple testing implied by looking at decisions separately. Well-being measures are aggregated. Additional control variables similar to Table 3.

Table B4b
Logit regressions for Mini-UG, SPD-P and DG-P with demographic controls & Big Five Inventory

	(1) SP-Punisher Dummy	(2) TP-Punisher Dummy *(low cost)*	(3) TP-Punisher Dummy *(high cost)*
Long-run HWB	-0.060	0.022	-0.119
	(0.120)	(0.136)	(0.142)
Short-run HWB	-0.064	-0.158	-0.086
	(0.135)	(0.158)	(0.187)
EWB	-0.008	0.119	0.396*
	(0.129)	(0.152)	(0.190)
Expenditures	0.000	-0.001	0.000
	(0.001)	(0.001)	(0.001)
Parents' Income	-0.265	0.085	0.108
	(0.198)	(0.173)	(0.221)
Cognitive Ability (CRT)	0.234	-0.822	-0.243
	(0.485)	(0.556)	(0.635)
Extraversion	0.004	0.064	-0.106
	(0.101)	(0.103)	(0.134)
Agreeableness	-0.058	-0.036	-0.288
	(0.140)	(0.154)	(0.207)
Conscientiousness	-0.131	0.199	0.001
	(0.111)	(0.125)	(0.112)
Neuroticism	-0.054	0.008	-0.045
	(0.131)	(0.125)	(0.149)
Openness	0.100	-0.032	0.247
	(0.137)	(0.140)	(0.169)
Big Five Inventory	✓	✓	✓
Additional controls (& Cons.)	✓	✓	✓
N	102	102	102
Pseudo R^2	0.093	0.154	0.156

Notes: Robust standard errors in parentheses;$^{+}$/*/** indicates significance at the 10%/5%/1% level. Well-being measures are aggregated. Additional control variables similar to Table 3.

Table B5a
Summary of (ordered) logit regression results for genuienly pro-social behavior

	(1) Pro-Sociality Index	SE	(2) Giver Dummy	SE	(3) Trustor Dummy	SE	(4) Cooperator Dummy	SE
Hedonic Well-Being (HWB)								
Long-run HWB	0.015	(0.072)	-0.024	(0.088)	0.171	(0.104)	-0.010	(0.089)
Short-run HWB	-0.064	(0.109)	0.051	(0.120)	-0.113	(0.162)	-0.002	(0.134)
Change in short-run HWB	**0.387***	(0.132)	0.074	(0.121)	0.152	(0.140)	0.562	(0.218)
Eudaimonic Well-Being (EWB)								
EWB	**0.258***	(0.099)	0.213	(0.113)	0.256	(0.138)	**0.394***	(0.134)

Notes: Ordered logit regression in (1). Logit regressions in (2)-(4). Robust standard errors in parentheses. $^{+}$/*/** indicates significance at the 10%/5%/1% level. Each coefficient represents one regression, in which the dummy of pro-social behavior is the dependent variable. *p*-values are adjusted (Hochberg 1988) to account for the multiple testing implied by using four WB categories and for analyzing individual decisions separately (2)-(4). WB measures are aggregated and all regressions include control variables similar to Table 3. All coefficients with *p*-values below 5 % are in bold type, n = 102.

Table B5b
Summary of logit regression results for punishment

	(1) SP-Punisher Dummy	SE	(2) TP-Punisher Dummy *(low cost)*	SE	(3) TP-Punisher Dummy *high cost)*	SE
Hedonic Well-Being (HWB)						
Long-run HWB	-0.074	(0.083)	0.004	(0.084)	-0.002	(0.088)
short-run HWB	-0.088	(0.115)	-0.082	(0.117)	-0.058	(0.149)
Eudaimonic Well-Being						
EWB	-0.090	(0.103)	0.095	(0.113)	0.204	(0.148)

Notes: Robust standard errors in parentheses. $^{+}$/*/** indicates significance at the 10%/5%/1% level. Each coefficient represents one regression, in which the (SP- or TP-) punishment dummy is the dependent variable. p-values are adjusted (Hochberg 1988) to account for the multiple testing implied by using three WB categories. WB measures are aggregated and all regressions include control variables similar to Table 3. All coefficients with *p*-values below 5 % are in bold type, n = 102.

on one dummy of social behavior using ordered logit. To follow the regression strategy of Table 3, I deviate from this approach by regressing one dummy of social behavior (or the pro-sociality index) on one WB measure. Noticeably, KE's approach could be considered as the more conservative since WB measures are treated as ordinal variables. Crucially, following their approach leads to qualitatively identical results.

Table B5a-b provides a summary of these (ordered) logit results, including similar controls as Table 3. Crucially, every coefficient represents a single regression, and coefficients of the control variables are not shown due to space constraints. Notably, the coefficients' *p*-values are adjusted for the fact that a single regression is run for each WB category by using Hochberg procedure. Moreover, in Table B5a, this correction additionally incorporates that we look at three different situations. The following three main results emerge from the tables: First, while the EWB coefficient is only significant in one of the individual decisions, it remains positive for all three individual decisions. One has to bear in mind that Table B5a provides a very stringent test as it controls both for multiple testing due to different WB concepts and three different decisions. Only the most pronounced relationship between EWB and a pattern of pro-social behavior (cooperation) remains significant. Importantly, the main specification 1 still fully supports the idea of a positive connection between EWB and pro-sociality. Second, Table B5a still provides some support for the idea that cooperating subjects observe an increase in their short-run

HWB. At least in the main specification, the effect remains significant. Third, long-run HWB clearly has a weaker relationship with pro-sociality than EWB. In Table B5a, no significance at all is observed. If there is a relationship at all, it is with trusting. Overall, this analysis suggest that – despite slightly weaker results for individual decisions (due to very stringent testing) – the observations of the main text remain reasonably robust.

