Autistic and non-autistic prosocial decision-making: The impact of recipient neurotype

Supplementary materials

Type II Wald Chi-Square Tests for Choice Model

Table S1

Analysis of Variance Results (Type II Wald Chi-Square Tests)

Effect	χ^2	***df***	***p***
Neurotype	0.30	1	.584
Recipient	828.68	2	< .001***
Effort	650.09	1	< .001***
Reward	434.47	1	< .001***
Neurotype × Recipient	13.77	2	.001**
Neurotype × Effort	11.35	1	< .001***
Recipient × Effort	0.38	2	.829
Neurotype × Reward	0.12	1	.732
Recipient × Reward	28.89	2	< .001***
Effort × Reward	18.03	1	< .001***
Neurotype × Recipient × Effort	10.22	2	.006**
Neurotype × Recipient × Reward	0.57	2	.754
Neurotype × Effort × Reward	2.38	1	.123
Recipient × Effort × Reward	6.49	2	.039*
Neurotype × Recipient × Effort × Reward	3.69	2	.158

Note. Response variable: Choice (chosen). Significant effects are indicated by asterisks.
*p < .05, **p < .01, ***p < .001

Choice model interactions not reported in manuscript

Effort sensitivity greatest at lowest reward levels:

An interaction between effort and reward, $\chi2(1) = 18.03$, $p < .001$ (Figure S1), revealed that the influence of effort on choice behaviour was greatest at the lowest reward levels (2 credits, $\chi2(1) = 243.02$, $p < .001$; 4 credits, $\chi2(1) = 244.73$, $p < .001$) and decreased as reward magnitude increased (6 credits, $\chi2(1) = 141.01$, $p < .001$), 8 credits, $\chi2(1) = 97.534$, $p < .001$).

