Supplemental Materials

To accompany

Physical effort exertion for peer feedback reveals evolving social motivations from adolescence to young adulthood

Rodman, Powers, Kastman, Kabotyanski, Stark, Mair, and Somerville

Contents:

1. Supplemental Methods
2. Secondary and Control Analyses
3. Plotting Methods
4. Supplemental Figures
5. Supplemental Tables
6. **Supplemental Methods**

*Participants*

One hundred nine healthy individuals were recruited from the greater Boston, MA area through online forums such as Craigslist, advertising in local newspapers, and flyers. Exclusion criteria included history of a neurological disorder and current psychiatric disorder. Seven participants were excluded from final analyses due to noncompliance during the calibration (N=2; ages 19.48 and 22.78 years (y)), squeezing for the entire trial against instruction (N=2; ages 20.77y and 20.88y), disbelief the task was real (N=1; age 21.98y), technical malfunction of dynamometer (N=1; age 12.38y), and exhibiting difficulty comprehending task instructions (N=1; age 21.63y). The remaining 102 participants were included in data analyses. Two participants did not complete the expectancy ratings (ages 12.16y and 17.65y) and three participants did not complete the desirability ratings (ages 14.52y, 16.28y, and 21.30y), resulting in a total of 100 participants and 99 participants in analyses examining predictors expectancy and desirability, respectively.

A Generalized Additive Mixed Effects Model (GAMM) is a semi-parametric model, where the smoothed terms are non-parametric and any non-smoothed terms (i.e., here, the covariates of non-interest) are the parametric terms (Wood, 2003, 2017). Thus, a traditional power analysis based on hypothesis testing within a parametric framework is not appropriate for the current analyses.

*Equipment*

The grip force task was presented using PsychoPy (v1.84) which interfaced with Biopac, Inc. (Goleta, CA) hand dynamometer hardware. Grip force was recorded using a hand dynamometer (Biopac TSD121B-MRI) made of two molded plastic cylinders that, when squeezed, compress an air tube. Air compression was converted into voltage proportional to the exerted force by a transducer. This value was sent to a Biopac DA100C module and converted from analog to digital signal using a custom-built 3.5mm breakout board connected to a National Instruments USB-6009 multifunction IO box. This digital signal was sampled at 60Hz and used a continuous input to PsychoPy to provide real-time visual feedback during the task, wherein participants saw a vertical bar on the screen rise and fall proportionally to their grip force (Figure 1A). Participants were instructed to use their dominant hand, which was stabilized using Velcro for the duration of the procedures to maintain uniform hand configuration.

*Maximum force calibration*

At the start of the study session, participants completed a stepwise calibration procedure to titrate the relative difficulty of trials to each individual’s hand strength. During the calibration, participants repeatedly attempted to reach the top of a vertical bar representing sequentially higher force levels on each successive attempt until they were no longer able to reach the threshold and their maximum grip strength was recorded. This maximum force value was used to calibrate the task thresholds proportionally to participants’ strength (easy threshold: 40% of maximum; hard threshold: 80% of maximum), as in prior work (Kurniawan et al., 2010). This procedure was repeated at the end of the session to quantify fatigue effects over time, so they could be controlled for during data analyses.

*Debriefing*

At the end of the study session, participants were debriefed using a funnel debriefing procedure, and any participant who did not believe the cover story was excluded from the final sample (N=1; age 21.98y). This procedure involves asking participants questions about their experience completing the task. These questions started as broad (i.e., did you notice anything off about the task?) and became increasingly specific and geared to assessing their belief in the cover story (i.e., did you believe the peers/feedback was real?). The interviewer then assigned a subjective rating from completely believed cover story (0) to did not believe cover story at all (5) Participants with a rating of 5 were excluded (n=1). The full debriefing form can be accessed on OSF here: <https://osf.io/yf7a6/?view_only=db215376d9de4074b2e1640d3612b666>.