In sum, the message of the first two robustness checks seems to be the following: There is only one relationship between WB and pro-sociality that remains – despite slightly weaker results for individual decisions – reasonably robust: the one with EWB. In addition, there is some *game* and *decision specific* evidence that cooperating increases one's mood, at least in relative terms. Finally, if anything, long-run HWB might be related to trust, but this appendix casts even doubt on that.

Unaggregated Analysis

To further illuminate why specific *aggregate* WB concepts, especially EWB, are correlated with pro-social behavior and to provide an additional robustness check, Table B6a-b provide regression results with non-aggregated WB measures in a similar fashion to Table B5a-b. Notably, I do not control for multiple testing in the following analysis as I do not have the power to do so. These results should therefore be treated with caution. Only consistently significant results across games should be treated as valid. In the tables, every coefficient again constitutes one regression and coefficients of control variables are dropped due to space constraints. Regressing WB measures on dummies of cooperation using ordered logit leads to very similar results.[4]

There are two main messages from this analysis of Table B6a-b: First, the significance of the aggregate EWB measure is mainly driven by the *Self Actualization Index* and *Social Well-Being*, not the *PWB Index*. This reflects that the latter measure already showed weaker results in the main text compared to KE. This may be explainable insofar as I only used a single-blind instead of KE's double-blind procedure (although it is unclear why only one measure is affected). Second, some aspects of pro-social behavior seem more related to the individual component of EWB (interestingly TP-punishment), others to the social component of EWB (Trust), and some to both components (Generosity, Cooperation).

Interestingly, the significance of the (aggregate) *change in short-run HWB* for the cooperating decision seems to be driven by both aspects of HWB: the affective and the cognitive-evaluative component. Both the *Mood Index Change* and the *Now Happiness Change* are significant. For long-run HWB measures, only for the punishment settings, two variables seem to be significant: Rejecting in the mini-ultimatum game seems to be

[4]Notably, a regression analysis similar to Table 3 that includes all *unaggregated* WB measures in one regression does not lead to similar – or at all plausible – results as in the main text. Using unaggregated WB measures largely intensifies the multicollinearity problem, impeding a meaningful analysis.

Table B6a
Summary of logit regression results for genuinely pro-social behavior

	(1) Giver Dummy	SD	(2) Trustor Dummy	SD	(3) Cooperator Dummy	SD
Hedonic Well-Being (HWB)						
Long-run HWB						
Overall Happiness (OH)	0.042	(0.241)	0.341	(0.224)	-0.119	(0.235)
Positive Affect Schedule (PAS)	0.020	(0.043)	0.064	(0.063)	0.026	(0.051)
Negative Affect Schedule (NAS)	0.027	(0.035)	-0.061	(0.040)	-0.001	(0.036)
Highest/ Lowest happiness						
Highest Happiness (HH)	0.217	(0.205)	0.223	(0.250)	0.006	(0.239)
Lowest Happiness (LH)	-0.204	(0.203)	**0.503**$^{+}$	(0.258)	**-0.463***	(0.224)
Life satisfaction						
Satisfaction with Life (SWL)	-0.021	(0.054)	-0.041	(0.047)	-0.012	(0.058)
Short-run HWB/ Mood						
Now Happiness (NH)	0.051	(0.162)	-0.027	(0.037)	0.159	(0.192)
Mood Index (MI)	0.004	(0.022)	-0.019	(0.029)	-0.018	(0.027)
Now Happiness Change (NHC)	0.113	(0.200)	0.086	(0.230)	**0.420**$^{+}$	(0.233)
Mood Index Change (MIC)	0.019	(0.027)	0.028	(0.030)	**0.146***	(0.053)
Eudaimonic Well-Being (EWB)						
PWB Index (PWBI)	0.028	(0.073)	**0.171**$^{+}$	(0.090)	0.000	(0.073)
Self-Actualization Index (SAI)	**0.109**$^{+}$	(0.063)	0.049	(0.056)	**0.200***	(0.081)
Social Well-Being (SoWB)	**0.062**$^{+}$	(0.033)	**0.072***	(0.035)	**0.123***	(0.041)

Notes: Robust standard errors in parentheses. $^{+}$/*/** indicate significance at 10/5/1% level. Each coefficient represents one regression, in which the dummy of pro-social behavior is the dependent variable. All regressions include control variables used in Table 3. All coefficients with p-values below 10 % are in bold type, n = 102.

Table B6b
Summary of logit regression results for punishment

	(1) SP-Punisher Dummy	SD	(2) TP-Punisher Dummy *(low cost)*	SD	(3) TP-Punisher Dummy *high cost)*	SD
Hedonic Well-Being						
Long-run HWB						
Overall Happiness (OH)	0.062	(0.201)	-0.031	(0.227)	-0.155	(0.222)
Positive Affect Schedule (PAS)	**-0.105***	(0.041)	0.072	(0.047)	**0.081**$^{+}$	(0.047)
Negative Affect Schedule (NAS)	0.035	(0.034)	0.054	(0.036)	0.042	(0.035)
Highest/ Lowest happiness						
Highest Happiness (HH)	-0.134	(0.205)	0.095	(0.205)	0.083	(0.243)
Lowest Happiness (LH)	0.175	(0.170)	**-0.410***	(0.207)	-0.196	(0.233)
Life satisfaction						
Satisfaction with Life (SWL)	-0.043	(0.046)	-0.005	(0.044)	0.005	(0.055)
Short-run HWB/ Mood						
Now Happiness (NH)	-0.155	(0.157)	-0.212	(0.162)	-0.116	(0.199)
Mood Index (MI)	-0.016	(0.022)	-0.015	(0.022)	-0.005	(0.027)
Eudaimonic Well-Being (EWB)						
PWB Index (PWBI)	-0.101	(0.067)	-0.028	(0.070)	0.045	(0.079)
Self-Actualization Index (SAI)	-0.017	(0.045)	0.028	(0.050)	**0.131***	(0.059)
Social Well-Being (SoWB)	-0.022	(0.027)	0.023	(0.028)	0.023	(0.030)