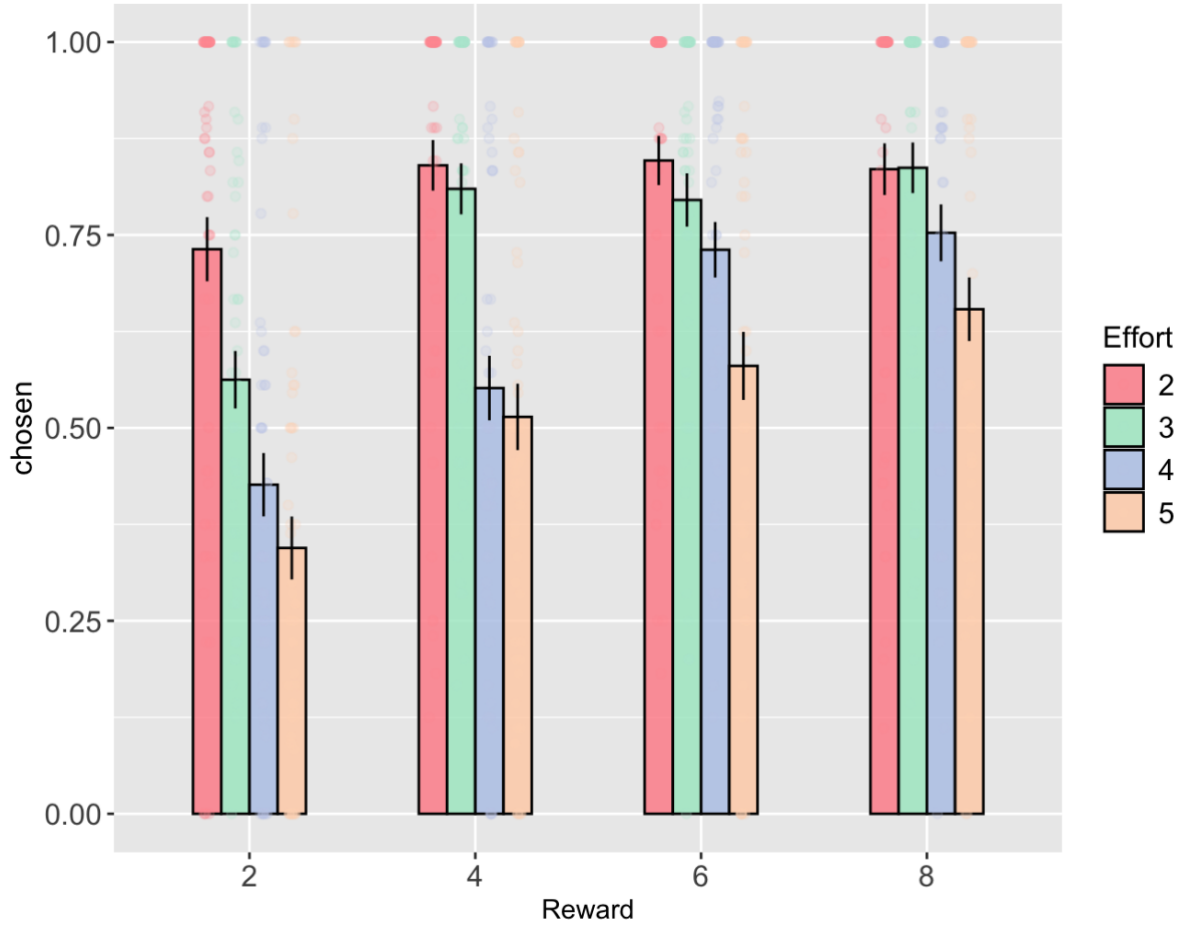

Figure S1. Proportion of trials where participants chose to exert effort rather than rest to gain rewards for a specified recipient as a function of the differential reward (x-axis) and effort (coloured bars) required to obtain that reward.

Greater self-benefit bias and other-benefit aversion as reward increases:

A significant 2-way interaction between recipient and reward, $\chi2(2) = 28.89$, $p < .001$, and a significant 3-way interaction between recipient, effort and reward, $\chi2(2) = 6.49$, $p = .039$, demonstrate a greater self-preference on choice behaviour as reward and effort increase. At the lowest reward level (2 credits), participants strongly preferred Self over Other-Different ($\beta = -4.08$, SE = 0.32, $z = -12.81$, $p < .001$) and over Other-Same ($\beta = -4.33$, SE = 0.32, $z = -13.57$, $p < .001$). This self-preference pattern strengthened as reward increased, with the largest differences observed at the highest reward level (8), where Self was strongly preferred over Other-Different ($\beta = -7.23$, SE = 1.01, $z = -7.17$, $p < .001$) and Other-Same ($\beta = -8.16$, SE = 1.01, $z = -8.10$, $p < .001$). Notably, no significant differences were found between Other-Different and Other-Same across any reward level (all $p > .05$), suggesting that participants did not distinguish between these two recipient categories regardless of reward magnitude.

Follow-up slope comparisons for the relationship between effort and choice within each recipient category and across reward levels reveal that across all reward levels, effort negatively predicted choice, such that participants were less likely to choose as effort increased. However, the extent of this negative relationship varied by recipient type. The estimated slopes for effort at each reward level are presented in Table S2. At reward level 2, effort significantly reduced choice across all recipient types: Other-Different ($\beta = -0.155$, SE = 0.0146, $p < .001$), Other-Same ($\beta = -0.171$, SE = 0.0152, $p < .001$), and Self ($\beta = -0.115$, SE = 0.0155, $p < .001$). Comparisons of these slopes revealed that the negative effect of effort was significantly stronger for Other-Same compared to Self, $t(2062) = -2.594$, $p = .0286$. However, there was no significant difference in the effect of effort between Other-Different and Other-Same ($p = 1.00$) or Other-Different and Self ($p = .17$). At reward level 4, a similar pattern emerged, with effort negatively predicting choice for all recipients: Other-Different (β

= -0.159, SE = 0.0132, $p < .001$), Other-Same (β = -0.156, SE = 0.0139, $p < .001$), and Self (β = -0.072, SE = 0.0138, $p < .001$). However, post-hoc comparisons indicated that effort had a significantly stronger negative effect for both Other-Different ($t(2041) = -4.527$, $p < .0001$) and Other-Same ($t(2039) = -4.320$, $p < .0001$) compared to Self, suggesting that participants were more likely to sustain effort for themselves than for others at this reward level.

