ADOLESCENT SOCIAL MOTIVES: MEASUREMENT AND IMPLICATIONS

by

JOHN C. FLOURNOY

A DISSERTATION

Presented to the Department of Psychology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

August 2018

DISSERTATION APPROVAL PAGE

Student: John C. Flournoy

Title: Adolescent social motives: measurement and implications

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Psychology by:

Jennifer Pfeifer	Co-Chair
Sanjay Srivastava	Co-Chair
Elliot Berkman	Core Member
Nicole Giuliani	Institutional Representative

and

Janet Woodruff-Borden	Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded August 2018

© 2018 John C. Flournoy
This work, including text and images of this document but not including supplemental files (for example, not including software code and data), is licensed under a Creative Commons
Attribution 4.0 International License.

DISSERTATION ABSTRACT

John C. Flournoy

Doctor of Philosophy

Department of Psychology

August 2018

Title: Adolescent social motives: measurement and implications

The study of decision making during adolescence has received considerable attention throughout the history of developmental psychology, justifiably, given the marked increases in morbidity and mortality that belie otherwise robust health. Although the dominant theories invoked to help explain decision-making during adolescence acknowledge the existence of motivations that are thought to be central to this developmental period, there is little work that investigates the effects of these motives, per se. In particular, motivations toward developing sexual and romantic relationships, as well as finding one's optimal position in peer status hierarchies have both been acknowledged as especially relevant for this period of development. Almost all research in this area focuses on self-report, and is heavily weighted toward the domain of status and popularity. A major gap in this literature is an understanding of how adolescent-relevant motivations affect basic behavioral processes, and of the consequences of individual differences in motivations.

The current investigation uses reinforcement learning to examine the effects of social motives on stimulus salience. This may allow both indirect, behavioral measurement of motivations, and is itself a potential mechanism by which motivations affect behavior via experience of the environment, and learning. Adolescent (N = 104) and college student (N = 230) participants learned four social-motive-relevant, and two baseline face-word associations. Learning was characterized using both proportion of optimal responses in the last half of the learning task, and a Rescorla-Wagner-like computational model. Results showed greater learning, and higher learning rates, in the social-motive conditions.

In order to explore the validity of behavior on the task as a measure of particular motivations, individual learning differences between social and baseline conditions were compared with developmental indices, self-report traits, and self-report health-relevant behaviors. Older participants were better at the learning task, but social-motive learning enhancement was constant across development. Measures of

social-motive effects on learning did not correlate with self-reported traits or health-related behaviors. The effects of motive-relevant words on learning may be due to factors unrelated to motivation, but research design may also be problematic. Self-report trait instruments performed well, but a more comprehensive taxonomy of motivational constructs and measures would be beneficial.

CURRICULUM VITAE

NAME OF AUTHOR: John C. Flournoy

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR, USA

University of California, Berkeley, CA, USA

DEGREES AWARDED:

Doctor of Philosophy, Psychology, 2018, University of Oregon

Master of Science, Psychology, 2014, University of Oregon

Bachelor of Arts, Cognitive Science, 2004, University of California, Berkeley

AREAS OF SPECIAL INTEREST:

Adolescent development

Development of individual differences

Quantitative methodology

PROFESSIONAL EXPERIENCE:

Graduate Research/Teaching Fellow, University of Oregon, 2012–2018
Research Coordinator, Stanford University School of Medicine, 2009–2012

GRANTS, AWARDS AND HONORS:

Gary E. Smith Summer Professional Development Award, University of Oregon, 2015
Clarence and Lucille Dunbar Scholarship, University of Oregon, 2014

PUBLICATIONS:

Thalmayer, A.G., Saucier, G., Flournoy, J. C., & Srivastava, S. (under review). Ethics-Relevant Values as Antecedents of Personality Change: Longitudinal Findings from the Life and Time Study. *European Journal of Personality*.

Thalmayer, A.G., Saucier, G., Srivastava, S., Flournoy, J. C., & Costello, C. K. (revised and resubmitted). Ethics-Relevant Values in Adulthood: Longitudinal Findings from the Life and Time Study. *Journal of Personality*.

de Macks, Z. A. O., Flannery, J. E., Peake, S. J., Flournoy, J. C., Mobasser, A., Alberti, S. L., ... & Pfeifer, J. H. (2018). Novel insights from the Yellow Light Game: Safe and risky decisions differentially impact adolescent outcome-related brain function. *NeuroImage*. 10.1016/j.neuroimage.2018.06.058

Flannery, J. E., Giuliani, N., Flournoy, J. C., & Pfeifer, J. H. (2017). Neurodevelopmental Changes Across Adolescence in Viewing and Labeling Dynamic Peer Emotions. *Developmental Cognitive Neuroscience, 25*, 113-127. 10.1016/j.dcn.2017.02.003

Giuliani, N. R., Flournoy, J. C., Ivie, E. J., Von Hippel, A., & Pfeifer, J. H. (2017). Presentation and Validation of the DuckEES Child and Adolescent Dynamic Facial Expressions Stimulus Set. *International Journal of Methods in Psychiatric Research, 26(1)*, e1553. 10.1002/mpr.1553

King, K. M., Littlefield, A., McCabe, C., Mills, K. L., Flournoy, J. C., & Chassin, L. (2017). Longitudinal modeling in developmental neuroimaging research: Common challenges, and solutions from developmental psychology. *Developmental Cognitive Neuroscience*. 10.1016/j.dcn.2017.11.009

Leonard, J. A., Flournoy, J. C., Lewis-de los Angeles, C. P. & Whitaker, K. J. (2017). How much motion is too much motion? Determining motion thresholds by sample size for reproducibility in developmental resting-state MRI. *Research Ideas and Outcomes, 3*, e12569. 10.3897/rio.3.e12569

Madhyastha, T., Peverill, M., Koh, N., McCabe, C., Flournoy, J. C., Mills, K. L., King, K. M., Pfeifer, J. H., & McLaughlin, K. A. (2017). Current methods and limitations for longitudinal fMRI analysis across development. *Developmental Cognitive Neuroscience*. 10.1016/j.dcn.2017.11.006

Matta, T., Flournoy, J. C., & Byrne, M. (2017). Making an unknown unknown a known unknown: Missing data in longitudinal neuroimaging studies. *Developmental Cognitive Neuroscience*. 10.1016/j.dcn.2017.10.001

Flournoy, J. C., Pfeifer, J. H., Moore, W. E., Tackman, A. M., Masten, C. L., Mazziotta, J. C., ... Dapretto, M. (2016). Neural Reactivity to Emotional Faces May Mediate the Relationship Between Childhood Empathy and Adolescent Prosocial Behavior. *Child Development, 87(6)*, 1691–1702. 10.1111/cdev.12630

Goodkind, M. S., Gallagher-Thompson, D., Thompson, L. W., Kesler, S. R., Anker, L., Flournoy, J., ... O'Hara, R. M. (2015). The impact of executive function on response to cognitive behavioral therapy in late-life depression. *International Journal of Geriatric Psychiatry, 31(4)*, 334-339. 10.1002/gps.4325

O'Hara, R., Marcus, P., Thompson, W. K., Flournoy, J., Vahia, I., Lin, X., ... Jeste, D. V. (2012). 5-HTTLPR Short Allele, Resilience, and Successful Aging in Older Adults. *American Journal of Geriatric Psych, 20(5),* 452-456. 10.1097/JGP.0b013e31823e2d03

ACKNOWLEDGEMENTS

The work contained herein may be understood as but a byproduct of the process of coming to understand two important lessons: that psychological science profoundly shapes the public discourse around perennially vulnerable groups; and that paths to knowledge in this science are beset on all sides by demons of obfuscation. No instrument can measure the immensity of my gratitude to my advisors, and mentors, who led me through these lessons. Jennifer Pfeifer was a profound and subtle guide as I sifted through a multiplicity of substantive interests, and then provided a rich environment where I could productively explore the content that eventually lead to this project; she was an editor of writing and concept par excellence; and she honed well my nascent, stubborn, predilections for new methods and new scientific culture. Sanjay Srivastava somehow laid bare, with calm, humor, or fury, as appropriate, vast sections of the machine we use to try to know things in this corner of science – how much of the machine is still hidden, I don't know, but he also provided me tools to keep meddling about; and he showed me how to keep a light on, and a true path set, just as dark clouds appeared on the horizon. I'm grateful to Elliot Berkman, too, who has from the beginning encouraged my strengths, given support whenever asked, and provided clear and thoughtful discussion on problems in methodology, content, and professional development. And thanks to Nicole Giuliani who was willing to collaborate with me as an early grad student, who didn't mind my cantankerous fastidiousness, and who helped turn it toward useful ends. Kate Mills gave invaluable feedback when the plan for this project was just coalescing, shepherded its completion as each section began to take shape, but more vital than that, offered me professional camaraderie in that liminal space between mentor and peer. I would also like to thank Ruth O'Hara for giving me an opportunity to return to psychological research based on little more than my earnest conviction of interest in the field, and for subsequently, incidentally but not unintentionally, exposing me to didactic content for a full post-doctoral training program before I had even applied to graduate school. I would not have arrived at this point so easily without that kindness.

My appreciation is everlasting for the numerous fellow graduate students who enhanced every aspect of my life and work. My thanks goes to Rose Hartman, a role model who translated complexity into fluency, and who introduced me to Bayes. I am grateful, also, to Rose, Nicole Lawless DesJardins, and Allison Tackman for teaching me statistics; and to Rose and Nicole for establishing and cultivating R Club, where once a week or so we could all spend an hour basking in scientific computing with other methods mavens. Thanks to Will Moore, who taught me everything he could about neuroimaging, and who saw its problems with clear eyes. I owe Shannon Peak my thanks for suggesting, with a knowing grin, that I might like computational models of cognition. I am deeply grateful to Dani Cosme for her adroit intellect as we delved into countless conversations that shaped both how I understand development during adolescence and why I would want to; for conversations on so many other topics; and for inspiring pragmatic idealism by example. I am thankful for Arian Mobasser's friendship, and for the tremendous structure of community and wellbeing he built that sustained me and many others during this process. I am also thankful for Jimena Santillán's friendship, her industrious example, her advice on how exactly to start and finish this thing, and for her ear, which somehow makes things clearer.

I owe Lori Olson an immense debt of gratitude for all the administrative support starting from when I submitted my application, and lasting until the last form was signed and submitted. She made it possible to complete this project amidst the continuation of the rest of life.

I am not the author of this dissertation but for the love and friendship of these coauthors of my life: Thanks to my sister, Michele Flournoy, who encouraged me to tag along with her on a career-changing adventure that reminded me how fun and rewarding concrete abstractions can be. Thanks to Andrei Boutyline, with whom I've hunted all manner of thought through nearly half of all conceivable terrains, for casually beckoning me through the many doors scattered about the hallways of the academy. Thanks to Daniel Ashby, with whom I started wondering about anything one might call philosophy or wisdom, and who has always been a companion for wandering about. Thanks to Rick Wood for every trip to the back shelf, bottom row of the middle school library where they kept the books about esoteric, and arcane aspects of the mind, and the years of bus-stop walk and Scouting outing conversations, all of which probably led us straight to Cog Sci 100. I am so grateful to my dad, Pete Flournoy, for showing me the justice in truth and necessity of critique; and I am so grateful to my mom, Sally Flournoy, for explaining (perhaps when I was a little too young) why something like $2x = 4$ is such an awesome concept, for advocating for my educational opportunities, and for showing me the necessity of optimism. Finally, to my wife, Melanie Berry, I want to express a level of gratitude which is as ineffable as the fact of existence. Her love and support has been the necessary precondition that makes any of the rest of these acknowledgments possible. She has been a stalwart and incisive collaborator on every iteration of every idea that has passed between my ears in the last eleven years. Her own academic pursuits inspired me to undertake this project. Since she has known me, she has helped me understand what I am capable of doing, and she convinced me, at every juncture and against my worse judgment, that I could complete it.

To who you are, and will never be. To Ozymandias. But most of all, to my wife, Melanie Berry.