Notes: Robust standard errors in parentheses. $^{+}$/*/** indicate significance at 10/5/1% level. Each coefficient represents one regression, in which the (SP- or TP-) punishment dummy is the dependent variable. All regressions include control variables used in Table 3. All coefficients with p-values below 10 % are in bold type, n = 102.

negatively related to long-run positive affect: *Bradburn's Positive Affect* and the *Positive Affect Schedule*. Hence, in this decision, only the affective not the cognitive-evaluative component of HWB matters.

Overall, however, the central findings of the analysis in the main text are supported by the unaggregated analysis. The most robust correlation emerges between pro-social behavior and EWB (abstracting from the weakness of the PWB Index - PWBI) as well as with the change in short-run HWB or mood after the cooperating decision.

B.3 Individual Data

This section presents a completely un-aggregated analysis, showing all individual well-being measures as well as the six individual choices in our games. Table B7a focuses on pro-social behavior while Table B7b refers to punishment. As in the last of the previous section, I dispense from controlling for multiple testing as I do not have sufficient power for it. In this sense, results should be treated with care. I am not interested whether one individual WB is significant for one particular choice. The question is whether some measures are consistently significant across settings. As previous results indicated this mainly true for measures of eudaimonic WB and in particular the Self-Actualization Index and Social Well-Being. In addition, similar to previous results, there is an indication that long-run well-being could be related to trust. If anything, punishing in the Mini-UG – i.e. second-party punishment – seems to be negatively related to WB measures.

B.4 Additional Analysis of the *short-run HWB Hyp.*

This section presents an explorative analysis of why behaving egoistically in the seq. prisoner's dilemma seems to lead to a (relative) decrease in mood.[5] Both short-run HWB items, *Now Happiness* and the *Mood Index* are only measured before and after the seq. prisoner's dilemma when subjects have made *both* decisions in this game. Nonetheless, only the *Cooperator* dummy is significant in Table 3 but not the *Trustor* dummy. Table B8 provides a hint why this might be the case. It separates both individual measures by different classifications as outlined in the table. In line with results from Table 3, there is no significant difference between Givers and Nongivers. In addition, as expected by previous results, the decrease in short-run HWB is stronger for *Free-riders* (3) than for *Nontrustors* (2). Crucially, those who neither trust nor free-ride (4) experience a similar decrease than *Free-riders* hinting at a potential explanation: Many of the *Nontrustors* cooperated (18 of 37), but only very few *Free-riders* trusted (4 out 23), implying that predominantly *Free-riders* did not behave pro-social at all.

There are two plausible mechanisms that I cannot distinguish: Either not behaving pro-socially at all makes people unhappy or free-riding has much worse effects than not trusting. Becchetti and Antoni (2010) find evidence in line with the first idea as they observe that using payoff enhancing power has an effect. In my game, sent money is doubled for both the first and the second mover. Not using this efficiency-enhancing power in both situations might decrease short-run HWB. However, it might still be true that free-riding itself makes people unhappy because returning a favor (cooperating) might

[5] Unlike in the main text, we do not control for multiple testing in this section. The idea of this section is not to establish a connection between the change in short-run HWB and pro-sociality. This section explores – under the assumption that there is a connection – why this connection might be game and decision specific

be considered to be a social norm whereas trusting might not.[6]

[6]Finally, it is possible to split up the the data provided by Table B8 into the two sequences of games implemented in the experiment. Either the dictator game or the seq. prisoner's dilemma was played first. This analysis reveals that the observed differences with respect to change of short-run HWB between pro-socially and egoistically behaving subjects are slightly more pronounced when the SPD is played as a second game. One plausible explanation for this observation is that playing the DG first highlights the efficiency-enhancing feature of the SPD in line with the idea of Becchetti and Antoni (2010).

Table B7a
Well-being for the dictator game and the sequential prisoner's dilemma - mean values