At reward level 6, the negative effect of effort was again evident for all recipients: Other-Different (β = -0.139, SE = 0.0134, $p < .001$), Other-Same (β = -0.122, SE = 0.0129, $p < .001$), and Self (β = -0.049, SE = 0.0125, $p < .001$). Pairwise comparisons revealed that effort had a significantly greater negative effect for Other-Different compared to Self ($t(2068) = -4.915$, $p < .001$) and for Other-Same compared to Self ($t(2066) = -4.098$, $p < .001$). No significant difference was found between Other-Different and Other-Same ($p = 1.00$). Finally, at reward level 8, effort remained a strong negative predictor of choice for Other recipients, but had a much weaker effect for Self: Other-Different (β = -0.094, SE = 0.0121, $p < .001$), Other-Same (β = -0.105, SE = 0.0123, $p < .001$), and Self (β = -0.018, SE = 0.0116, $p < .001$). Comparisons revealed that effort had a significantly stronger effect on Other-Different compared to Self ($t(2057) = -4.501$, $p < .001$) and on Other-Same compared to Self ($t(2058) = -5.120$, $p < .001$), but no significant difference between Other-Different and Other-Same ($p = 1.00$).

Overall, these results suggest that effort consistently reduces choice, but this effect varies across recipient types and reward levels. At lower reward levels (2 or 4 credits), the negative effect of effort was relatively similar across recipient types. However, at higher reward levels (6 or 8 credits), effort had a significantly weaker effect on choices for Self, indicating that participants were more willing to exert effort when the reward was personally beneficial (see Figure S2). This pattern suggests a self-benefit bias, where participants were

less deterred by high effort when choosing for themselves but remained effort-averse when choosing for others, as reward levels increased.

Table S2.

Estimated Slopes for Effort by Recipient and Reward Level

Reward Level	Other-Different (β)	Other-Same (β)	Self (β)	Other-Different - Other-Same (*p*)	Other-Different - Self (*p*)	Other-Same - Self (*p*)
2	-0.155	-0.171	-0.115	1.00	.170	.029
4	-0.159	-0.156	-0.072	1.00	<.001	<.001
6	-0.139	-0.122	-0.049	1.00	<.001	<.001
8	-0.094	-0.105	-0.018	1.00	<.001	<.001