TABLE OF CONTENTS

Chapter Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Adolescence, the period of development between childhood and adulthood, is a time of both vigorous health and increased health risk. An extensive body of literature describes trajectories in the rates of health-risking behaviors like binge drinking (Chassin, Hussong, & Beltran, 2009; Patrick & Schulenberg, 2011), and risky sexual behavior (Stevenson, Zimmerman, & Caldwell, 2007) that, along with rates of unintentional injury, peak during late adolescence and early adulthood (Centers for Disease Control and Prevention, 2003; Willoughby, Good, Adachi, Hamza, & Tavernier, 2013). A major focus of the field of adolescent developmental psychology and neuroscience is to describe what gives rise to this phenomenon. A variety of theories current in this field may capture many features of adolescent cognition and behavior that have relevance for health-related decisions, but most do not focus on how such cognition and behavior may serve adaptive functions. Here, adaptivity is concptualized in the evolutionary sense, where reproductive fitness over the course of evolutionary history is the criterion. Taking a more explicitly functional approach will help to strengthen and unify our understanding of behavior during this transition into adulthood.

One early attempt to explain adolescent decision making proposes that youth make decisions based on imperfect information, or exaggerated expectations of invincibility (Elkind, 1967). However, on most questions, adolescents show no particular bias toward over or underestimating probabilities, though their estimation is not perfect; the probabilities they estimate also generally correlate with prospective outcomes (de Bruin, Parker, & Fischhoff, 2007; Fischhoff et al., 2000; Fischhoff, 2008). Adolescents do, however, overestimate their probability of dying (Fischhoff, de Bruin, Parker, Millstein, & Halpern-Felsher, 2010), which may be functionally important if it enhances the value of immediate rewards, especially if those rewards are relevant for thriving and reproductive fitness in short-term time frames.

Some more recent approaches use evidence from neuroscience to describe an imbalance between action-promoting, and action-inhibiting processes that is thought to be a cause of risk-taking behavior. Two distinct theories are variations on this idea. The dual systems model suggests that during adolescence, the system that encodes socially and emotionally rewarding stimuli increases its sway on behavior more rapidly than cognitive control does (Shulman et al., 2016). The imbalance model proposes that the miscalibration is not just between two systems, but rather that an optimal balance between multiple systems is still developing (Casey, Galván, & Somerville, 2016), and that this global imbalance leads to risky behavior. No approach in this vein deals directly with why certain behaviors are more likely to be selected by the psychological processes or neural systems that hold an excess of power during adolescence.

Two other theories do address certain functional aspects of adolescent decision making. Social reorientation theory suggests that during adolescence, individuals shift their dominant social network from the family to peers (Nelson, Jarcho, & Guyer, 2016). This shift is in part driven by developmentally appropriate changes in the salience of interactions with peers and potential romantic partners, which helps motivate adolescents to enter adulthood. Health-risking behavior increases are seen, in part, as a by-product of differences between what the family/caregiver network values and what the new peer group values, though the reasons for these value differences are not fully explicated.

Fuzzy trace theory also has relevance for explaining the function of possible decreased sensitivity to uncertainty (i.e., risk) during adolescence. This theoretical framework posits that adolescents have yet to develop heuristics that tend to govern more experienced, adult decision making, and so are more willing to explore the full problem space (perhaps, in part, to develop these adult heuristics; Romer, Reyna, & Satterthwaite, 2017). If some possible outcomes of this exploration are negative health-consequences, then this exploration obviously constitutes health-risking behavior. The fuzzy trace perspective, and related views, hold that this exploration is an essential developmental task in adolescence, allowing adolescents to build expertise and wisdom about the world they must engage with as adults (Blakemore & Mills, 2014; Crone & Dahl, 2012; Romer et al., 2017).

A model of decision-making borrowed from economics can help integrate functional perspectives, and does not require problematic neurodevelopment to explain adolescent-specific changes in risk behavior. One of the biggest challenges for imbalance models (but not the two, more functional approaches) is that adolescent health risking behavior may involve either impulsive *or* highly cognitively controlled behaviors. The value-based choice view goes beyond the ideas of competing systems, describing decision-making as the integration of expected-value inputs from myriad sources that culminates in the selection of behaviors with large expected values (Pfeifer & Berkman, 2018). The expected-value inputs may come from basic appetitive processes like desire for food or sex, as well as from self concept, social norms, peer expectations, emotional states, and theoretically any other sources that have influence over decision-making. From the vantage point of this perspective, it becomes clear that adolescent refinement of motivations for fitness-relevant behaviors related to status and mate-seeking could have a significant impact on value computations. In other words, if the function of adolescence is to complete the development of the organism to the fully adult form, motivations related to adult social concerns like status and mate-seeking are likely to have increasing weight as inputs to a value-based choice process.

In order to understand health-risking behavior, and decision making during adolescence more generally, it is important to begin to try to understand the motivational inputs that are related to evolutionarily relevant developmental tasks. As is described by fuzzy trace theory, adolescents do not yet have decision-making heuristics in place regarding their behavior in these motivational domains, and so they must explore, often using a decision-making approach that makes more fine-grained use of outcome probabilities (Romer et al., 2017). Differences in various motivations will increase or decrease the expected value of both exploration and exploitation behaviors. A substantial body of research documents broad motivational changes in adolescence as captured by the sensation-seeking construct (Harden & Tucker-Drob, 2011; Steinberg et al., 2008; Ruch & Zuckerman, 2001; Zuckerman, Eysenck, & Eysenck, 1978), but very little work examines adolescent motivation at a finer resolution. Since learning is the process by which decision-making is influenced by previous experience gained during exploration, examining learning processes directly may be able to provide a window into effects of motivations on decision-making during adolescence.

Motivations from an Evolutionary Psychology Perspective

The impact of evolutionarily relevant motivations should be considered especially carefully because selection pressure is likely to be a powerful force shaping biological predispositions, and individual differences, during this phase of development (Baltes, 1997; Mõttus et al., 2018). Relative to other primates, human childhood is extremely protracted, essentially representing a 'pause' in development, with coincident more-rapid post-pubertal development (Del Giudice, Angeleri, & Manera, 2009). From a birds-eye-view, this places the psychology and behavior of the adolescent phase at a transition point between a relatively quiescent period, and period of acceleration into an adult motivational frame. Particularly striking changes during this period of development include increasing propensity for adult-like behaviors, such as the inception of sexual and romantic behavior (Cavazos-Rehg et al., 2009; Harden, 2014), and navigation of shifting social networks and status hierarchies (Cairns, Leung, Buchanan, & Cairns, 1995; Steinberg &

Morris, 2001) with possible consequences for adult success (Shi & Moody, 2017). As adolescents develop more adult-like motivations for mate-seeking and status, they also begin to gain access to adult-like means for enacting those motivations, many of which come with health-relevant risks.

To some extent, the impact on adolescent behavior of motivations for status and mates has been either acknowledged or studied. With regard to sex, authors often acknowledge that the transition to sexual maturity is a defining biological feature of adolescence,. However, little contemporary adolescent research investigates directly the increased motivation to explore and engage in romantic and sexual behavior (Suleiman, Galván, Harden, & Dahl, 2017), though this may be in part due to the political sensitivity of this topic (Harden, 2014). Status is studied more directly, with many articles focusing on the antecedents of popularity (LaFontana & Cillessen, 2002; Li, Xie, & Shi, 2012; Owens & Duncan, 2009; Rose, Swenson, & Waller, 2004), as well as on possible consequences (Allen, Porter, McFarland, Marsh, & McElhaney, 2005; Cillessen, Mayeux, Ha, de Bruyn, & LaFontana, 2014; Meuwese, Cillessen, & Güroğlu, 2017). Even so, though there is a long history of research on motivation in psychology, very little contemporary adolescent research situates mate-seeking and status in a broader motivational framework.

The fundamental social motives (FSM) framework proposed by Kenrick, Neuberg, Griskevicius, Becker, and Schaller (2010) may be helpful for understanding differential effects of both domains, as it parses motives into a small set based on adaptive problems humans are likely to face. These domains might be said to be fundamental in the sense that they are, plausibly, evolved psychological inputs on value computations during decision making processes that have relevance for survival and reproduction. The evolutionary-psychological framework is especially important for understanding adolescence, a developmental period that may be considerably sensitive to selection pressure (Baltes, 1997; Ellis et al., 2012; Suleiman & Harden, 2016). Using the FSM framework allows integration of findings from this work into the broader evolutionary psychology conversation that is beginning to arise in adolescent research.

Another advantage of the FSM approach is that it parses motivations by the functional importance of the motivation with regard to likely adaptive problems. There are many traditions (Wiggins, 1991) that would place status motives and mate-seeking motives within one of two broad domains, respectively, such as agency and communion (Bakan, 1966), power and intimacy (McAdams, 2009). However, those traditions are not able to capture the distinct challenges inherent in various social relations, which map onto important relationships at different developmental stages (specifically, mate-seeking as a result of puberty; also, for example, kin care in later adulthood). The fact that the FSM approach organizes the measurement of motivation around broad relationship categories, may make it especially useful for thinking about adolescence, which is proposed to be marked by a restructuring of an adolescent's place in the social world (Nelson, Leibenluft, McClure, & Pine, 2005; Nelson et al., 2016).

Literature on risk-taking behavior directly implicates both mate-seeking and status motivation domains in risky sexual behaviors and a variety of potentially status-related health-risking behaviors such as substance use. For example, in a study which assessed perceptions of the sexual-self among adolescents, there was an association between increasing sexually-motivated self-perceptions and various forms of sexual risk-taking (Buzwell & Rosenthal, 1996). With regard to status, there is some evidence consistent with the idea that adolescents who have high popularity as rated by peers may try to maintain their status through health-risking behaviors (Mayeux, Sandstrom, & Cillessen, 2008). Beyond health-risking behaviors, these motives may also be crucial ingredients in understanding a whole class of risk-taking that is more positively valued, including but not limited to risks taken in service of extending one's social and physical environment beyond one's childhood home (Sercombe, 2014). It is unfortunate that although the above literature acknowledges motivation, and measures related constructs, mate-seeking and status *motives* are rarely a direct target of measurement. There is a need for research investigating how adolescent learning and decision-making is shaped in the presence of stimuli which are motivationally relevant to their developmental stage.

Status motive. Status during adolescence has been studied from a variety of perspectives, including perceptions, structure, strategies, and outcomes (Allen et al., 2005; Jarvinen & Nicholls, 1996; LaFontana & Cillessen, 2010; Savin-Williams, 1979; Selfhout et al., 2010; Underwood, Kupersmidt, & Coie, 1996). Although the motivation to gain status (as distinct from social status, per se) has been addressed, the literature on status-related social goals typically relies on self-report. For example, using self-report measures, LaFontana and Cillessen (2010) found that prioritizing status over other domains (e.g., achievement, friendship) peaks in late adolescence. This kind of work often also employs peer-nomination of popularity or sociometric methods, and has generally found that self-reported motivation to pursue popularity, or related agentic or dominance goals, is only weakly associated with social status from middle childhood (Rodkin, Ryan, Jamison, & Wilson, 2013) through adolescence (Caravita & Cillessen, 2012; Cillessen et al., 2014; Jarvinen & Nicholls, 1996; Ojanen, Grönroos, & Salmivalli, 2005). This highlights the important distinction between an intrinsic motivation to learn about and navigate status hierarchies, and the ability (or lack thereof) to successfully actualize that motive. This same literature also implicates both popularity per se, *as well as* status-related social goals, in negative behaviors like agression (Cillessen et al., 2014; Ojanen et al., 2005) and bullying (Caravita & Cillessen, 2012), indicating that it is crucial to study motives to fully understand processes and outcomes related to status in adolescence. However, reliance on self-report is not always optimal, especially in a domain with strong social norms about the value of having and attaining high status may plausibly contaminate these measures. It would be useful to be able to make use of a behavioral measure of status motives that is less susceptible to self-report biases.

Mate-seeking motive. Puberty marks the beginning of sexual maturation of an individual, and the beginning of the adolescent developmental period (Susman & Dorn, 2009). Though there has been extensive research on the importance of romantic and sexual relationships during adolescence, peer and parent influences, risky patterns of sexual behavior, and health-relevant outcomes (Capaldi, Stoolmiller, Clark, & Owen, 2002; Schalet, 2011; Suleiman & Deardorff, 2015), there has been very little research on sexual development as it relates to motivation. Indeed, while the study of risky sexual behavior is clearly important, there is so little research on the normative aspects of sexual development during adolescence (for example as a developmentally appropriate motivation) that a recent review called it “the ‘elephant in the room’ of adolescent … development” (Suleiman et al., 2017). Most work to date has focused on risk factors for unsafe sex unrelated to motivation (e.g., poverty, peer influence, parental monitoring; Carlson, McNulty, Bellair, & Watts, 2014; James, Ellis, Schlomer, & Garber, 2012; Santelli & Beilenson, 1992). This neglects the importance of sexual motivation (i.e., desire) in the development of an empowered and responsible sexual identity, impeding efforts to reduce health-risks through sexual education (Fine & McClelland, 2006). Although sexual desire is recognized rhetorically as an important component in emerging adolescent romantic and sexual behavior, its direct interrogation through either behavioral or self report measures is limited. Likewise, the role of mate-seeking motives in the exploratory behavior characteristic of adolescence is also under-studied. As with status, reliance on self-report may not be optimal, especially given the extremely sensitive nature of this topic both socially and politically.