Well-being measures	**1) Dictator game**				**2) Seq. prisoner's dilemma - first mover**				**3) Seq. prisoner's dilemma - second mover**			
	Nongivers	**Givers**	p-value	% SD	**Nontrustors**	**Trustors**	p-value	% SD	**Free-riders**	**Cooperators**	p-value	% SD
Hedonic Well-Being (HWB)												
Long-run HWB												
Overall Happiness (OH)	6.21	6.44	0.15	21%	6.05	6.52	**0.02***	41%	6.39	6.34	0.57	4%
Positive Affect Schedule (PAS)	32.54	34.44	**0.06**$^{+}$	32%	32.46	34.43	**0.05**$^{+}$	33%	32.70	34.01	0.17	22%
Negative Affect Schedule (NAS)	21.69	22.14	0.61	6%	23.76	20.95	**0.04***	37%	22.09	21.94	0.46	2%
Highest/ lowest happiness												
Highest Happiness (HH)	7.26	7.59	**0.08**$^{+}$	29%	7.27	7.57	0.10	26%	7.35	7.49	0.29	13%
Lowest Happiness (LH)	3.67	3.41	0.80	17%	3.35	3.60	0.20	17%	3.91	3.39	0.93	36%
Life satisfaction												
Satisfaction With Life (SWL)	23.92	24.41	0.33	9%	23.89	24.42	0.32	10%	24.57	24.13	0.63	8%
Short-run HWB/ Mood												
Now Happiness (NH)	5.41	5.59	0.25	13%	5.59	5.65	0.43	4%	5.61	5.63	0.47	2%
Mood Index (MI)	39.41	40.63	0.29	11%	41.43	39.66	0.79	16%	41.83	39.86	0.77	18%
Now Happiness Change (NHC)	0.00	0.10	0.35	8%	-0.16	0.06	0.16	20%	-0.43	0.10	**0.02***	48%
Mood Index Change (MIC)	-1.26	-0.94	0.41	4%	-3.00	0.37	**0.02***	42%	-5.48	0.49	**0.00****	75%
Eudaimonic Well-Being (EWB)												
PWB Index (PWBI)	27.46	28.67	**0.07**$^{+}$	31%	27.16	28.80	**0.02***	42%	28.00	28.27	0.39	7%
Self-Actualization Index (SAI)	41.54	43.98	**0.01***	46%	41.51	43.92	**0.01***	45%	40.78	43.71	**0.01***	55%
Social Well-Being (SoWB)	58.03	61.98	**0.01***	44%	56.57	62.69	**0.00****	68%	55.43	61.94	**0.00****	73%
Material well-being												
Monthly expenditures (MWB 1)	595.00	688.33	**0.07**$^{+}$	30%	613.51	674.92	0.17	20%	622.17	661.52	0.29	13%
Parents' income (MWB 2)	2.85	3.38	**0.06**$^{+}$	31%	3.03	3.26	0.25	14%	3.04	3.22	0.33	10%
Cognitive ability												
Cognitive Reflection Test (CRT)	1.33	1.76	**0.03***	38%	1.59	1.60	0.49	0%	1.30	1.68	**0.08**$^{+}$	34%
N	39	63			37	65			23	79		

Notes: $^{+}$/*/** indicates significance at the 10/5/1% level according to t-tests of the null hypothesis that subjects behaving pro-socially are better off than non-pro-social subjects (e.g. that they have higher positive affect, lower negative effect, higher income, etc.). % *SD* indicates how big the difference between those who show pro-social behavior and those who do not is in percent of the standard deviation. All *p*-values below 10% are in bold type, n=102.

Table B7b
Well-being for the Mini-Ultimatum Game (Mini-UG), the Sequential Prisoner's Dilemma (SPD-P) and the Dictator Game (DG-P) with punishment - mean values

Well-being measures	**1) Mini-UG**				**2) SPD-P**				**3) DG-P**			
	Nonpun.	**SP-Punisher**	p-value	% SD	**Nonpun.**	**TP-Punisher** (low cost)	p-value	% SD	**Nonpun.**	**TP-Punisher** (high cost)	p-value	% SD
Hedonic Well-Being (HWB)												
Long-run HWB												
Overall Happiness (OH)	6.35	6.35	0.50	0%	6.36	6.35	0.52	1%	6.35	6.36	0.49	1%
Positive Affect Schedule (PAS)	35.33	32.28	**0.00****	52%	32.84	34.56	**0.07**$^{+}$	29%	33.16	35.44	**0.05***	39%
Negative Affect Schedule (NAS)	20.73	23.07	**0.06**$^{+}$	31%	21.06	22.85	0.88	24%	21.95	22.04	0.52	1%
Highest/ lowest happiness												
Highest Happiness (HH)	7.58	7.35	0.15	20%	7.38	7.54	0.24	14%	7.42	7.60	0.24	16%
Lowest Happiness (LH)	3.46	3.56	0.63	7%	3.72	3.31	0.92	28%	3.51	3.52	0.48	1%
Life satisfaction												
Satisfaction With Life (SWL)	25.00	23.54	**0.09**$^{+}$	27%	24.20	24.25	0.48	1%	24.13	24.52	0.37	7%
Short-run HWB/ Mood				0%								
Now Happiness (NH)	5.71	5.44	0.17	19%	5.66	5.48	0.74	13%	5.61	5.44	0.70	12%
Mood Index (MI)	40.10	38.39	0.20	17%	39.48	38.92	0.61	5%	39.30	38.88	0.57	4%
Eudaimonic Well-Being (EWB)												
PWB Index (PWBI)	28.90	27.59	**0.05***	33%	28.24	28.17	0.53	2%	27.92	29.08	0.10	30%
Self-Actualization Index (SAI)	43.42	42.72	0.26	13%	42.72	43.37	0.27	12%	42.21	45.64	**0.00****	64%
Social Well-Being (SoWB)	61.25	59.78	0.20	16%	59.10	61.79	**0.07**$^{+}$	30%	59.77	62.64	**0.08**$^{+}$	32%
Material well-being												
Monthly expenditures (MWB 1)	652.60	652.69	0.50	0%	670.50	635.48	0.71	11%	644.29	678.40	0.32	11%
Parents' income (MWB 2)	3.46	2.93	**0.06**$^{+}$	31%	3.30	3.06	0.76	14%	3.16	3.24	0.42	5%
Cognitive ability												
Cognitive Reflection Test (CRT)	1.67	1.54	0.28	12%	1.66	1.54	0.71	11%	1.62	1.52	0.65	9%
N	48	54			50	52			77	25		

Notes: $^{+}$/*/** indicates significance at the 10/5/1% level according to t-tests of the null hypothesis that TP-Punisher are better off (e.g. that they have higher positive affect, lower negative effect, higher income, etc.) whereas SP-Punisher are worse off. % *SD* indicates how big the difference between those who punish and those who do not is in percent of the standard deviation. All *p*-values below 10% are in bold type, n=102.