*Outcome measures*

Timeseries data acquired in the grip force task permit the isolation of several different components of the overall grip response: peak grip force, speed, and opt out behavior. These three dependent variables were chosen because response vigor (i.e., strength and speed) and choice behavior (i.e., opting out) are the most common behavioral constructs examined in the motivation/effort literature (Arulpragasam et al., 2018; Botvinick et al., 2009; Kurniawan et al., 2010; Niv et al., 2007; Pessiglione et al., 2007). We computed the standard measure of physical effort exertion–*peak grip force*–defined for each trial as the highest intensity of grip force exerted (excluding opt-out trials, described below), expressed in units of percent of the participant’s maximum force calibration value.

Second, we measured the *speed of effort exertion* operationalized as time (in milliseconds from the start of the grip phase) to reach the threshold for each successful trial. Together, peak grip force and speed reflect different components of response vigor exerted during pursuit of a desired outcome, a readout of incentive motivation (Niv et al., 2007; Salamone et al., 2009). Trials that were unsuccessful (i.e., participant did not reach the threshold) were not included in analyses.

Although participants were instructed to complete each trial, initial data inspection revealed a small proportion of trials for which participants did not attempt to approach the threshold, suggesting that they chose not to exert force on that trial. We quantified and tallied these *opt-out* trials, which may reflect strategic decisions to forego all effort when the target was not of sufficient value. For easy trials, a trial was considered an opt-out if the grip force did not exceed half the distance to the threshold, or 20%. For hard trials, a trial was tallied as an opt-out if the grip force did not exceed the average peak grip force for easy trials, or 54.85%. These definitions are consistent with prior work (Rodman et al., 2021), though the computed thresholds were based on the current reported data.

For all three dependent variables, we inspected the distributions of the residuals of each regression model. Data points identified as outliers (+/- 3SD) were winsorized. The dependent variable of speed required a log transform to correct for non-normality of residuals.

1. **Secondary and Control Analyses**

*Raw ratings of expectancy and desirability.*

Prior to the dynamometer task, participants were given a prompt to imagine they were starting at a new school and based on the photographs of peers, answered the questions “*how much would this person want to be friends with you?*” – a measure of *acceptance expectancy* – and “*how much would you want to be friends with this person?” –* a measure of *peer desirability –* for each peer stimulus. Responses were made by clicking on a continuous scale from “not at all” (0) to “very much” (100).

For acceptance expectancy, participants indicated that they expected to be rated about midway between not at all and very much (M=46.40, SD=23.46), and participants’ rating of how much they would like the peers was about the same (M=46.76, SD=23.96). Expectancy and desirability ratings were positively correlated (*r*(97)=0.53, *p*<.001), sharing about 28% of variance.

Previous work has shown that there are gender differences in how adolescents process social evaluation (Guyer et al., 2012). To determine whether the reported effects in the present study were especially pronounced in males or females, we conducted several analyses investigating the effects of gender. Here we operationalize participant gender as their self-identified gender, which in this sample fell into categories of male and female. We operationalized peer gender as the expressed gender identity of the stimulus images used in the study (without access to information about their assigned sex), categorized as male or female.

First, we built LME models to examine whether subjective ratings of acceptance expectancy or peer desirability differed by age, participant gender, or peer gender (findings reported below). Second, we recomputed all key models reported in the main document with gender as a covariate to examine whether a) reported findings hold accounting for gender, and b) there were effects of gender on task behaviors. All key findings described in the main document held and no main effects of gender on effort exertion (i.e., peak grip force, speed, or opting out) were found.

*Raw ratings by gender and age.*

We examined whether gender qualified how favorably participants evaluated their peers and how favorably they expected to be evaluated by each peer. For both expectancy and desirability, GAMM with smoothed age terms solved linear functions; thus, all analyses report linear effects. We first computed an LME model to examine the main effects of age, participant gender, and peer gender on expectancy and desirability. Findings examining these effects on expectancy revealed significant effects of age (*B*=1.217, *SE*=0.393, *p=*0.003), such that younger participants expected to be rated less favorably. There was also a main effect of peer gender (*B*=-3.804, *SE*=0.479, *p*<0.001), with participants expecting to be rated more favorably by female peers compared to male peers. While the effect of participant gender on acceptance expectancy was not significant (*B*=5.199, *SE*=2.752, *p*=0.062), it showed a pattern whereby males expected to be liked more often than females.