Reinforcement learning

As noted above, research on status and mate-seeking motives during adolescence is either largely absent, or reliant on self-report. The field would benefit greatly from a new measure that provides an alternative method to capture individual and developmental differences in the salience of these motives across the transition from childhood through adolescence and into adulthood. A primary focus of this work is therefore the design and validation of a new behavioral measure to do so. The measure, described below, uses a standard reinforcement learning paradigm and experimentally manipulates the extent to which learning targets are relevant to different motivational domains. The basic reinforcement learning paradigm has been well described in both humans and rodents (A. G. E. Collins & Frank, 2012; Glimcher, Camerer, Fehr, & Poldrack, 2009; Jones et al., 2011). The salience of stimuli in such a paradigm directly impacts learning (Grossberg, 1975; Mackintosh, 1975; Cunningham & Brosch, 2012). The premise of this new measure is that linking a learning target to motivationally relevant information will affect the salience of that target in proportion to the strength of the motivation. By measuring how learning rates are potentiated by this motivational framing, it may be possible to measure a behavioral indicator of psychologically basic status and mate-seeking motives. Moreover, this would demonstrate that status and mate-seeking motives directly affect reinforcement learning, and so, perhaps, shape exploratory behavior, learning, and decision-making during adolescence.

In designing an appropriate reinforcement learning task, it is important to consider that the optimal parameters governing the accumulation of information will differ depending on the stability of the association between the response to a stimulus and reward receipt (assuming the association is probabilistic). For example, in the context of a task with a stable association, after some number of successful pairings, the learner should no longer be tempted to switch her response, thus ensuring she is not fooled by a random run of no-reward feedback. However, if the association is not stable, and does actually switch (as is the case in reversal learning paradigms), the optimal learner should have a lower threshold for switching her association between a particular response and expected reward. For the learning strategy captured by the the proposed task to have relevance for our understanding of real-world behavior, its association stabilities should roughly correspond to the stability of status-relevant and mate-seeking-relevant associations in the environment.

Both sociometric status and romantic relationships are fairly rank-order stable throughout adolescence. In a study of roughly 300 adolescents with high sociometric status, estimates of the year-to-year stability from 6th-12th grade ranged between .51 and .72, with stability across all except a single 1-year interval above .6 (Cillessen & Borch, 2006). The percent of adolescents reporting romantic relationships lasting 11 months or longer, an indicator of stability relevant for this task, increases from 20% at an age of 14 years, to roughly 60% at 16 years (W. A. Collins, 2003). Given the brief temporal scale of the proposed task, it is reasonable to assume that participants would not prioritize learning strategies that are especially sensitive to changing associations.

Aims

Motivations have been under-studied in efforts to understand changes in decision-making during adolescence, many of which have consequences for health and well-being. Relevant motivations that are often only peripherally identified in the literature are well captured by the evolutionary-psychological approach codified in the FSM framework. Mate-seeking and status motives are especially relevant for the adolescent development period. Because learning and exploration are thought to be a crucial aspect of the importance of behavioral changes in adolescents, examining these motives using reinforcement learning provides the opportunity to better understand the motives, themselves, through a behavioral method, and also to understand the impact of motives on a mechanism related to learning and exploration. Broadly speaking, the goal of this work is to investigate the appropriateness of this reinforcement learning task, as well as the ways in which these motives (measured in multiple ways) relate to attitudes and behaviors.

Aim 1a: Does framing reinforcement learning targets with (mate-seeking and status) motive-related words increase stimulis salience and potentiate learning? Relevance of the stimuli to individuals' underlying motives should affect the salience of stimuli, and thus also the rate at which individuals learn. On average, it is expected that information related to status and mate-seeking motives will be motivationally relevant, and will potentiate learning over and above information in a baseline condition.

Aim 1b: How does motive potentiation of learning covary with development? It is likely that both status and mate-seeking motives increase between childhood and adolescence. If differences in learning rates between stimuli with social and minimally-social relevance reflect underlying motives, then it is expected that learning potentiation due to motive-framing will increase from early to late adolescence.

Aim 2: Does the motive potentiation of learning relate to a constellation of self-report measures that theoretically ought to demonstrate convergence? Specifically, (i) what is the relation between motive potentiation and self-report social motives; (ii) what is the relation between self-report social motives and motive-congruent health-related outcomes; and (iii) what is the relation between motive potentiation and *motive-congruent health-related outcomes.* It is expected that there will be a positive correlation between self-report measures of motives and outcomes, and differences in learning between social-motive conditions versus the baseline condition.

CHAPTER II

OVERVIEW OF SAMPLE CHARACTERISTICS

Data from four samples are used in these analyses. The first two samples comprise foster-care-involved adolescents (FCA) and community adolescents (CA) who were recruited from the greater Eugene/Springfield, Oregon, area for an ongoing longitudinal study investigating decision-making related to health-risking behavior (funded by NIDA, P50 DA035763). The third and fourth samples consisted of college student young adults who completed the study for course credit either in person (CSYA sample), or online (CSYA-O sample). The full sample consists of 334 participants (118 male) with ages between 12.5 and 42 years (see Table 1 for full details). A single participant who was over the age of 25 years was excluded from analyses. The majority of participants in all samples identified as white, though in both college samples this majority did not exceed 60% (Table 2). Some participants in each sample did not provide data on all focal measures (see Table 3), and so sample sizes for each analysis will differ to some extent. Examining the Table 3, it is clear that the proportion of missing data is small.

Adolescent sample. The CA sample was recruited through Craigslist, flyers, classroom presentations, class field trips to the University of Oregon, and word of mouth. Foster-care involved youth were recruited in collaboration with Oregon Department of Human Services caseworkers. For the parent study, the target sample size for each subsample was 80 participants based on power-analyses for the primary fMRI measure at wave 1. The primary measures of interest for the purposes of this dissertation were collected at wave 2, approximately 18 months after wave 1. Inclusion criteria include age at wave 1 of between 11.0-17.9 years and English fluency. Exclusion criteria included diagnosed psychiatric, conduct, or developmental disabilities, and MRI contraindications. Although describing differences between adolescents with and without a history of childhood adversity is not an aim of this dissertation, combining

Table 1. Descriptive statistics on age, and PDS for each sample

			Age				PDS			
	N		Male		Female		Male		Female	
Sample	Male	Female	M	SD	M	SD	M	SD	M	SD
FCA	19	20	15.6	1.5	17.0	1.7	2.9	0.6	3.8	0.2
CA	29	36	15.5	1.5	15.6	1.7	2.9	0.6	3.5	0.4
CSYA	33	52	20.9	4.2	19.3	1.3	3.6	0.3	3.8	0.2
CSYA-O	37	104	19.8	1.4	19.3	1.2	3.5	0.5	3.8	0.3

FCA: foster-care-involved adolescents; CA: community adolescents; CSYA: college students; CSYA-O: college students, online. PDS: Pubertal development scale.

Table 2. Race/ethnicity information by sample

Percent (%)	All	FCA	CA	CSYA	CSYA-O
AmIndn	0.3	0.0	0.0	0.0	0.7
Asian	10.5	0.0	0.0	12.9	16.6
Black/AA	2.4	5.1	0.0	3.5	2.1
Latinx/Hsp	7.2	0.0	6.2	11.8	6.9
Pcfc Islnd	1.8	0.0	0.0	1.2	3.4
Other	6.9	15.4	13.8	2.4	4.1
White	68.0	74.4	76.9	68.2	62.1
Wht/Hsp	0.9	0.0	3.1	0.0	0.7

FCA: foster-care-involved adolescents; CA: community adolescents; CSYA: college students; CSYA-O: college students, online. AmIndn: American Indian and Alaska Native; Asian: Asian; Black/AA: Black or African American; Latinx/Hsp: Latinx/Hispanic; Pcfc Islnd: Native Hawaiian and Other Pacific Islander; Other: Other; White: White; Wht/Hsp: White/Hispanic.

Table 3. Missing data from each sample for age, PDS, and gender

		Age		PDS		Gender		Task	
Sample	N	N_{miss}	%	N_{miss}	%	N_{miss}	%	N_{miss}	%
FCA	39	-		1	0.3	-		13	3.9
CA	65	-		-		-		4	1.2
CSYA	85	-		7	2.1	-		3	0.9
CSYA-O	145	-		5	1.5	4	1.2	1	0.3
All	334	-		13	3.9	4	1.2	21	6.3

FCA: foster-care-involved adolescents; CA: community adolescents; CSYA: college students; CSYA-O: college students, online. PDS: Pubertal development scale.

across the two samples is intended to improve power and provide increased variability on measures of motives as well as attitudes and outcomes. Nesting of participants within sample is encoded in hierarchical linear models as appropriate.

Young adult sample. The CSYA and CSYA-O samples comprise late adolescents and emerging adults age 18-25 years recruited from the UO human subjects pool using the web-based SONA scheduling system. Students sign up for time slots most convenient to them without knowing study details beyond inclusion and exclusion criteria. There are no exclusion criteria for this sample. The target sample size for this sample was 160 for $>$ 90% power to detect a small ($\beta_{\text{std on y}} = 0.1$) interaction between motive condition, and a continuous covariate such as the self-report Fundamental Social Motives Inventory (described below), assuming an intraclass-correlation $>$ .25. Note that unlike the adolescent samples, the young adult sample provided their age as an integer (in years).

For the online young adult sample only, attention check items were included in the questionnaire to help assess whether individuals were participating in good faith. Participants were excluded from the study and not awarded class credit if their data conformed to a pattern that strongly indicated inattentive and random responding. The primary criterion for making this judgment was extremely brief task duration (total time < 18 minutes) combined with scores on the task that indicated chance-level responding (roughly around 500). This means that those who took a long time but showed little learning, and those who learned both well and quickly were not excluded. Secondary criteria used to confirm the pattern indicated by task data included erroneous scores on several attention check questions. These attention check items were also used to test for acquiescence bias on self-report scales that did not contain reverse-keyed items. A total of six individuals were excluded from the study for not participating in good faith. Participants were not otherwise excluded for not passing attention check items both because further inspection often showed that responding was *not* random with respect to some portions of the questionnaire, and because some authors have reported that eliminating participants based on attention check items introduces a source of demographic bias (e.g., Anduiza & Galais, 2016).

Pubertal Development. To assess pubertal development, all participants completed the Pubertal Development Scale (PDS; Petersen, Crockett, Richards, & Boxer, 1988), with the mean of five items used to indicate progress through puberty (range [1,4]). While the full sample includes a wide range of pubertal development, there is notable heterogeneity. It is not surprising that almost all participants in the college samples report that puberty is complete. Notably (though also unsurprisingly), females in the the adolescent samples report being further through puberty than males (Figure 1A), even though the mean age of males and females in these samples is roughly equivalent (Table 1). Most participants in the college samples have PDS scores indicating puberty is complete, or nearly completed (Figure 1A).