Table B8
Now Happiness Change and *Mood Index Change* – Dictator Game and Seq. Prisoner's Dilemma

1) Dictator Game

	Nongivers	Givers	T-Test
Now Happiness Change	0.00	0.10	0.35
Mood Index Change	-1.26	-0.94	0.42

2) Seq. Prisoner's Dilemma

	Nontrustors	Trustors	T-Test
Now Happiness Change	-0.16	0.06	0.16
Mood Index Change	-3.00	0.37	**0.02**

3) Seq. Prisoner's Dilemma

	Free-riders	Cooperators	T-Test
Now Happiness Change	-0.43	0.10	**0.02**
Mood Index Change	-5.48	0.49	**0.00**

4) Seq. Prisoner's Dilemma

	Nontrustors & Free-riders	Trustors or Cooperators	T-Test
Now Happiness Change	-0.42	0.07	**0.04**
Mood Index Change	-5.95	0.31	**0.00**

Notes: The T-Test column provides p-values according to t-tests of the null hypothesis that subjects who behave pro-socially experience a higher NHC and MIC. All p-values below 5% are in bold type, n=102.

C Instructions

This section provides the instructions for all six games (DG, SPD, JOY-G, Mini-UG, SPD-P, DG-P), translated from the German original. Additionally, for the more complicated games (SPD, SPD-P, DG-P), control questions were asked that are provided at the end of this section.

C.1 Instructions - DG (Dictator Game)

In this section of the experiment the **situation** is as follows:
At the beginning, Person A gets 20 EUR. Person B gets 0 EUR. Then, Person A has the opportunity to send money to Person B. However, Person A is not obliged to send money to Person B.

The Details:

So **Person A** gets **20 EUR** and has the following two alternatives: Sending no money to B (**0 EUR**) or sending money to B (either **2, 4, 6, 8, 10, 12, 14, 16, 18, or 20 EUR**).

Person B gets **0 EUR** at the beginning and knows that Person A has the described alternatives. Although B will be informed about A's decision, B does not make any decision him- or herself and has to accept A's decision.

Hence, we have the following **payoffs**:
For **Person A**: 20 EUR - money send to B
For **Person B**: Money send from B

Four examples:

1. If A sends **0 EUR**, A's payoff is 20 EUR and B' payoff is 0 EUR.
2. If A sends **10 EUR**, A's payoff is 10 EUR and B' payoff is 10 EUR.
3. If A sends **20 EUR**, A's payoff is 0 EUR and B' payoff is 20 EUR.
4. If A sends **4 EUR**, A's payoff is 16 EUR and B' payoff is 4 EUR.

Your decisions:

In the following, you will have to decide in the role of Person A.
If the computer chooses this section for payment, the computer will randomly match you with another participant. Additionally, it will be randomly determined who will be assigned to which role. For the payment of **both** participants, only the decision of the participant who is assigned to the role of Person A will be of interest.

C.2 Instructions - SPD (Seq. Prisoner's Dilemma)

In this section of the experiment the **situation** is as follows:
Person A and Person B both have an endowment of 10 EUR at the beginning and they have to decide how to use their endowment. Both can either keep their money or send it to the other person. If money is sent, this money is doubled by the computer.

The Details:

This section of the experiment consists of two consecutive stages:
In the **first stage**, Person A has two alternatives: sending **0 EUR** or **10 EUR** to Person B. In doing so, sent money is doubled. This decision determines how much money Person B has at the beginning of the second stage. We have two cases:

1. If A sends **10 EUR**, this amount is doubled. Because B already owns 10 EUR, B has 30 EUR in total now (and A has 0 EUR).

2. If A sends **0 EUR**, B does not get any additional money and owns his or her initial endowment of 10 EUR (as A does).

In the **second stage**, Person B has the following six different alternatives: sending **0, 2, 4, 6, 8 or 10 EUR** back to Person A. The amount that B chooses is again doubled. Person B can make his or her decision conditional on A's choice.

In **case 1)**, Person B owns 30 EUR at the beginning of the second stage whereas B owns 10 EUR in **case 2)**. For calculating B's final payoff, one still has to subtract the amount sent by B. Because the amount sent by B is also doubled, A can earn between **0 and 20 EUR** in this stage. In case 2), Person A gets an additional 10 EUR form stage 1, in case 1) A gets no additional money.

Two examples:

1. A sends **10 EUR**. Hence for a start, B has 30 EUR. If B also sends **10 EUR**, B finally earns 20 EUR, as does A. If B on the contrary sends 0 EUR, B finally earns 30 EUR and A earns 0 EUR.

2. A sends **0 EUR**. Hence for a start, B has 10 EUR. If B also sends **0 EUR**, B finally earns 10 EUR, as does A. If B on the contrary sends 10 EUR, B finally earns 0 EUR and A earns 30 EUR.

Mathematically we can express these **payoffs** as follows:

For **Person A**: 10 EUR - transfer to B + 2 * transfer from B

For **Person B**: 10 EUR - transfer to A + 2 * transfer from A

For further calculations, you can use the implemented payoff calculator

Your decisions:

In the following, you will have to decide both in the role of Person A as well as in the role of Person B. In the role of Person B, you will have to make a decision for both possible alternatives of Person A (0 EUR vs. 10 EUR).