When examining peer desirability ratings, findings from the additive LME model again showed significant main effects of age (*B*=0.912, *SE*=0.387, *p*=0.020) with older participants more likely to rate peers more favorably. A main effect of peer gender (*B*=-7.389, *SE*=0.504, *p*<0.001) indicated participants rated female peers more favorably than male peers, and there was no effect of participant gender (*B*=0.721, *SE*=2.731, *p*=0.792). These main effects were qualified by interactions described below (see Figure S5 for age main effects of age).

Linear mixed-effects regressions also examined the effect of participant gender, peer gender, age, and their interaction on subjective ratings of acceptance expectancy and peer desirability with a random effect of subject. Findings revealed a significant three-way-interaction of participant gender by peer gender by age on subjective ratings of acceptance expectancy (*B*=-2.684, SE=0.268, *p*<.001). Males and females at younger ages believed they would be rated more favorably by the same gender than the opposite gender and this effect was especially pronounced for female participants. However, as adolescents transitioned to young adulthood, this effect was diminished, and predictions became less yoked to peer gender (Figure S6A).

When examining the subjective ratings of peer desirability, a three-way-interaction also emerged (*B*=-2.067, SE=0.281, *p*<.001). Female participants rated female peers more favorably than male peers consistently across age. Male participants, on the other hand, rated male peers more favorably when they were younger, but around late adolescence, began rating female peers more favorably which continued into adulthood (Figure S6B).

***Inclusion of fatigue measure in social vs. non-social analyses.***

In the current study, the task with money targets always preceded the task with social targets. Because our interests center on interacting age-effects, ordering effects would affect participants of all ages equally. Nonetheless, we completed control analyses to account for possible order effects and related fatigue across tasks.

First, we included total trial number across both tasks (i.e., trials 1-32 for money task and 33-93 for the social task) as a nuisance covariate in all analyses comparing the money and social tasks. Second, we repeated the maximum grip calibration procedure at the end of the session and calculated the difference score of the maximum grip calibration measured at the start and end of the entire study session to index fatigue across tasks. We re-computed all GAMMs examining social vs. non-social targets including this fatigue parameter as a nuisance covariate (Table S1). To examine the interaction between age and target type, we visualized the age interaction fit and calculated a 95% CI, which showed that all initially reported findings comparing monetary reward to peer feedback targets held. Therefore, findings identifying age-related differences in the effect of social vs. non-social target type on effort exertion are unlikely to be explained by fatigue- or order-effects.

***Inclusion of random slope of Task***

In the first set of analyses, we examine how relative effort exertion for money and peer feedback differs by age. For each outcome variable, we completed secondary analyses that included a random slope of task, given the amount of variance across participants. When examining the relative difference in marginal effects of effort exertion for money compared to social targets across age, all original findings hold. Similar to the original findings, adults showed either equivalent (i.e., peak, opt-out) or greater motivation (i.e., speed) for money compared to peer feedback, whereas adolescents showed a diminished or reversed pattern (i.e., greater force, less opt-out, and equivalent speed for social vs. money targets; see Figure S7).

1. **Plotting Methods**

*Figure 2. Age-effects of social vs. nonsocial targets on effortful behavior*

When examining whether money or social target type impacted effort exertion across age, the traditional summary output of the *mgcv’s* (Wood, 2019) *gamm* function using “by=factor” to elect the interacting factor term does not output formal statistics for an isolated interaction term, or the relative difference between these two spline functions. To explicitly test the relative difference in target type across age, we used the *plot\_diff* function of the extension package *itsadug* (Rij et al., 2020)*.* This tool plots the relative difference spline functions across age for the money vs. social targets and calculates a 95% Confidence Interval, which account for cross-covariance between the two spline functions. Using this approach, we interpreted the fit of the function and accompanying 95% CI to determine the interaction effect of task type and age.