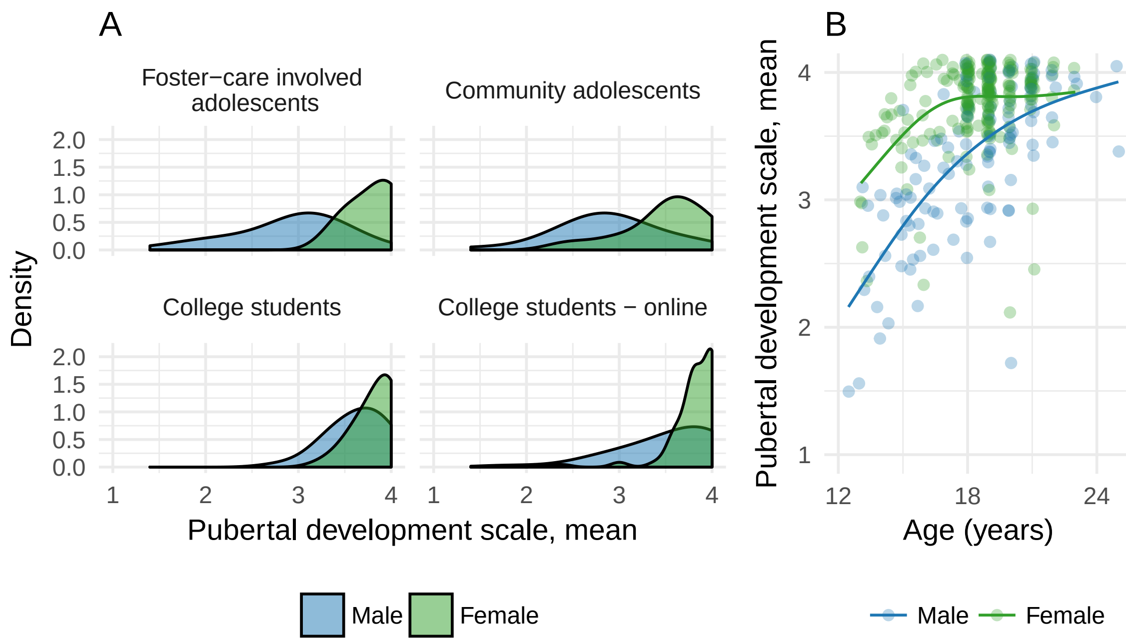

Figure 1. Pubertal development by sample and age. Distribution of pubertal development by sex and sample (A), and the relation between pubertal development and age (B). Best fit line is a generalized additive model for descriptive purposes only.

CHAPTER III

AIM 1: EFFECT OF MOTIVE FRAME ON REINFORCEMENT LEARNING

Background

Adolescence is a time of social reorientation during which individuals begin expressing romantic and sexual interests (for an overview, see: W. A. Collins, 2003; W. A. Collins, Welsh, & Furman, 2009), and spend more time with peers (at least, relative to time spent with family; R. Larson & Richards, 1991; R. W. Larson, Richards, Moneta, Holmbeck, & Duckett, 1996; Richards, Crowe, Larson, & Swarr, 1998). The social reorientation hypothesis in developmental cognitive neuroscience proposes that these changes are the result of developing motivations that shift goals toward increasing social experiences in peer and romantic social networks (Nelson et al., 2005; Nelson et al., 2016). It further proposes that neurodevelopmental changes in perceptual, motivational, and executive control brain regions are the mechanisms that cause transitions through sensitive periods in development in which these systems are tuned to relevant social experiences (both in the sense of being receptive to, and motivated toward them).

The identification of period-relevant social goals mirrors work in evolutionary psychology that has lead to a taxonomy of social motives which are broadly anchored by evolutionarily important social relationships (e.g., caregivers, peers, potential mates, mates, and offspring; for a thorough overview, see Neel, Kenrick, White, & Neuberg, 2015). This work identifies social status and mate-seeking as distinct social motives that are not merely aspects of (a perhaps earlier-developing) affiliation motive, and provides a complimentary framework for thinking about the distinct fitness challenges adolescents begin to face as they transition beyond the juvenile period. The development of these social motives, and their role in learning and decision making, is important to understand given their likely contribution to many types of decision making and, as suggested by Nelson et al. (2016), to psychopathology.

One implication of the reorientation thought to arise from changes in social motives is that there should be concomitant changes in the value and salience of motive-relevant stimuli. While it may be the case that pubertally triggered neurobiological changes begin to directly alter the salience of cues that are evolutionarily stable predictors of period-relevant rewards, it is also probably the case that environmentally constrained changes in reward contingencies also alter the value of certain cues. In other words, changes in stimulus-salience may arise from feedback between the congenital changes in the salience of particular stimuli, and learning that links those stimuli to particular outcomes. In adolescence, specifically, this should result in increases in the salience of stimuli relevant to the milieux of the peer group and of potential romantic partners (for example, stimuli related to social status, and dating availability, respectively).

Salience and learning. The reciprocal relation between stimulus salience, motivation, and learning is an important foundation for the idea that developmental changes in motivational systems should be responsible, in part, for the social reorientation thought to occur during adolescence. A substantial body of work on learning and attention has shown that more salient cues enhance learning, and that learned predictive value modifies cue salience. In laying out a programmatic approach to learning and attention, Mackintosh (1975) states learning rates are different depending on the stimulus because the "probability of attending to a stimulus determines the probability of learning about that stimulus" (p. 294). Grossberg (1975) summarizes a body of literature indicating that if two cues are paired with an unconditioned stimulus, the more salient cue will be conditioned more quickly, even to the extent that it may block the conditioning of the less salient cue.

Cue salience can both modulate, and be modulated by, blocking effects. For example, using cues of colored dots and modulating salience by the density of dots, Denton and Kruschke (2006) showed that it is harder to block a new cue if that new cue is more salient (i.e., has higher dot density) than the already conditioned cue. In other blocking paradigms, learning an association between a highly predictive cue and outcome blocks an organisms ability to form an association between that outcome and a new cue that is paired with the initially learned cue (see Shanks, 2010, for an overview). This has been interpreted as indicating that the "blocked" stimulus essentially becomes less salient than the initially learned stimulus. Though many models have been developed to account for this phenomenon, one explanation is that the new stimulus provides no new information for predicting the outcome (which is already perfectly predicted by the prior associated stimulus).

From an evolutionary perspective, the modulation of learning by salience, and salience by learning, is harmonious with both the idea that the environment is not perfectly stable, as well as with the idea that the goals, social or otherwise, of an organism change across development. There are at least two possibilities for why co-modulation of learning and attention may have arisen as a part of our cognitive apparatus. First, it may be a way to solve the problem of how to constrain perceptual capacity, and second, and more simply, it may just be the optimal strategy for navigating environments that are not perfectly stable. Hullinger, Kruschke, and Todd (2015) provide support for the second possibility via simulations of evolved learning systems. They found that attention evolves as an optimal strategy even in environments that did not overwhelm agents' perceptual capacity, but only when there was some amount of change in the value of different stimuli across the agents' lifespans. This gradual change of the diagnostic value of particular cues is very roughly analogous to what a developing organism might experience as their motives toward different kinds of social experience change. For this reason, reinforcement learning paradigms present an ideal (and idealized) paradigm for investigating changing motivations during development. Even though they have been tested extensively throughout the lifespan, we know of no studies investigating how social-motive-relevant information alter stimulus salience and learning in these paradigms.

Using reinforcement learning to investigate social motives. In reinforcement learning paradigms, incorporating social information that is relevant to adolescent-emerging social goals should enhance learning for cues that are associated with that information. Specifically, in this study, adolescents and young adults try to learn the association between targets (images of faces) and both minimally social descriptors (Hungry, Thirsty), as well as descriptors relevant to romantic behavior (Dating, Looking [for someone to date]), and to social hierarchy (i.e., status; Popular, Unpopular). We expect that our participants will differ in their motivation to learn the face-descriptor associations in proportion to the relevance of those descriptors to their social goals (e.g., mate-seeking, or status). We presume that knowing whether someone is dating, or looking for someone to date would, generally, be relevant information for someone who wants to engage in romantic or sexual behavior. We also presume that status information is relevant to people who find themselves navigating increasingly complex social hierarchies. Finally, we presume that learning about another person's hunger or thirst is not especially relevant to adolescent-emerging socially motivated goals. Past work has shown consistently that performance on reinforcement learning tasks improves with age (we return to this in the discussion, but see, e.g., Peters & Crone, 2017). For this reason, any developmental-motivation effect of stimulus salience on learning should manifest over-and-above the increase in performance that would be expected from early adolescence to young adulthood. Given the above, we hypothesized that, relative to younger participants, older participants would show faster learning of the mate-seeking and status descriptors relative to the minimally-social descriptors, and that pubertal status would also show this relation.

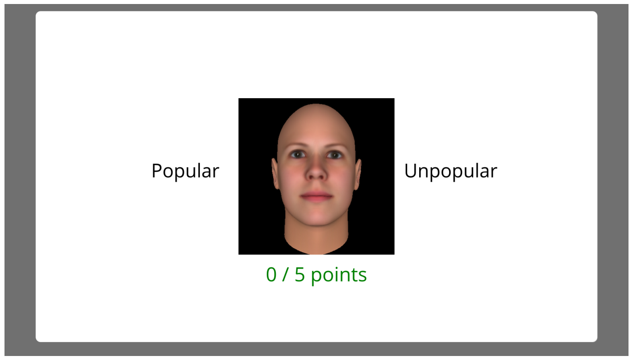

Figure 2. Example of a trial on the SPLT. Image shows post-response feedback phase on an incorrect trial where the participant could have earned 5 points for a correct answer, but earned 0 points.

Method

Social Probabilistic Learning Task. The Social Probabilistic Learning Task (SPLT) is a standard reinforcement learning paradigm using several stimulus-word pairings that are grouped by motive relevance. Common probabilistic reinforcement learning paradigms focus on estimating parameters that govern learning associations between stimuli using abstract images and nonsocial categories as the conditioned stimulus (e.g., a weather prediction of 'sunny' or 'rainy'; Knowlton, Mangels, & Squire, 1996). In the SPLT, motive context is manipulated by pairing computer-generated representations of human faces with state or trait words related to mate-seeking and status motives. The purpose of this manipulation is to examine how this motive framing alters learning, especially across adolescent development, as well as how individual differences in learning are related to the measures of attitudes and behavior outlined below.

On each trial, the participant sees one of six faces, along with two labels, and is asked to classify the face using one of the two labels. The labels are: *Hungry* versus *Thirsty* (baseline condition that is nonsocial, or only minimally social), *Dating* versus *Looking* for someone to date (mate-seeking), and *Popular* versus *Unpopular* (status). In order to help reduce task difficulty, "Hungry", "Dating", and "Popular" always appear on the left, while "Thirsty", "Looking", and "Unpopular" always appear on the right. After choosing the label, the participant is given feedback that they were correct and earned either 1/1 point or 5/5 points, or that they were incorrect and earned 0/1 point or 0/5 points (Figure 2). Each face is probabilistically associated with one label with P(Correct | Choice = Label) = .80. That is, there is a probabilistically optimal choice response, which precludes a memorization strategy and results in more automatic reinforcement learning (Knowlton et al., 1996). The participant is instructed that "the same word

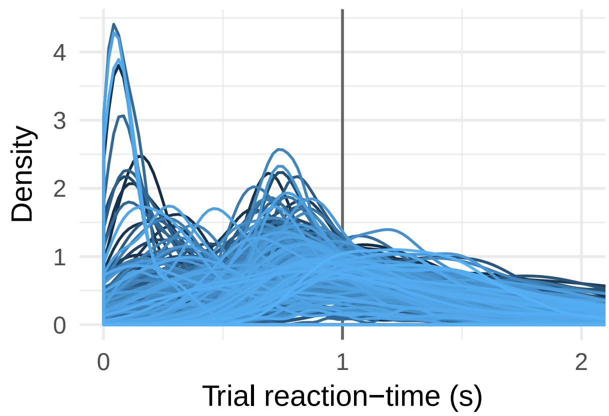

Figure 3. Distribution of reaction times for all participants. The vertical line is set at 1s, which is close to the mean reaction time across all samples.

goes with the same picture most of the time, but not always," and to "try to guess correctly as often as you can to get the most points."

To ensure engagement with the stimulus faces, 3 unambiguously male and 3 unambiguously female faces were drawn from a set of faces evaluated by Todorov (2008). Target faces were drawn from a subset of those highest in likeability and attractiveness ratings, and were selected to have roughly equal euclidean distance from one another on the other dimensions on which they had been rated (e.g., trustworthiness). Essentially, the faces were selected to be similarly salient, as well as similarly distinct from one another such that one or two particularly unique faces did not overwhelm the learning effect of the motives. On each run of the task (i.e., for each participant), one male and one female face was randomly assigned to each label within condition (baseline, mate-seeking, status). On each trial, the participant had 3.5 seconds to respond, and was shown response feedback for 1 second. The task comprises a total of 384 trials across 8 blocks.