If the computer chooses this section for payment, the computer will randomly match you with another participant. Additionally, it will be randomly determined who will be assigned to which role. For your payment, only your decision in this role (either as Person A or B) will be important.

C.3 Instructions - JOY-G (Joy-of-Destruction Game)

In this section of the experiment the **situation** is as follows:

Person A and Person B both get 11 EUR and have the opportunity to reduce the other's payoff but they do not have to do this.

The Details:

Both **Person A** and **Person B** are initially endowed with 11 EUR and then **simultaneously** make the following decision: Leaving the other person's income **unchanged** or reducing it by **1, 2, 3, 4, or 5 EUR**. Neither leaving the other's income unchanged nor reducing it costs money.

Importantly however, only in **two out of three** cases A's and B's decision will really determine payoffs. In **one out of three** cases these decisions are irrelevant and the computer will "overwrite" them by reducing both incomes either by **0, 1, 2, 3, 4, or 5 EUR**, where all amounts are equally likely.

Importantly, neither Person A nor Person B will be informed about which case has occurred. If for example own income is reduced, it is not evident whether the other person **or** the computer is responsible for this reduction.

Three examples:

1. If both A and B decide to leave the other's payoff unchanged and if the computer does not alter this, both A and B finally earn **11 EUR**. If the computer overwrites these decisions and chooses 2 EUR as a reduction, both A and B finally earn **9 EUR**.

2. If both A and B decide to reduce the other's payoff by 5 EUR and if the computer does not alter this, both A and B finally earn **6 EUR**.

3. If A decides to reduce B's payoff by 1 EUR and if B leaves A's payoff unchanged, A finally earns **11 EUR** and B earns **10 EUR** (in case the computer does not alter these decisions).

Your decisions:

Because Person A and B have symmetric roles, you will only have to decide on the next screen whether you want to reduce your counterpart's payoff or whether you want to leave it unchanged.
If the computer chooses this section for payment, the computer will randomly match you with another participant.

C.4 Instructions - Mini-UG (Mini Ultimatum-Game)

In this section of the experiment the **situation** is as follows:
Person A has two alternatives to split 20 EUR between him- or herself and Person B. Person B can accept or reject this choice.

The Details:

This section of the experiment consists of two consecutive stages:
In the **first stage**, Person A has to choose between two proposals on how to split the 20 EUR:

- **Proposal 1:** Person A and Person B both get **10 EUR**.
- **Proposal 2:** Person A gets **18 EUR** and Person B gets **2 EUR**.

In the **second stage**, Person B has two alternatives: B can either **accept** or **reject** A's choice.
If Person B accepts A's choice, payoffs are according to the chosen proposal. If Person B rejects A's choice, both A and B receive **0 EUR**.

Your decisions:

You will have to decide both in the role of Person A and in the role of Person B. In the role of Person B, you will have to decide for both proposals whether you would like to accept or reject them.
If the computer chooses this section for payment, the computer will randomly match you with another participant. Additionally, it will be randomly determined who will be assigned to which role. For your payment, only the decision in the assigned role (either as Person A or B) will then be of interest.

C.5 Instructions - SPD-P (Seq. Prisoner's Dilemma with Punishment)

In this section of the experiment the **situation** is as follows:
Similar to Section 2) [SPD] Person A and Person B get 10 EUR and have to decide how to use this money. Additionally however, there is a Person C. C gets 20 EUR and has the opportunity to reduce B's payoff.

The Details:

This section of the experiment consists of three consecutive stages:
In the **first stage**, Person A has to choose between two alternatives: Sending either **0 EUR** or **10 EUR** to Person B. We have two cases:

- **Case a**: If A sends **0 EUR, all other following stages are omitted** and the final payoffs are 10 EUR both for A and B and 20 EUR for C.
- **Case b**: If A sends **10 EUR**, this money is doubled. Because B already owns 10 EUR initially, B then temporarily owns 30 (and A 0 EUR).

In the **second stage**, Person B has two alternatives: Sending either **0 EUR** or **10 EUR** to Person A. The amount that Person B chooses is again doubled by the computer.
Hence, we have the following temporary payoffs:

- **Case 1**: If B sends **0 EUR**, B gets 30 EUR and A gets 0 EUR (unchanged payoffs of case b)
- **Case 2**: If B in contrast sends **10 EUR**, both A and B get 20 EUR each.

In the **third stage**, however, Person C has the opportunity to use part of his or her endowment to reduce Person B's payoff: This is possible in steps of 50 cents: **50, 100, 150, . . . , 600 cents**. Person C can condition his or her decision on B's decision.
Fifty cents invested by C reduce B's payoff **250 cents**, or 2.50 EUR. If C e.g. invests 200 cents (in order to reduce B's payoff), Person B will loose 10 EUR compared to the second stage. B's minimal payoff, however, can in the worst case only be reduced to 0 EUR (and does not get negative). If Person C does not invest anything in reducing B's payoff, all payoffs are the same as at the end of the second stage.

Three Examples:

1. If Person A sends **0 EUR**, stages two and three are dropped. A and B get 10 EUR and C gets 20 EUR (compare **case a**)
2. If Person A sends **10 EUR**, if B sends **0 EUR**, and if C does **not** reduce, A receives 0 EUR, B 30 EUR and C 20 EUR (compare **case 1**). If in contrast C invests e.g. **400 cents** (=4 EUR), A receives 0 EUR, B 10 EUR and C 16 EUR.
3. If Person A sends **10 EUR**, if B sends **10 EUR**, and if C does **not** reduce, A receives 20 EUR, B 20 EUR and C 20 EUR (compare **case 2**). If in contrast C invests e.g. **50 cents**, A receives 20 EUR, B 17.50 EUR and C 19.50 EUR.