*Figures 3 and 4. Effect of expectancy/desirability on effort exertion for peer feedback across age*

To visualize the effects of an interaction tensor, a 3-dimensional functional plane to characterize nonlinear patterns of continuous interacting variables, we used the *fvisgam* function from the package *itsadug* (Rij et al., 2020) in R (R Core Team, 2020). This approach produced a heat map of marginal effects, whereby the y and x-axes display the effect of the predictor (i.e., expectancy/desirability) and age, respectively and the color gradient from light to dark indicates a scale of less to more effort exertion of the dependent variable (i.e., peak grip, speed, or opt-out behavior). For these plots, the predictors expectancy and desirability were winsorized (+/- 3SD) so the displayed patterns were not overly impacted by outliers (NB: there was one exception for the plot displaying the effect of age and desirability on opt-out behavior, whereby the winsorized value resulted in a nonlinear pattern, deviating from the original model estimation of a linear pattern; thus, the use of the non-winsorized value of desirability was retained).

1. **Supplemental Figures**

**Chart, histogram

Description automatically generated**

Figure S1. Histogram of Age distribution.

Graphical user interface

Description automatically generated with low confidence

Figure S2. Example illustration of smoothed interaction terms in Generalized Additive Mixed-effect Models. Smoothing spline regressions use a data-driven approach to estimate nonlinear trajectories and are internally cross validated to avoid overfitting. A. For interactions between continuous and categorical predictors (e.g., the analysis of a categorical variable with two levels A and B by age), this approach fits a nonlinear function for each categorical level. B. For interactions between two continuous predictors (e.g., analyses of a continuous variable X by age), we used a smoothing tensor, which fits a 3-dimensional functional plane to characterize nonlinear patterns of continuous interacting variables.

**Graphical user interface

Description automatically generated**

Figure S3. Effort exertion for peer feedback varies by acceptance expectancy. Across all indices of effort exertion, participants exerted greater peak grip force (blue; A), grip speed (green; B), and opted out less (red; C), for peer feedback when they had greater expectations of being liked. For outcome measure peak grip force (blue; A), participants also exerted greater effort for peer feedback when participants believed they would be rated very unfavorably. Figures display the mean-centered fitted effect of the main effect of acceptance expectancy from the GAMM, with effort exertion outcome measures (peak grip force in units of % of maximum grip; speed in units of *log*(ms), and opt-out behavior in units of probability) along the y-axis and acceptance expectancy (z-scored) along the x-axis. Shading indicates standard error of the mean.

**Graphical user interface

Description automatically generated**

Figure S4. Effort exertion for peer feedback varies by peer desirability. Across all indices of effort exertion, participants exerted greater peak grip force (blue; A), grip speed (green; B), and opted-out less (red; C), for peer feedback from peers they rated more highly. Figures display the mean-centered fitted effect of the main effect of peer desirability, with effort exertion outcome measures (peak grip force in units of % of maximum grip; speed in units of *log*(ms), and opt-out behavior in units of probability) along the y-axis and peer desirability (z-scored) along the x-axis. Shading indicates standard error of the mean.

A picture containing text

Description automatically generated

Figure S5. Raw ratings of peer desirability and acceptance expectancy across age. Compared to younger participants, older participants expected to be rated more favorably by peers (A) and rated peers more favorably (B). Raw ratings are displayed along the y-axis and age along the x-axis. Display shows linear trendlines and standard error of the mean in the shaded region.