The receipt of the reward for choosing the optimal response (that is, the response that most often generated a reward) was probabilistic. Across conditions and samples, the observed probability of reward receipt for an optimal choice was very close to the generative $p = .80$ (Table 4). Reaction times for all participants can be seen in Figure 3, with an overall mean of 1.06s (SD = 0.61s). Most participants completed all trials with only occasional missed responses (Table 5). For a small number of participants, far fewer trials were obtained (min = 294, with n = 7 having fewer than 350 trials) either as a result of early termination, or because of computer error during data file saving.

Table 4. Probability that an optimal choice results in reward

Condition	Sample	Probability of reward
Hungry/Thirsty	FCA	0.791
	CA	0.790
	CSYA	0.789
	CSYA-O	0.794
Dating/Looking	FCA	0.792
	CA	0.797
	CSYA	0.791
	CSYA-O	0.788
Popular/Unpopular	FCA	0.796
	CA	0.792
	CSYA	0.798
	CSYA-O	0.800

FCA: foster-care-involved adolescents; CA: community adolescents; CSYA: college students; CSYA-O: college students, online. PDS: Pubertal development scale.

Table 5. Number of trials completed

	N_{trials}			
	Male		Female	
Sample	M	SD	M	SD
FCA	378	4	379	4
CA	379	8	378	15
CSYA	379	8	379	8
CSYA-O	381	2	378	12

FCA: foster-care-involved adolescents; CA: community adolescents; CSYA: college students; CSYA-O: college students, online.

Modeling approach. The focus of this analysis is on describing and estimating differences in learning that are proposed to be related to differences in motives that develop during adolescence. The goal of such estimation is to provide reliable information about the magnitude and sign of relations between observed phenomena, as well as about the uncertainty of that information. A Bayesian approach provides a robust framework for specifying and estimating hierarchical, nonlinear models with many parameters while constraining inferences in a way that reduces the rate of making incorrect claims with confidence (Gelman, Hill, & Yajima, 2012; Gelman & Carlin, 2014). These are key concerns for the current work for two reasons. First, convergence of reinforcement learning model can be difficult to achieve, and Bayesian estimation helps by allowing one to constrain parameter values for these models in a principled and intuitive way using parametric prior distributions. Second, in contrast to a confirmatory experiment, this analysis seeks to explore relations between a number of theoretically related constructs; as such, the hierarchical Bayesian approach helps constrain estimates in ways that increase their efficiency (Gelman et al., 2012).

A model for reinforcement learning. In the context of this task, where the relation between the optimal response and the stimulus is constant, a simple model of the degree of learning could rely on a simple proportion of optimal responses P_{ok} for each condition k. The test of the hypothesis of the effect of framing would then be the difference between conditions in P_o. This simple model may sacrifice some precision for simplicity, and so we will also use a reinforcement learning model with several parameters that can account for deviations from a strict Rescorla-Wagner (RW) process. This increases the number of possible comparisons we are able to make between conditions, which may generate useful information about how motive-domain framing affects the learning process (conditional on the model), but which also increases the complexity of patterns between conditions and parameters that must be interpreted. It will be helpful to keep in mind that the framing can only be said to potentiate learning if, regardless of its effect on any model parameters, it also results in higher proportions of optimal responding.

The model used for these analyses is a Rescorla-Wagner model implemented by Ahn, Haines, and Zhang (2017), which they label as their go-no-go model number 2. Their original model handles binary decisions (button-press or no button-press) in response to four different cues. However, the form of the learning algorithm is generalizable to other binary choices in response to cues. The go-no-go models used by Ahn et al. (2017) were derived from work by Guitart-Masip et al. (2012). Their most flexible reinforcement learning model generates the probability of an action for each trial via five parameters: the learning rate, ε; the effective size of reinforcement, ρ; a static bias parameter, b; an irreducible noise parameter, ξ; and a Pavlovian learning parameter, π. In the SPLT, trial feedback does not vary by valence (responses result in reward, or no reward, but never punishment), so we use the model that does not include this Pavlovian component.

Reinforcement learning model for the SPLT. The pattern of an individual j's decision to press right ($a_{\rightarrow} = 1$) versus left ($a_{\rightarrow} = 0$) is distributed Bernoulli with $a_{\rightarrow j} \sim \text{Bernoulli}(\theta_j)$. Each individual's vector of θ is just the probability of pressing the right arrow key, $P(a_{\rightarrow t}|s_t)_t$, for each trial t, conditional on whatever stimulus s is present on that trial. This probability is calculated using a logistic transformation of the action weight for pressing the right arrow key minus the action weight for pressing the left arrow key. This probability is then adjusted by a noise parameter, $0 \leq \xi_{jk} \leq 1$, for each participant j during condition k. The noise parameter modulates the degree to which responses are random. When $\xi = 1$, $P_t = .5$, and because each individual has a unique noise parameter for each condition, this allows the model to account for participants who do not learn during any particular condition. The full equation is:

$$P(a_{\rightarrow t}|s_t)_t = \text{logit}^{-1}\big(W(a_{\rightarrow t}|s_t) - W(a_{\leftarrow t}|s_t)\big) \cdot (1 - \xi_{jk}) + \frac{\xi_{jk}}{2}.$$

The action weight is determined by a Rescorla-Wagner (RW) updating function,

$$W_t(a, s) = \begin{cases} Q_t(a, s) + b_{jk} & \text{if } a = \text{press } \rightarrow \\ Q_t(a, s) & \text{otherwise} \end{cases},$$

where b_{jk} encodes the degree of bias, for participant j during condition k, toward pressing the right arrow key ($\rightarrow$). The function Q encodes instrumental learning and is governed by the individual's learning rate for that condition, ε_{jk}, and the inverse temperature parameter, ρ_{jk}, that scales the effective size of the possible rewards $r_t \in \{0, 1, 5\}$:

$$Q_t(a_t, s_t) = Q_{t-1}(a_t, s_t) + \varepsilon_{jk}\big(\rho_{jk} r_t - Q_{t-1}(a_t, s_t)\big).$$

Hierarchical Parameters. Each parameter $(\varepsilon, \rho, b, \xi)$ varies by condition $k \in 1 : K$, and by participant $j \in 1 : J$ nested in sample $m \in 1 : M$. The structure of the hierarchical part of the model is the same for each parameter, so the following description for ε will serve as a description for all of the parameters. For each individual j, $\beta_{\varepsilon j}$ is a K-element row of coefficients for parameter ε for each condition:

$$\beta_{\varepsilon j} \sim \mathcal{N}(\delta_{\varepsilon mm[j]}, \Sigma_{\varepsilon})$$

where $\delta_{\varepsilon mm[j]}$ is a column of K means for individual j's sample M, as indexed in the vector mm, and Σ_ε is a $K \times K$ matrix of the covariance of individual coefficients between conditions.

Finally, across all M samples, the means for each condition k are distributed such that:

$$\delta_{\varepsilon k} \sim \mathcal{N}(\mu_{\varepsilon k}, \sigma_\varepsilon)$$

where $\mu_{\varepsilon k}$ is the population mean for parameter ε in condition k, and σ is a slightly regularizing scale parameter for these means across all conditions and samples. The priors for these final parameters are:

$$\mu_\varepsilon \sim \mathcal{N}(0, 5)$$

$$\sigma_\varepsilon \sim \text{exponential}(1).$$

All β are transformed to be on the appropriate scale before being used to compute θ_t. For ε and ξ, values are transformed to be on $[0, 1]$ using the fast approximation to the unit normal cumulative distribution function, $\varphi_{\text{approx}}(x)$. The parameter ρ is transformed to be on $[0, \infty)$ using the exponential function, e^x. The bias parameter b does not need any transformation.

Simulating data. Before modeling the task data, we confirmed that this model can recover known parameters from simulated data. Using RStan (version 2.17.3; Stan Development Team, 2018), we simulated 100 data sets based on the structure of the sample data, using the same number of participants per sample, as well as precisely the same task structure. For this analysis, it is important to be able to recover all $\mu_{\theta k}$ for $\theta \in \{\varepsilon, \rho, b, \xi\}$ and $k \in \{1, 2, 3\}$, where 1 = Hungry/Thirsty, 2 = Popular/Unpopular, and 3 = Dating/Looking. Based on interactive simulation (online at https://jflournoy.shinyapps.io/rw_model/; Flournoy, 2018b), reasonable parameter values for the control condition might be $\mu_{\varepsilon'} = -1.65$ and $\mu_{\rho'} = -0.3$ (note that these are the *raw* parameter values, on the scale the priors are defined, and which are transformed before being used in the learning model).

One early indication that a model may not be well suited to describing the data is that when generating from the prior distribution, simulated data-sets either do not adequately cover the range of reasonable values, or cover ranges that are implausible (Gabry, Simpson, Vehtari, Betancourt, & Gelman, 2017). The simulated data do generally cover the range of the actual data when we look just at the proportion of optimal presses over time (Figure 4), and importantly do not show implausible behavior (e.g., most estimates around extreme values like 0, 1, or .5).

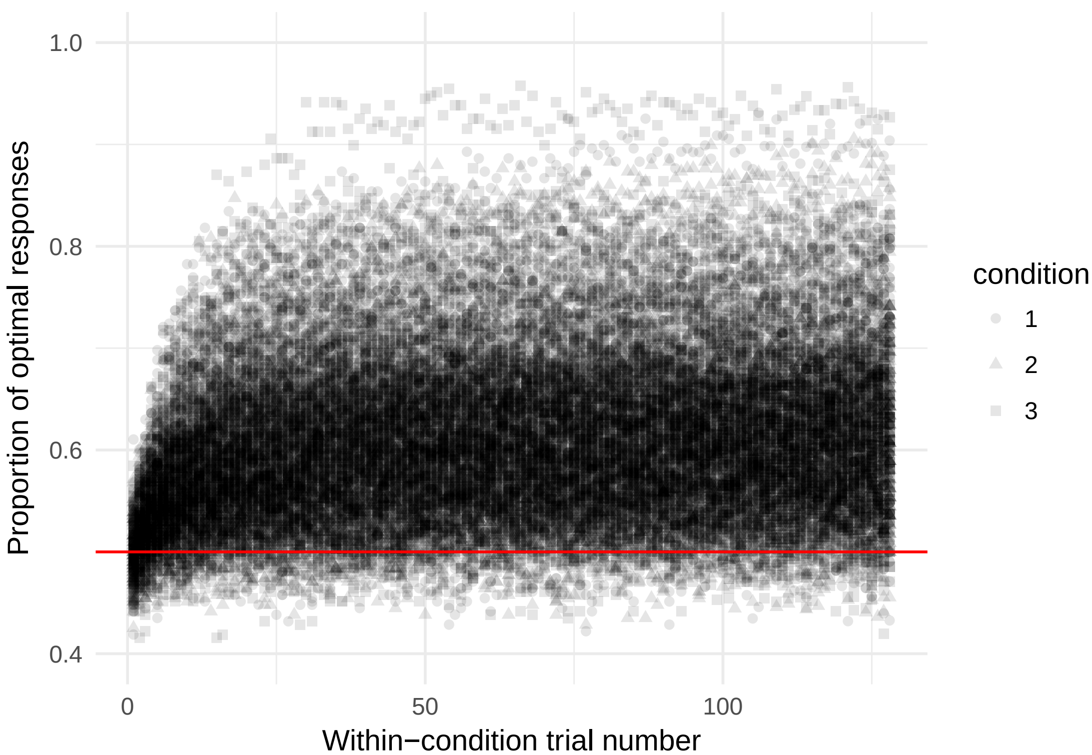

Figure 4. Model-simulated task data. Each point is the mean of optimal presses on that trial, for the indicated condition, from 1 of 100 simulated data sets using just the model prior distributions. The red line is at .5, corresponding to random responding. It is apparent that the model priors allow coverage of the entire range of realistic responses.

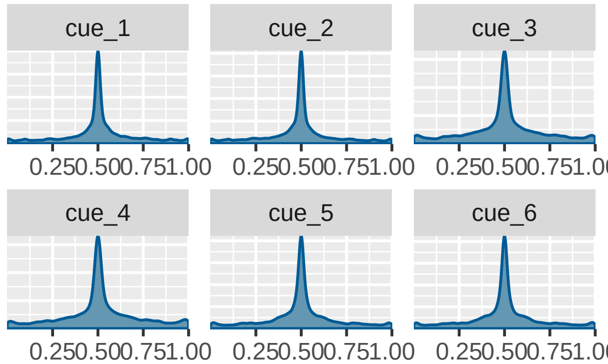

Figure 5. Prior density of task-end press-right probability. Across 100 simulated data sets, prior probabilities on learning-model parameters favor low rates of learning to discriminate optimal choices. There is, however, mass across the entire range of realistic values.