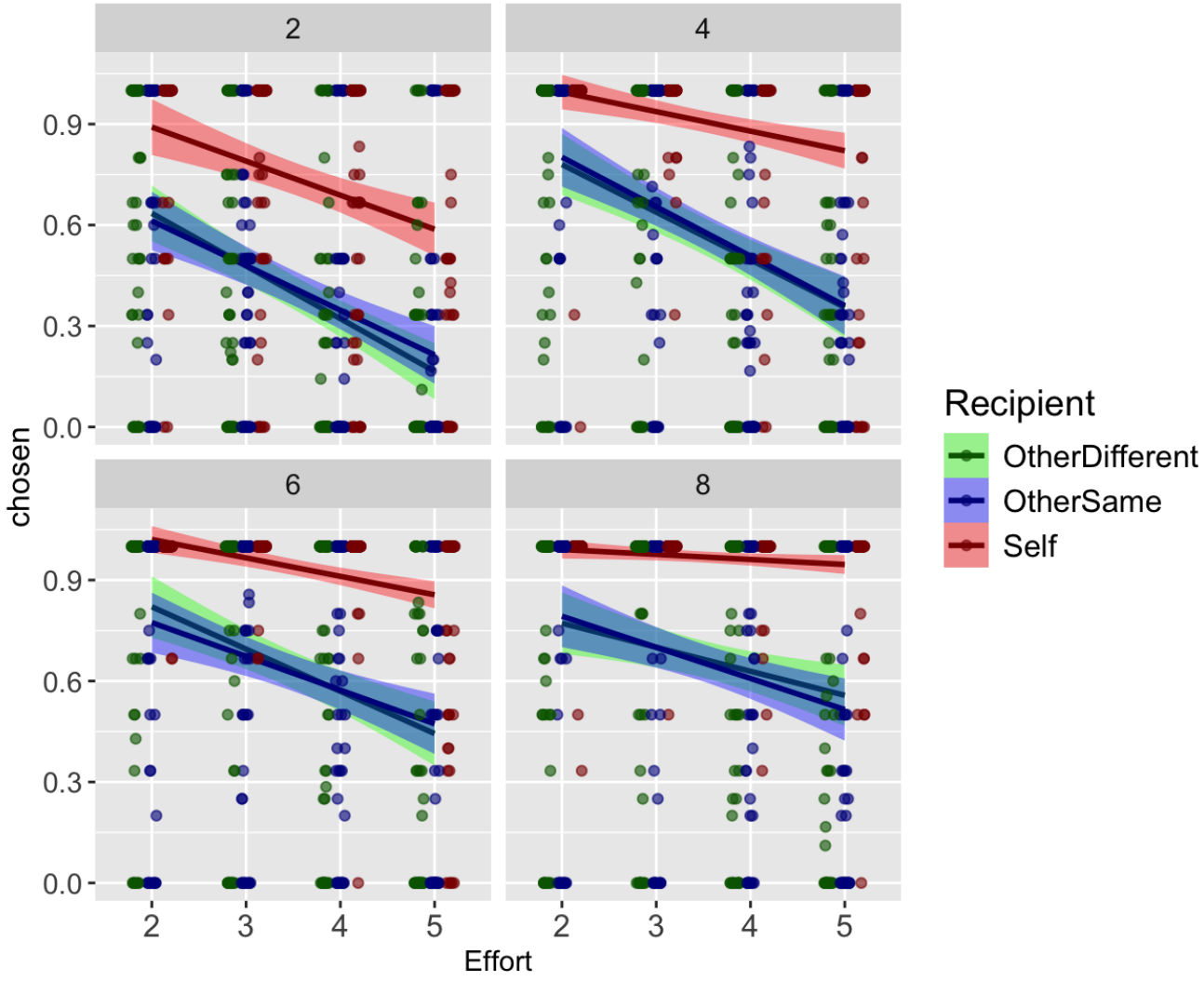

Figure S2. Proportion of trials where participants chose to exert effort rather than rest to gain rewards for others of a different neurotype (green slopes), others of the same neurotype (blue slopes) and themselves (red slopes) across varying effort (x-axis) and reward levels (facet panels).

Type II Wald Chi-Square Tests for Force Model

Table S3

Analysis of Variance Results for Force Production (Type II Wald Chi-Square Tests)

Effect	χ^2	***df***	***p***
Neurotype	0.82	1	.366
Recipient	159.25	2	< .001***
Effort	4372.27	1	< .001***
Reward	39.17	1	< .001***
Neurotype × Recipient	13.07	2	.001**
Neurotype × Effort	3.15	1	.076
Recipient × Effort	7.57	2	.023*
Neurotype × Reward	0.46	1	.498
Recipient × Reward	1.79	2	.409
Effort × Reward	5.39	1	.020*
Neurotype × Recipient × Effort	1.68	2	.432
Neurotype × Recipient × Reward	3.80	2	.149
Neurotype × Effort × Reward	0.10	1	.746
Recipient × Effort × Reward	8.09	2	.017*
Neurotype × Recipient × Effort × Reward	4.85	2	.089

Note. Response variable: Normalised Force (Force_norm). Significant effects are indicated by asterisks. *p < .05, **p < .01, ***p < .001

Force model interactions not reported in manuscript

Two-way effort x reward interaction:

An interaction between effort and reward, $\chi2(1) = 5.39$, $p = .020$, reveals that this influence of effort on force exerted was greatest at the highest reward level (8 credits, $\chi2(1) = 1374.6$, $p < .001$ and decreased as reward magnitude decreased (6 credits, $\chi2(1) = 1286.3$, $p < .001$; 4 credits, $\chi2(1) = 1161.1$, $p < .001$; 2 credits, $\chi2(1) = 486.81$, $p < .001$). See Figure S3.