For further calculations, you can use the implemented payoff calculator

Your decisions:

In the following, you will have to decide in all three roles (Person A, Person B, and Person C). In the role of Person C, you will have to make decision for both alternative actions of Person B.
If the computer chooses this section for payment, the computer will randomly match you with another participant. Additionally, it will be randomly determined who will be assigned to which role.

C.6 Instructions - DG-P (Dictator Game with Punishment)

In this section of the experiment the **situation** is as follows:
At the beginning, Person A and Person B get 5 EUR each. Then, Person A additionally gets 10 EUR and he or she can send money to Person B, but does not have to do this. In addition, Person C can reduce A's payoff.

The Details:

This section of the experiment consists of two consecutive stages:
In the **first stage**, **Person A** has three alternatives: Sending **0, 2.50 or 5 EUR** to Person B.
Person B knows about Person A's alternatives but does not make any decision him- or herself and has to accept A's decision.

In the **second stage**, **Person C** can reduce Person A's payoff. C can invest either **0, 1, 2, 3, 4, or 5 EUR**. One euro invested by C reduces A's payoff exactly by one euro. Person C can condition his decision with respect to A's decision.

Hence, we have the following **payoffs**:

- **Person A**: 15 EUR - amount sent to B - reduction by C
- **Person B**: 5 EUR + amount sent by A
- **Person C**: 20 EUR - reduction of A's payoff

Two Examples:

1. If Person A sends **0 EUR** and if C does **not** reduce, we have the following payoffs: A 15 EUR, B 5 EUR, and C 20 EUR. If in contrast C invests e.g. **3 EUR**, payoffs change to: A 12 EUR, B 5 EUR, and C 17 EUR.
2. If Person A sends **5 EUR** and if C does **not** reduce, we have the following payoffs: A 10 EUR, B 10 EUR, and C 20 EUR. If in contrast C invests e.g. **1 EUR**, payoffs change to: A 9 EUR, B 10 EUR, and C 19 EUR.

Your decisions:

In the following, you will have to make deicsions in the role of Person A and C. In the role of C, you will have to decide for all three alternatives of A (0, 2.50, or 5 EUR).
If the computer chooses this section for payment, the computer will randomly match you with another participant. Additionally, it will be randomly determined who will be assigned to which role.

C.7 Control Questions (SPD, SPD-P, DG-P)

SPD - Seq. Prisoner's Dilemma

1. Assume the following: A sends 10 EUR. Hence, A **temporarily** owns 0 EUR and B owns 30 EUR. In this situation, B sends 4 EUR.
   - In this situation, how large is the payoff of Person A?
   - In this situation, how large is the payoff of Person B?
2. Assume the following: A sends 10 EUR. Hence, A **temporarily** owns 0 EUR and B owns 30 EUR. In this situation, B sends 6 EUR.
   - In this situation, how large is the payoff of Person A?
   - In this situation, how large is the payoff of Person B?

SPD-P - Seq. Prisoner's Dilemma with Punishment

1. Assume the following: Case 2 has occurred and C invests **400 cents** (also compare example 3).
   - How large is A's payoff?
   - How large is B's payoff?
   - How large is C's payoff?
2. Assume the following: Case 1 has occurred and C invests **600 cents** (also compare example 2).
   - How large is A's payoff?
   - How large is B's payoff?
   - How large is C's payoff?

DG-P - Dictator Game with Punishment

1. Assume the following: Person A sends **2.50 EUR** and C invests **5 EUR**.
   - How large is A's payoff (in EUR)?
   - How large is B's payoff (in EUR)?
   - How large is C's payoff (in EUR)?

D Well-being Questionnaire

This section presents the WB items used in my questionnaire. Nearly all items are also used by KE, only the Social Well-Being (SoWB) Scale is added.

Hedonic Well-Being (HWB)

Long-run HWB

- **Overall Happiness (OH)**
  A 9-point-scale ranging from "extremely unhappy" to "extremly happy" is used:
  OH: *Overall*, how would you describe yourself?

- **Watson, Clark and Tellegen's (1988) Postive Affect (PAS) and Negative Affect (NAS) Schedules**
  Subjects use a 5-point-scale ranging from "very slightly or not at all" to "extremely" to indicate to what extent they have felt the way the words suggest during the past few weeks:
  - **PAS items**: interested, alert, excited, inspired, strong, determined, attentive, active, enthusiastic, proud
  - **NAS items**: irritable, distressed, ashamed, upset, nervous, guilty, scared, jittery, hostile, afraid

Highest/ Lowest Happiness

- **Highest/ Lowest Happiness (HH/ LH)**
  A 9-point-scale ranging from "extremely unhappy" to "extremly happy" is used:
  - **HH**: Over the past week, what is the *highest* level you experienced?
  - **LH**: Over the past week, what is the *lowest* level you experienced?

Life Satisfaction

- **Diener, et al.'s (1985) Satisfaction with Life Scale (SWL)**
  A 7-point-scale ranging from "strongly disagree" to "strongly agree" is used:
  1. In most ways my life is close to my ideal.
  2. The conditions of my life are excellent.
  3. I am satisfied with my life.
  4. So far I have gotten the important things I want in life.
  5. If I could live my life over, I would change almost nothing.