The priors also generate reasonable ranges for the model parameters across these 100 simulations. Notably, the mass of the prior distributions is distributed across the full range of reasonable values at each level and will tend to shrink posterior distributions toward zero if the data do not overwhelm the prior distributions. Ultimately, draws from the prior distributions transformed into predicted behavior based on the learning model described above generate a prior distribution over an agent's hypothetical final probability (at task-end) of choosing the word on the right-hand side of the screen. In a model that is well calibrated to this task, but that does not bias estimates, we should observe that these prior probabilities cover the full range, [0,1], for each of the six cues. These distributions cover the full range, with most mass around the null value of the final probability = .5 for choosing the left or right option (Figure 5). If the data does not overwhelm the prior distribution, we can be assured that the prior is not biasing the estimate away from what would be expected under the null that assumes participants do not learn, and respond randomly.

Recovery of population parameters. After generating these simulated data sets and known parameter values, the model was then fit to these simulated data to evaluate its ability to recover the parameters. The model was fit using 6 chains with 1200 warm-up iterations and 334 sampling iterations per chain (for a total of 2004 samples). An example from one simulation of estimated posterior densities for the final probability of a participant choosing the right-hand label shows that nearly all posterior 95% credible intervals capture the underlying generating parameter (Figure 6). This was also the behavior seen for individual parameters unless the simulated data came from a prior draw where the variation between individuals was extremely low, or a particular parameter (like the noise parameter, ξ) overwhelmed the other parameters in leading to the simulated behavior.

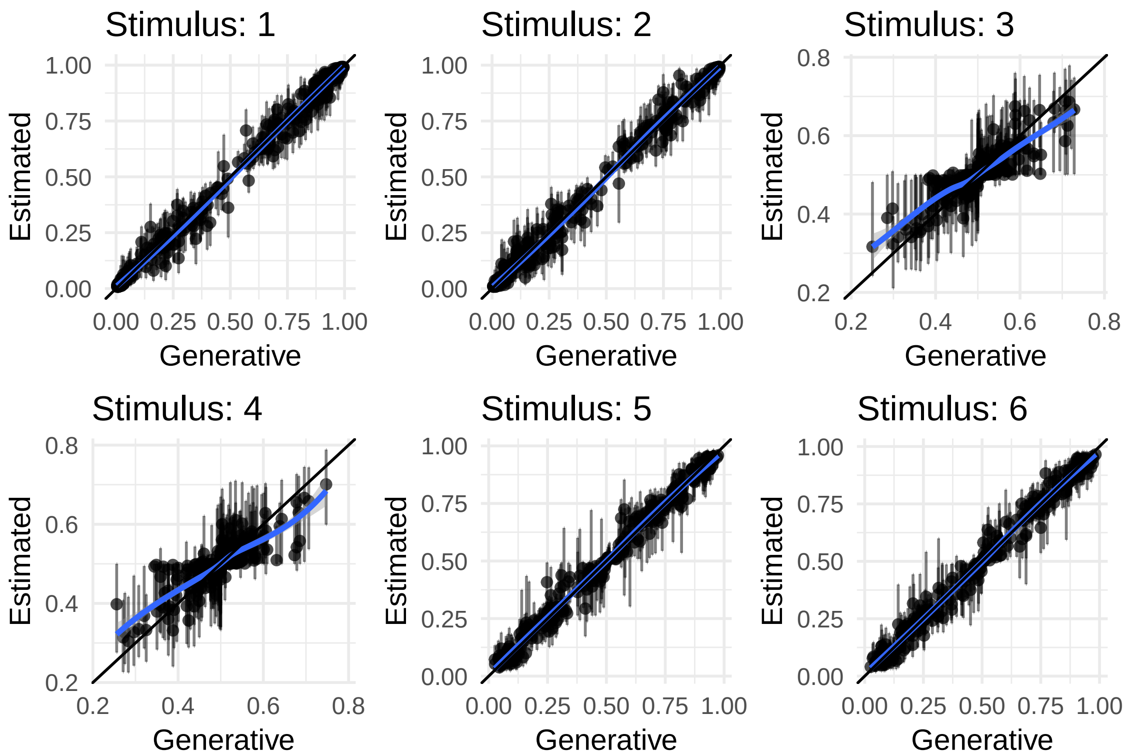

Figure 6. Correspondence of simulation generated and estimated behavior. The X-axis indexes the simulation-generated final-trial right-hand button-press probability ($P_{\text{r-final}}$), and the Y-axis indexes the estimated posterior median $P_{\text{r-final}}$. Whiskers indicate the posterior 2.5% and 97.5% quantiles.

While the above example provides some reassurance that the model performs adequately on simulated data, it is important to evaluate whether it can perform well across a number of simulated data sets. To examine the performance of all of the simulations, after the posteriors for each simulation estimate were sampled, the empirical cumulative density function was composed for the distribution of each parameter. The cumulative density *at the generating value* (from the simulation) was then found. This procedure is the same as finding the *p*-value for a test statistic using the probability density function for the relevant distribution (e.g., Student's *t*-distribution or the normal distribution). If the posterior is reasonable, then the generating parameter should be a random draw from that posterior, and if the posterior consistently tends toward excluding the generating value, we would expect the generating value to fall more often on one side or the other of the posterior. As such, the distribution of the *p*-value of the generating parameters (in relation to the posteriors) should be uniform (for the same reason that *p*-values are expected to be uniform if the null hypothesis is true). This is indeed what we find both from a visual inspection of the

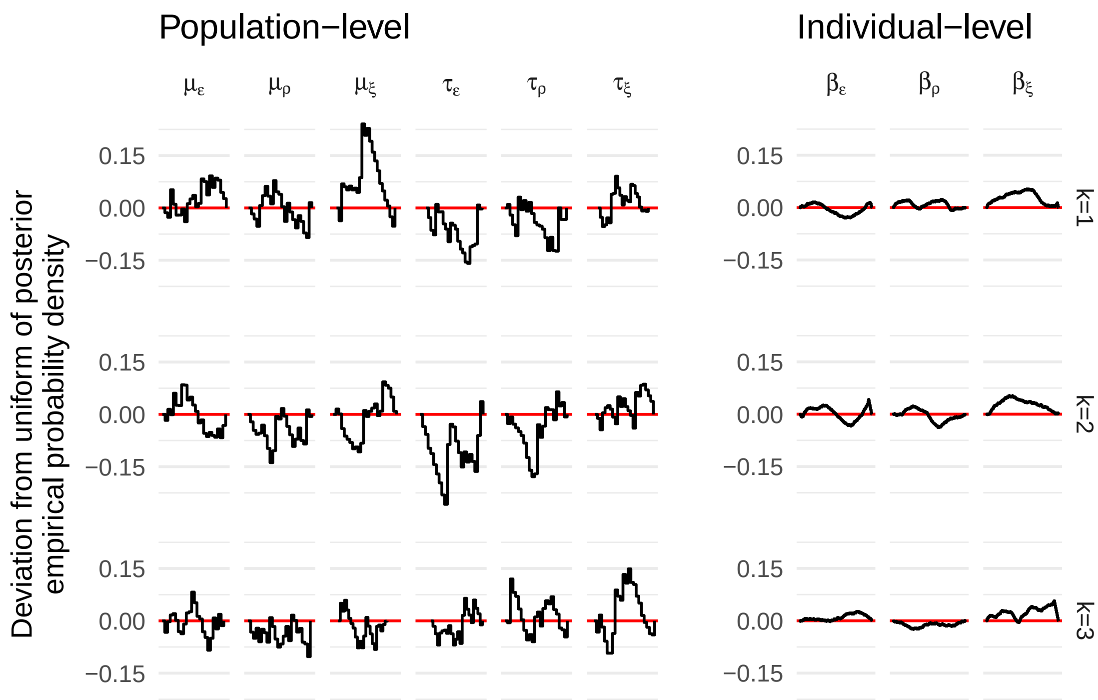

Figure 7. Posterior probabilities of generating values versus uniform. The empirical density of the posteriors probabilities of the data generating values (that is, the probability of the generating value assuming the posterior density is what it is drawn from) do not seem to deviate greatly or systematically from the cumulative uniform density function. Kolmogorov-Smirnov tests did not reject the hypothesis of no difference between these posterior probabilities and the uniform distribution on [0,1].

deviation of the cumulative density of the posterior probabilities of generating values (Figure 7) and from Kolmogorov-Smirnov tests, which do not reject the null hypothesis that there are no differences between the observed *p*-values and the uniform distribution on [0,1]. Note, however, that power to reject this null hypothesis (at $\alpha = .05$) may be somewhat limited for the estimated population parameters from just a small number (N = 100) of simulations.

Target model and estimation of sample parameters. All models were fit in RStan (version 2.17.3; Stan Development Team, 2018) using R (version 3.4.4; R Core Team, 2018). Samples were drawn (per model) over 1534 iterations from six chains, with the first 1200 iterations for each chain discarded as a warm-up period (total iterations per model = 2004). To compare models and choose the one with the best expected out-of-sample prediction we used the R package loo (Vehtari, Gabry, Yao, & Gelman, 2018) to compute an estimate of the leave-one-individual-out cross validation information criterion (Vehtari, Gelman, & Gabry, 2017).

From the winning model, we then extracted model parameters, as well as the final probability of choosing the optimal descriptor for a particular cue. Means of the posterior densities for all parameters are used in subsequent analyses. In addition, we used the average number of optimal choices in the first and last half of the run as a more easily calculable measure of optimal performance. Associations between all model parameters, measures of performance, age, and pubertal development were computed using Spearman's rank order correlation and standardized coefficients from linear mixed effects models. Linear mixed effects (LME) models were computed using the R package lme4 (version 1.1.17; Bates, Mächler, Bolker, & Walker, 2015). These LME models were computed in order to aid descriptions of plotted data and best-fit generalized additive model (GAM) lines, and were used in place of correlations to account for non-independence of data when, for example, a parameter for all three conditions was regressed on age. All such models have the form $y_{ij} = \gamma_{00} + \gamma_{10}\mathrm{x}_j + u_{0j} + e_{ij}$, where y is a target outcome like the learning rate ϵ, and x is a predictor like age or score on the pubertal development scale. The standardized γ_{10} coefficient (labeled as β in the tables below) is then used to help describe the magnitude of relations seen in the plots.

Results

Descriptive data. Examination of GAM trends in trial-by-trial average number of optimal presses indicates that learning occurs in each sample, and condition (Figure 8A). There is also an indication that learning occurs more quickly and optimal presses are generally more frequent in the two social-motive conditions relative to the minimally social condition (Figure 8B).

Model estimated parameters.