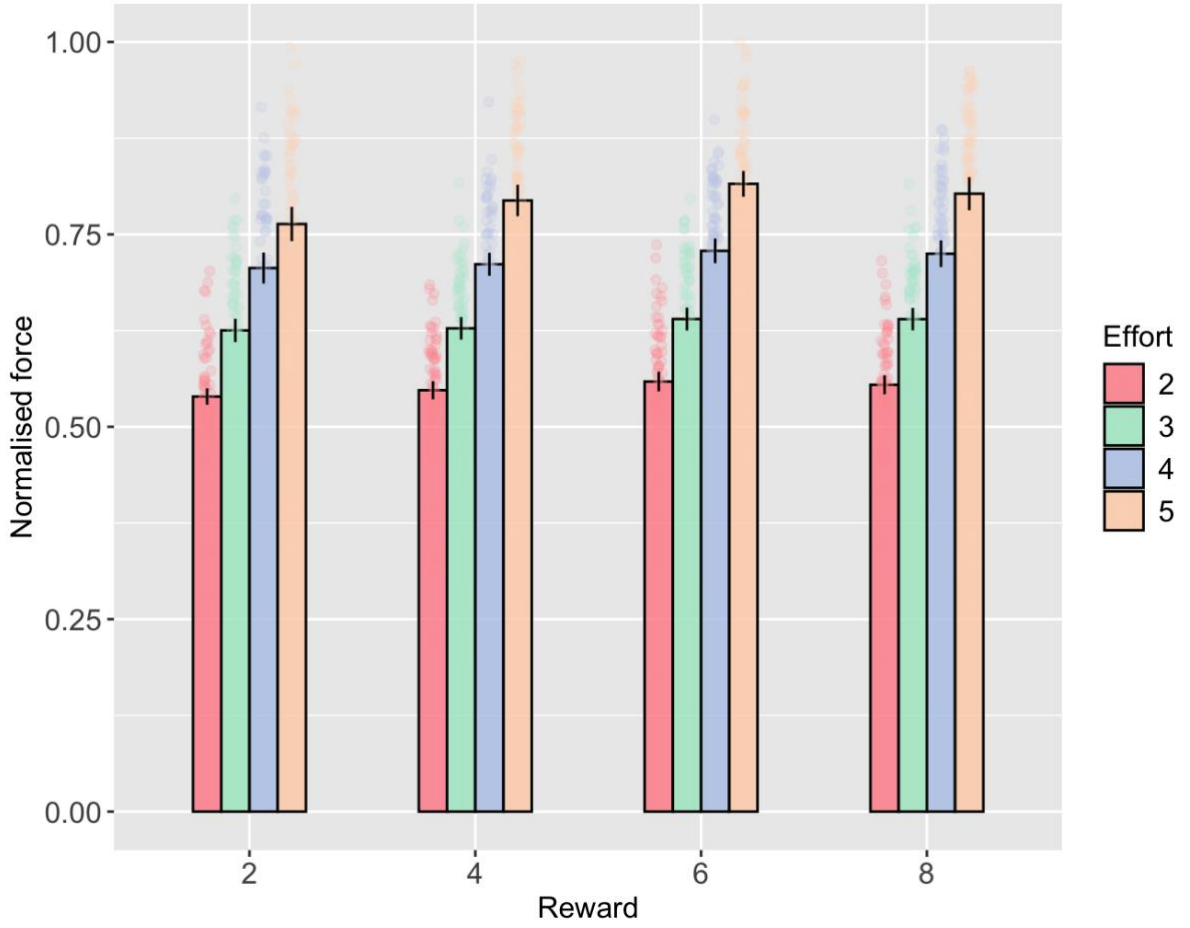

Figure S3. Normalised force applied to force-grip device to gain rewards for a specified recipient as a function of the differential reward (x-axis) and effort (coloured bars) required to obtain that reward.