Short-run HWB

- **Now Happiness (NH)**
  A 9-point-scale ranging from "extremely unhappy" to "extremly happy" is used:
  NH: *Right now*, how would you describe yourself?

- **Batson, et al.'s (1988) Mood Index (MI)**
  On a 9-point-scale, subjects have to express their current mood for several pairs of adjectives

  - **Mood items**: bad mood-good mood, sad-happy, depressed-elated, dissatisfied-satisfied, gloomy-cheerful, displeased-pleased, sorrowful-joyful
  - **Fillers**: nervous-calm, tense-relaxed, uncomfortable-comfortable, apathetic-caring, lethargic-energetic, unconfident-confident, unresponsive-emotional, passive-active

Eudaimonic Well-Being (EWB)

- **Ryff's (1995) Scales of Psychological Well-Being (SPWB)**
  A 6-point-scale ranging from "strongly disagree" to "strongly agree" is used:

  1. I tend to be influenced by people with strong opinions.*
  2. In general, I feel I am in charge of the situation in which I live.
  3. I think it is important to have new experiences that challenge how you think about yourself and the world.
  4. Maintaining close relationships has been difficult and frustrating for me.*
  5. I live life one day at a time and don't really think about the future.*
  6. When I look at the story of my life, I am pleased with how things have turned out.
  7. I have confidence in my opinions, even if they are contrary to the general consensus.|
  8. The demands of everyday life often get me down.*
  9. For me, life has been a continuous process of learning, changing and growth.|
  10. People would describe me as a giving person, willing to share my time with others.
  11. Some people wander aimlessly through life, but I am not one of them.|
  12. I like most aspects of my personality.
  13. I judge myself by what I think is important, not by the values of what others think is important.
  14. I am quite good at managing the many responsibilities of my daily life.|
  15. I gave up trying to make a big improvements or changes in my life a long time ago.*
  16. I have not experienced many warm and trusting relationships with others.*|
  17. I sometimes feel as if I've done all there is to do in life.*
  18. In many ways, I feel disappointed about my achievements in life.*|

  * indicates reverse scored items. | indicates items selected for Index of PWB (PWBI)

- **Jones and Crandall's (1986) Self-Actualization Index (SAI)**
  A 4-point-scale ranging from "disagree" to "agree" is used:

  1. I do not feel ashamed of any of my emotions.
  2. I feel I must do what others expect me to do.*

3. I believe that people are essentially good and can be trusted.
4. I feel free to be angry at those I love.
5. It is always necessary that others approve of what I do.*
6. I don't accept my own weaknesses.*
7. I can like people without having to approve of them.
8. I fear failure.*
9. I avoid attempts to analyze and simplify complex domains.*
10. It is better to be yourself than to be popular.
11. I have no mission in life to which I feel especially dedicated.*
12. I can express my feelings even when they may result in undesirable consequences.
13. I do not feel responsible to help anybody.*
14. I am bothered by fears of being inadequate.*
15. I am loved because I give love.

* indicates reverse-scored items.

- **Keyes' Social Well-Being (SoWB)**
  A 6-point-scale ranging from "strongly disagree" to "strongly agree" is used:

1. The world is too complex for me.*
2. I don't feel I belong to anything I'd call a community.*
3. People who do a favor expect nothing in return.
4. I have something valuable to give to the world.
5. The world is becoming a better place for everyone.
6. I feel close to other people in my community.
7. My daily activities do not produce anything worthwhile for my community.*
8. I cannot make sense of what's going on in the world.*
9. Society has stopped making progress.
10. People do not care about other people's problems.*
11. My community is a source of comfort.
12. I find it easy to predict what will happen next in society.
13. Society isn't improving for people like me.*
14. I believe that people are kind.
15. I have nothing important to contribute to society.*

* indicates reverse-scored items.

References

Abbink, K. and B. Herrmann (2011). The moral costs of nastiness. *Economic Inquiry 49(2)*, 631–633.

Anderson, J. E., S. V. Burks, J. P. Carpenter, L. Götte, K. Maurer, D. Nosenzo, R. Potter, K. Rocha, and A. Rustichini (2013). Self-selection and variations in the laboratory measurement of other-regarding preferences across subject pools: evidence from one college student and two adult samples. *Experimental Economics 16*(2), 70–89.

Batson, C., J. Dyck, J. Brandt, J. Batson, A. Powell, M. McMaster, and C. Griffitt (1988). Five studies testing two new egoistic alternatives to the empathy-altruism hypothesis. *Journal of Personality and Social Psychology 55(1)*, 52–77.

Diener, E., R. Emmons, R. Larsen, and S. Griffin (1985). The Satisfaction with Life scale. *Journal of Personality Assessment 49(1)*, 71–75.

Falk, A., E. Fehr, and U. Fischbacher (2003). On the nature of fair behavior. *Economic Inquiry 41(1)*, 20–26.

Hoff, K., M. Kshetramade, and E. Fehr (2011). Caste and punishment: The legacy of caste culture in norm enforcement. *The Economic Journal 121*, F449–F475.

Jones, A. and R. Crandall (1986). Validation of a short index of self-actualization. *Personality and Social Psychology Bulletin 12(1)*, 63–73.

Keyes, C. (1998). Social well-being. *Social Psychology Quarterly 61(2)*, 121–140.

Ryff, C. (1989). Happiness is everything, or is it? exploarations on the meaning of psychological well-being. *Journal of Personality and Social Psychology Vol. 57, No. 5*, 1069–1081.

Watson, D., A. Clark, and A. Tellegen (1988). Development and validation of brief mesarues of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology 54(6)*, 1063–1070.