Model comparisons. Six models were fit to the data and compared using an estimate of leave-one-participant-out cross-validation prediction accuracy. In the first set of three models, parameters were estimated for individuals, nested within sample (resulting in parameter estimates for the population, each sample, and each individual). The second set of three models dropped the sample grouping. The maximal model for each group included parameters (for each condition) for the learning rate (ϵ), reward modifier (also referred to as inverse temperature; ρ), irreducible noise (ξ), and right-arrow bias (b). A second model in each group dropped the parameter, b, while the third model dropped both b and ρ. While the maximal model provides the highest expected predictive accuracy, there is little difference between models that account for grouping of participants within sample versus those that don't – simply allowing each

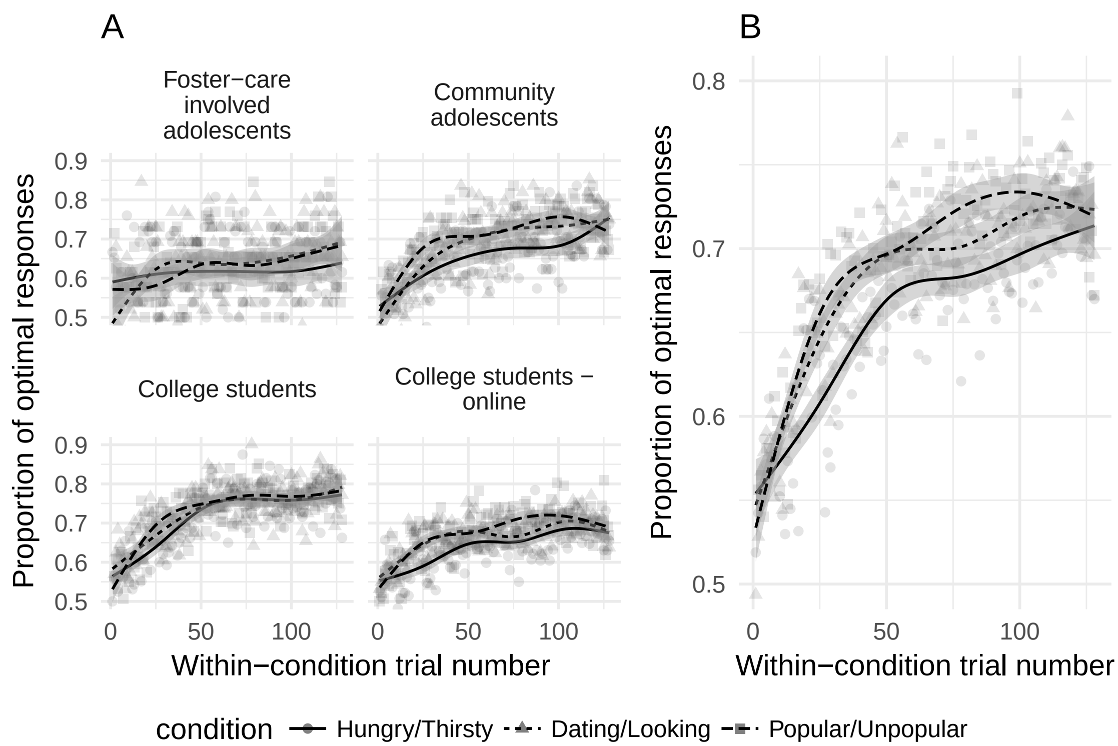

Figure 8. Average optimal behavior over trials. Each point is the average across participants for that trial for the indicated condition. Best fit lines and confidence bands are estimated using generalized additive models.

Table 6. Model comparison

Model	Δ ELPD	LOOIC
3-level, maximal	0.00	134588.73
2-level, maximal	-1.03	134590.78
3-level, without parameter b	-646.38	135881.48
2-level, without parameter b	-651.55	135891.82
3-level, without parameters b, ρ	-1150.56	136889.85
2-level, without parameters b, ρ	-1162.89	136914.50

Δ ELPD is the change in expected log predictive density from the best fitting model. LOOIC is the leave-one-out information criterion – like other information criteria, smaller values are better.

individual's parameters to differ is sufficient for capturing any overall difference between samples. We can also verify that parameter estimates change very little between the two- and three-level specifications, showing near perfect correlation (Figure 9). Correlations of individual-level estimates of the same parameters between all two-level models were also very high ($r_{\varepsilon 1,2,3} > .89, r_{\xi 1,2,3} > .87, r_{\rho 1,2} = .79$). In the maximal, two-level model, correlations between parameters were generally quite small, although there is a somewhat more systematic relation when looking at the noise parameter, ξ (Figure 10).

Validation of parameters with respect to behavior. In order to build further confidence in the plausibility of the estimated individual-level parameters, it may be helpful to examine how they relate to the proportion of optimal choices during the last half of the task run, when learning appears to be reaching its completion, and performance, on average, is plateauing. First, we should see that performance should be worse for people with high estimated ξ (noise) parameters relative to participants with similar values of other parameters. Second, we should also see that, all else equal, both higher estimates for ρ (reward modifier), and ε (learning rate), should be associated with better performance up to an optimal point, from which performance decreases. This curvilinear relation is due to the fact that because the task is probabilistic, the participant sometimes receives feedback in favor of the non-optimal choice. When ρ, ε, or both are very high, the learner may behave as if early but misleading feedback takes precedence over later feedback, or they may behave as if each new piece of feedback is the most important information for guiding their next decision. This is precisely what we see in these data. For each panel in Figure 11, ρ increases along with the overall proportion of optimal responses; as ξ increases (mapped from red to blue, and also tracked by a split indicated by solid and dotted lines), optimal responding is diminished no matter what the estimate of the other parameters; and though we do not see the sweet spot for ρ, we do see that

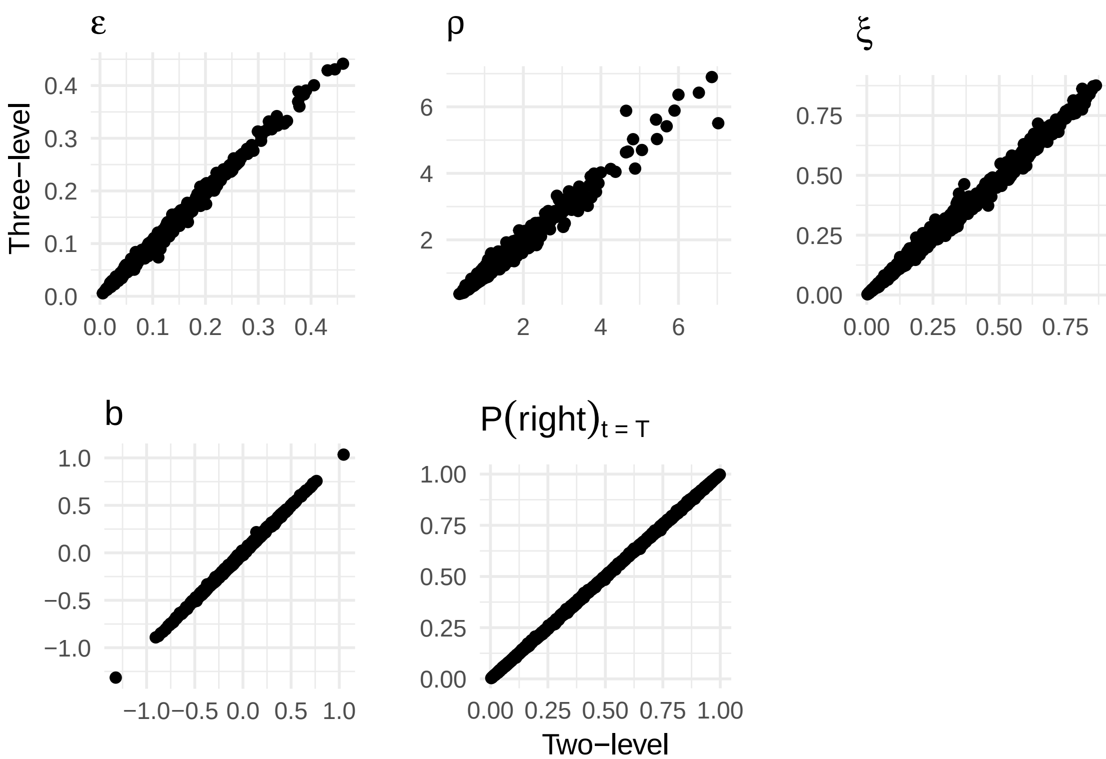

Figure 9. Correlation between three- and two-level model parameters. Each individual-level parameter for the three-level maximal model is plotted against that for the two-level maximal model (for all conditions). The values of $P(\text{right})_{t=\text{T}}$ are the probabilities of pressing the right arrow for the final trial, T, for each participant for each condition.

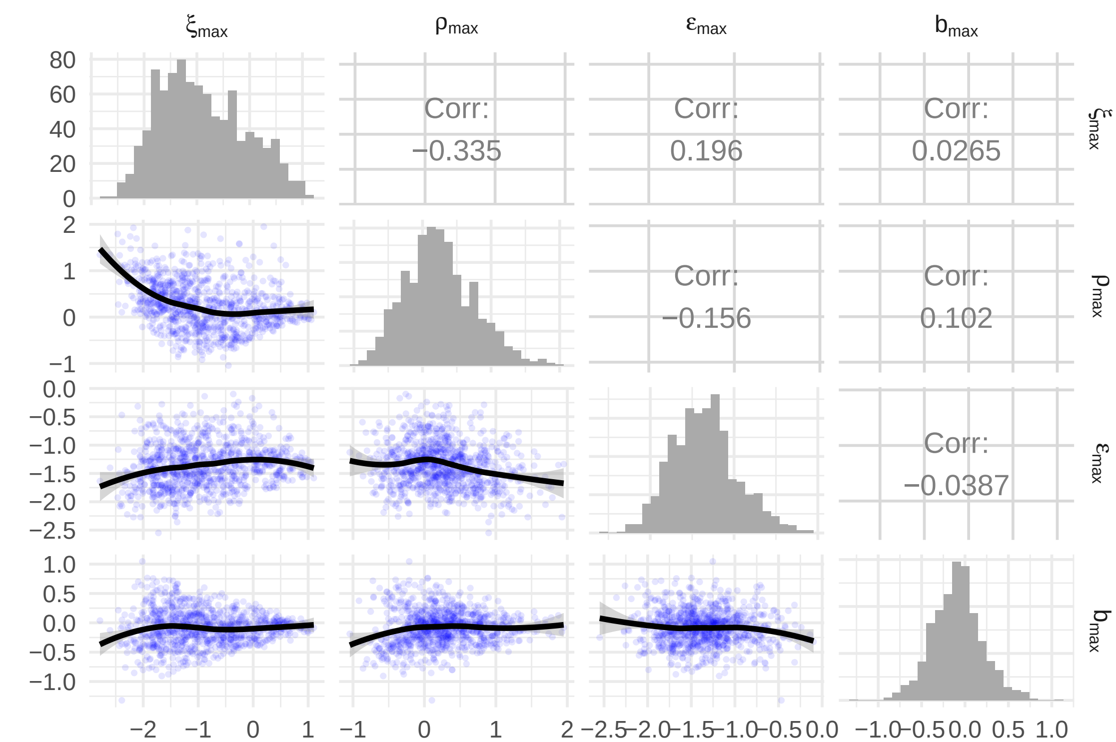

Figure 10. Correlation between parameters in 2-level maximal model. Points are poster means of individual parameter estimates. All parameters are shown on their untransformed (i.e., unconstrained) scales. Black lines are LOESS curves. The diagonal shows the count, while the upper triangle shows Spearman correlation coefficients. ξ: noise; ρ: reward modifier; ϵ: learning rate; b: right-arrow bias.

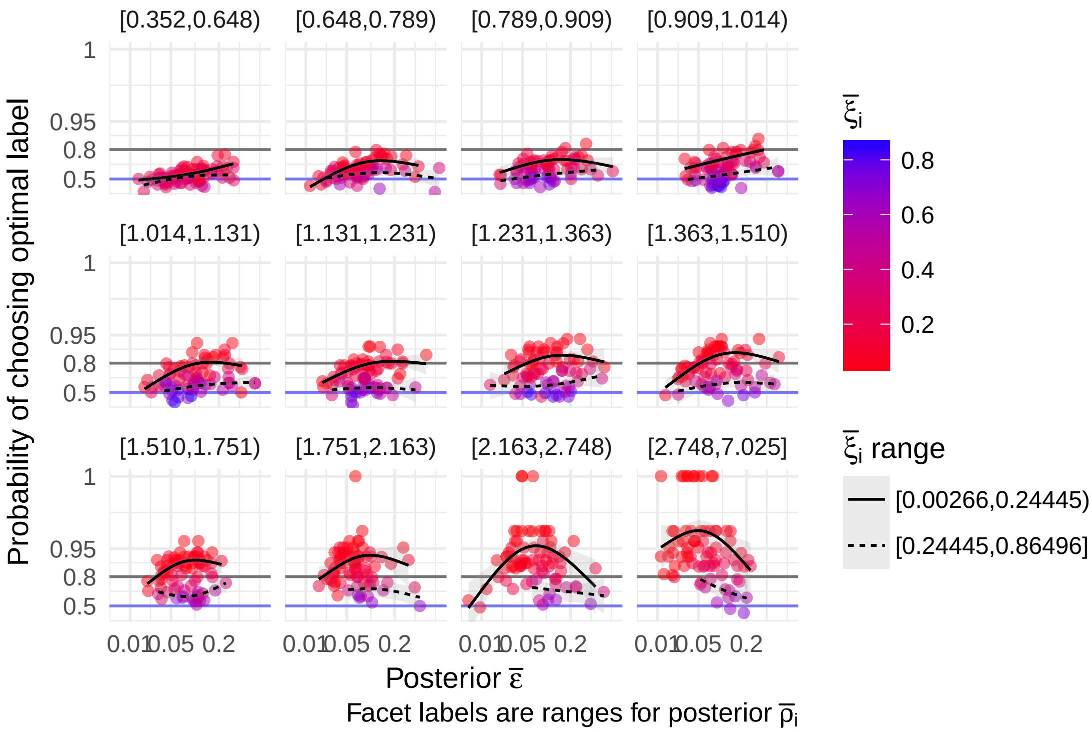

Figure 11. Run-end optimal press behavior and RW learning parameters. Each point is data from one condition for a single participant. As the learning rate parameter, ε, increases (X-axis), probability of making an optimal choice increases to a maximum and then decreases. Panels bin increasing values of the reward modifier, ρ, which also shows increasing optimal choice with larger values. However, as expected, higher values of the noise parameter, ξ (dotted line), suppress the associations between optimal choice behavior and the other two parameters.

as ρ increases, the sweet spot for ε becomes more pronounced, with optimal responding increasing until, approximately, $\varepsilon > .1$.