Self-bias in action energisation is greater at highest effort levels

A significant interaction between recipient and effort was observed on action energisation, $\chi 2(2) = 7.57$, $p = .023$ (Figure S4). As effort levels increased, self-biases on force exerted became more pronounced. For example, the normalised force difference between Self and Other-Same increased from the lowest effort level ($\beta = -0.04$, $SE = 0.008$, $t = -5.60$, $p < .001$), to the highest effort level ($\beta = -0.063$, $SE = 0.018$, $t = -3.60$, $p < .001$). Similarly, the normalised force difference between Self and Other-Different increased from the lowest effort level ($\beta = -0.057$, $SE = 0.008$, $t = -7.18$, $p < .001$) to the highest effort level ($\beta = -0.083$, $SE = 0.017$, $t = -4.74$, $p < .001$). Critically, no significant differences were found in action energisation when actions benefited others of the same or different neurotype to oneself across any effort level (all $p > .05$).

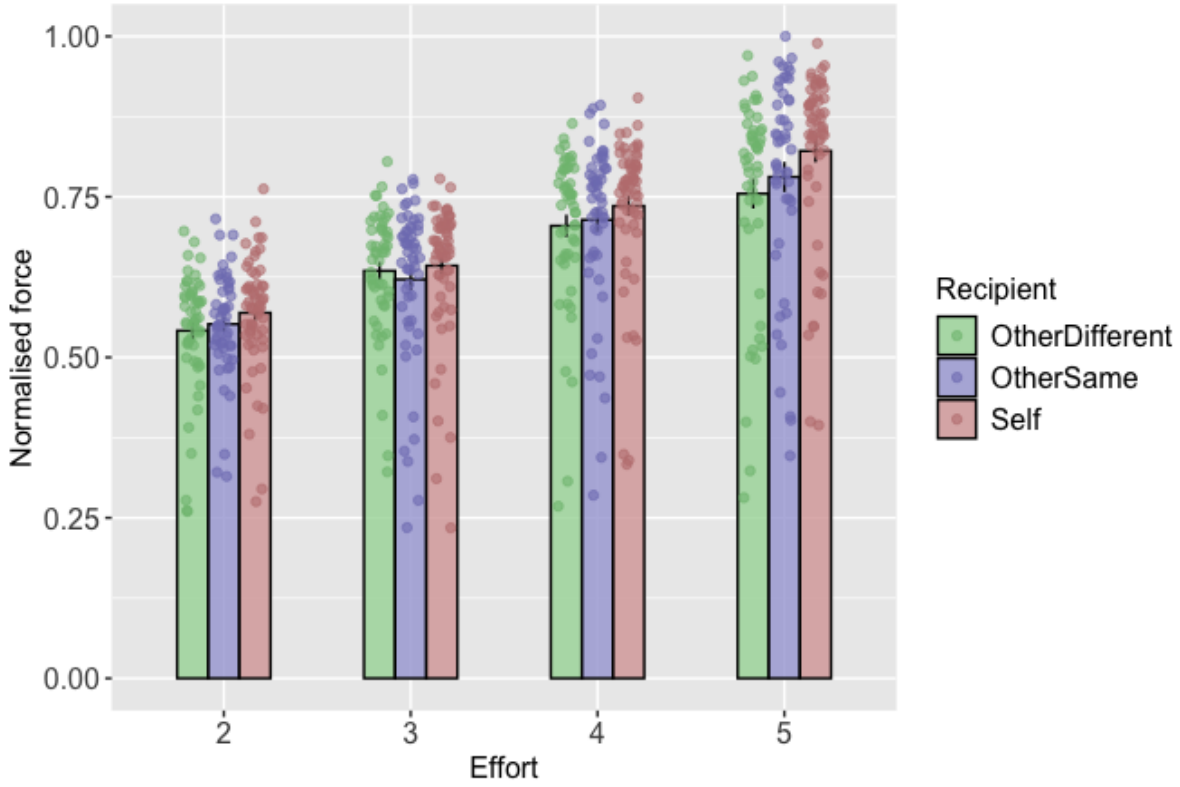

Figure S4. Normalised force applied to gain rewards for others of a different neurotype (green bars), others of the same neurotype (blue bars) and self (red bars) as a function of effort.