Graphical user interface, application

Description automatically generated

Figure S6. Raw ratings of peer desirability and acceptance expectancy broken down by participant and peer gender across age. A. When considering acceptance expectancy, female participants expected to be liked more favorably by female peers than male peers and this difference diminished with age. Male participants showed a similar, albeit mitigated pattern during adolescence, but this pattern reversed in young adulthood where male participants expected to be rated more favorable by female peers than male peers. B. When considering peer desirability, female participants rated female peers more favorably than male peers, and this was stable across age. Male participants, on the other hand, rated other male peers more favorably than female peers at younger ages, but this pattern reversed from mid adolescence to adulthood. Raw ratings are displayed along the y-axis and age along the x-axis. Display shows linear trendlines and standard error of the mean in the shaded region.

Icon

Description automatically generated with medium confidence

Figure S7. Age-effects of social vs. nonsocial targets on effort exertion including random slope of Task. Figures demonstrate the relative difference in marginal effects of effort exertion for money compared to social targets across age. Similar to the original findings, adults showed either equivalent (peak, A; opt-out, C) or greater motivation (speed, B) for money compared to peer feedback, whereas adolescents showed the reversed pattern. The x-axis displays continuous age and the y-axis displays the difference in marginal effects of effort exertion for social vs. money tasks (peak grip force in units of % of maximum grip; speed in units of *log*(ms), and opt-out behavior in units of probability). Shading indicates 95% Confidence Interval and dotted lines identify sections where the 95% CI does not contain zero.

1. **Supplemental Tables**

****

**Interaction effects derived from 95% CI plots are consistent with findings in the manuscript**

****

****

****

**References:**

Arulpragasam, A. R., Cooper, J. A., Nuutinen, M. R., & Treadway, M. T. (2018). Corticoinsular circuits encode subjective value expectation and violation for effortful goal-directed behavior. *Proceedings of the National Academy of Sciences*, *115*(22), E5233–E5242. https://doi.org/10.1073/pnas.1800444115

Botvinick, M. M., Huffstetler, S., & McGuire, J. T. (2009). Effort discounting in human nucleus accumbens. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(1), 16–27. https://doi.org/10.3758/CABN.9.1.16

Guyer, A. E., Choate, V. R., Pine, D. S., & Nelson, E. E. (2012). Neural circuitry underlying affective response to peer feedback in adolescence. *Social Cognitive and Affective Neuroscience*, *7*(1), 81–92. https://doi.org/10.1093/scan/nsr043

Kurniawan, I. T., Seymour, B., Talmi, D., Yoshida, W., Chater, N., & Dolan, R. J. (2010). Choosing to Make an Effort: The Role of Striatum in Signaling Physical Effort of a Chosen Action. *Journal of Neurophysiology*, *104*(1), 313–321. https://doi.org/10.1152/jn.00027.2010

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520. https://doi.org/10.1007/s00213-006-0502-4

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., & Frith, C. D. (2007). How the Brain Translates Money into Force: A Neuroimaging Study of Subliminal Motivation. *Science*, *316*(5826), 904–906. https://doi.org/10.1126/science.1140459

R Core Team. (2020). *R: A language and environment for statistical computing.* (3.5.2) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rij, J. van, Wieling, M., Baayen, R. H., & Rijn, H. van. (2020). *itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs* (2.4) [Computer software]. https://CRAN.R-project.org/package=itsadug

Rodman, A. M., Powers, K. E., Insel, C., Kastman, E. K., Kabotyanski, K. E., Stark, A. M., Worthington, S., & Somerville, L. H. (2021). How adolescents and adults translate motivational value to action: Age-related shifts in strategic physical effort exertion for monetary rewards. *Journal of Experimental Psychology: General*, *150*(1), 103–113. https://doi.org/10.1037/xge0000769

Salamone, J. D., Correa, M., Farrar, A. M., Nunes, E. J., & Pardo, M. (2009). Dopamine, Behavioral Economics, and Effort. *Frontiers in Behavioral Neuroscience*, *3*. https://doi.org/10.3389/neuro.08.013.2009

Wood. (2019). *Package “mgcv”: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation* (1.8-31) [Computer software]. https://cran.r-project.org/web/packages/mgcv/

Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *65*(1), 95–114. https://doi.org/10.1111/1467-9868.00374

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R, Second Edition*. CRC Press.