Effort-reward force trade-off only for others of a different neurotype

Following a significant 3-way interaction between recipient, effort and reward, $\chi^2(2) = 8.09$, $p = .017$, separate linear mixed-effects models revealed a significant interaction between effort and reward only for trials where actions benefited others of a different neurotype to oneself ($\chi^2(1) = 10.05$, $p = .002$) and not for trials where actions benefited the self ($p = .645$) or others of the same neurotype to oneself ($p = .255$). See Figure S5 for a visual depiction of these findings.

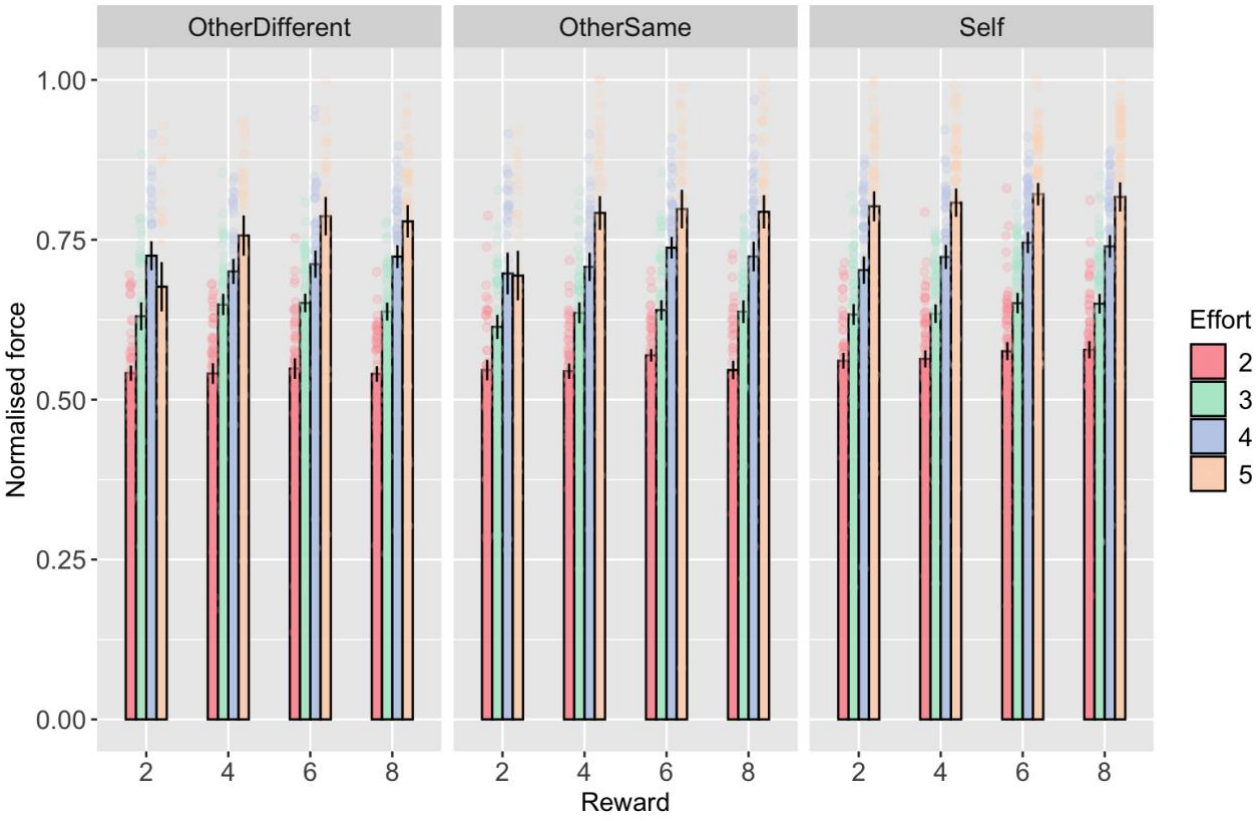

Figure S5. Normalised force applied to gain rewards for others of a different neurotype (left facet), others of the same neurotype (middle facet) and self (right facet) as a function of different reward (x-axis) and effort (coloured bars) required.