Very similar patterns are apparent between the parameters and the confidence participants had in their choices. At the end of each of 8 blocks (48 choices each), participants were asked, generally, how well they thought they knew which word went with which face. Scatter-plots and generalized additive model best-fit lines (Figure 12) show striking relations between global confidence ratings and both higher ρ ($r = .37, 99.5\%$ CI $= [.22, .50]$), and lower ξ ($r = -.34, 99.5\%$ CI $= [-.48, -.19]$) that mirror the relations to optimal choices seen above (ε also shows a smaller negative relation to confidence; $r = -.12, 99.5\%$ CI $=$ $[-.27, .05]$). This implies, at least, that the participants are aware of the general reward-optimality of their behavior, so that the model parameters, which show systematic relations to that behavior, also correlate with confidence ratings.

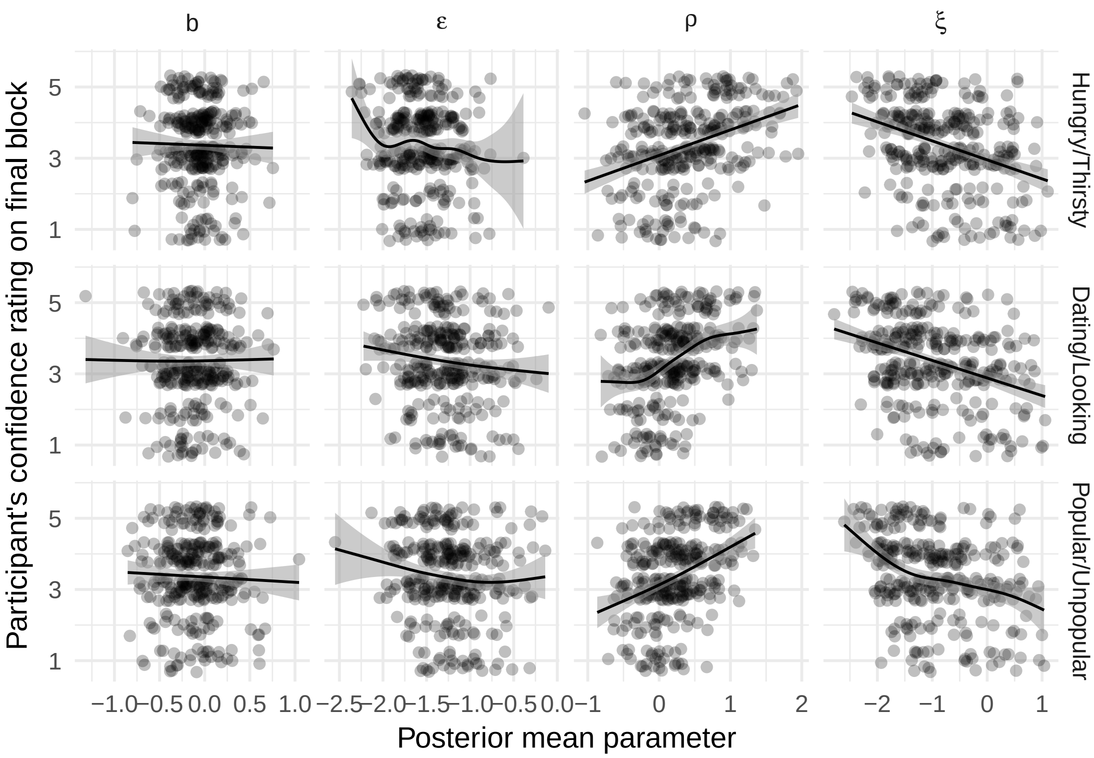

Figure 12. Confidence in learning and RW learning parameters. Each point is the confidence rating for a single participant at the end of all trials (and because each participant has only one rating, a participant's value on the Y-axis is the same across all panels). A clear association between the reward modifier, ρ, and confidence, and the noise parameter, ξ, and confidence can be seen. The right-arrow bias, b, and learning rate, ε, do not show clear associations with confidence.

Population-level parameters. Overall, learning was potentiated in the Dating/Looking, and Popular/Unpopular conditions. Comparing the proportion of optimal choices in the last half of the run, paired sample Yuen *t*-tests (computed using the PairedData package; Champely, 2017), which are robust to non-normality (Yuen, 1974), reject the null of no difference between the two conditions of interest and the control condition (Dating/Looking: $t(185) = 2.69$, $p = 0.008$, $\bar{D} = 0.029$, $\mathrm{CI}_{.95} = [0.008, 0.051]$; Popular/Unpopular: $t(185) = 4.54$, $p =$ 1e-05, $\bar{D} = 0.047$, $\mathrm{CI}_{.95} = [0.027, 0.068]$). The differences in observed behavior are also reflected in some parameter estimates. The 99.5% credible intervals for the difference in learning rate, ε, between the two social conditions and the minimally social condition are both well above 0, while the difference between the two social conditions has much of its mass close to 0 (Figure 13). This indicates that, on average, individuals' learning rates in the social conditions are larger relative to the minimally social condition, but are not very different between them (although there is some not insubstantial probability that the Popular/Unpopular learning rate is of the greatest magnitude). Differences in the population mean of the reward modifier, ρ, have large portions of the posterior probability mass both above and below 0 for all contrasts (Figure 14). Although there is a bit more variability in the contrasts between the social conditions and the reference condition, the social manipulation does not have a very strong or systematic effect on this model parameter. The estimated difference in the irreducible noise parameters, ξ, is also not convincingly positive or negative for any of the contrasts, although the majority of the probability mass for the social versus reference contrasts does lie below 0 (Figure 15). This suggests that there may be less random responding in the social conditions relative to the reference condition. Finally, differences in right-bias, b, are also not convincingly either positive or negative, although there is a suggestion of greater tendency toward choosing the left social option (Dating, median $b = -0.11, \mathrm{CI}_{.995} = [-0.20, -0.02]$; Popular, median $b = -0.10, \mathrm{CI}_{.995} = [-0.19, -0.01]$) than the left reference option (Thirsty, median $b = -0.05, \mathrm{CI}_{.995} = [-0.12, 0.034]$; Figure 16). The descriptors, "Hungry", "Dating", and "Popular" always appeared on the left, so the credibly negative medians for each of the social conditions, as well as the suggestion of difference between the social and reference conditions, may indicate that participants favored choosing the socially desirable descriptors (as opposed to "Looking" and "Unpopular") for some reason.

Given the dynamic nature of the task and non-intuitive contributions of model parameters to probabilities of choosing one option over another, the consequences of these differences between conditions may captured more clearly by an example of the kinds of behavior these parameters could give rise to.

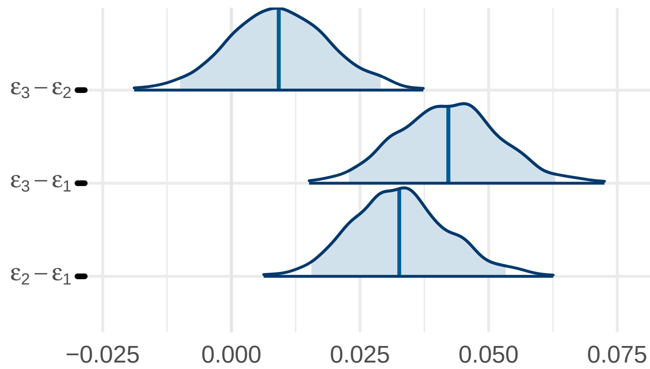

Figure 13. Condition contrasts for learning rate, ε. Subscript number denotes condition. Condition 1 is Hungry/Thirsty, Condition 2 is Dating/Looking, Condition 3 is Popular/Unpopular. Shaded region is 95% credible region, and tails extend to 99.5% credible region.

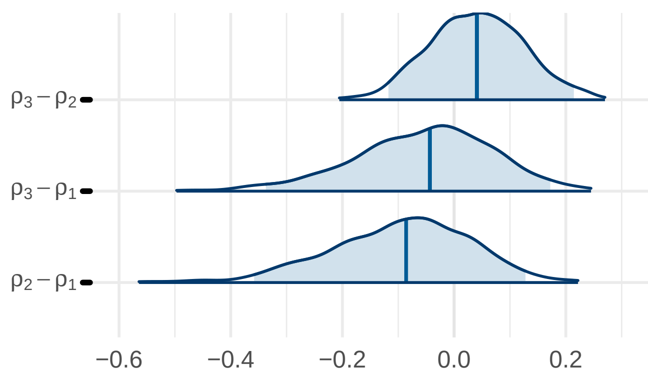

Figure 14. Condition contrasts for reward modifier, ρ. Subscript number denotes condition. Condition 1 is Hungry/Thirsty, Condition 2 is Dating/Looking, Condition 3 is Popular/Unpopular. Shaded region is 95% credible region, and tails extend to 99.5% credible region.

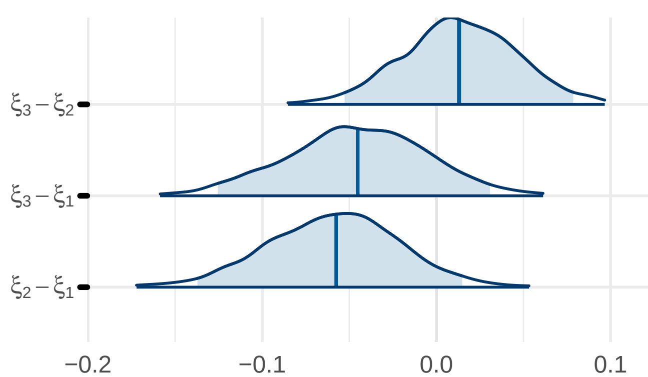

Figure 15. Condition contrasts for noise, ξ. Subscript number denotes condition. Condition 1 is Hungry/Thirsty, Condition 2 is Dating/Looking, Condition 3 is Popular/Unpopular. Shaded region is 95% credible region, and tails extend to 99.5% credible region.

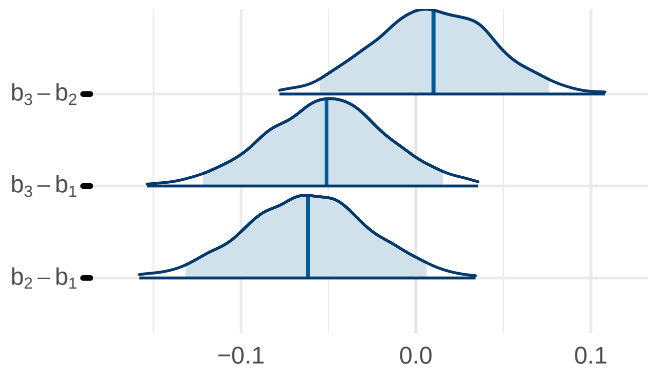

Figure 16. Condition contrasts for right-arrow bias, *b*. Subscript number denotes condition. Condition 1 is Hungry/Thirsty, Condition 2 is Dating/Looking, Condition 3 is Popular/Unpopular. Shaded region is 95% credible region, and tails extend to 99.5% credible region. Positive values indicate bias toward right-button responding (i.e., ”Thirsty”, ”Looking”, ”Unpopular”), negative values toward left-button responding (i.e., ”Hungry”, ”Dating”, ”Popular”).

A helpful approach might be to generate behavior over many identical runs using a hypothetical actor that takes on these population-level mean parameters (specifically, their posterior means). Doing so, we can look at the many possible paths through the paradigm that might be traversed by this actor who represents an “average” participant (though there may be no such real participant in existence, so this is merely illustrative). Looking at the plot of 313 simulated runs in Figure 17 reveals a very slight effect of the bias toward choosing the left option (i.e., “Hungry”, “Dating”, “Popular”; visible most easily at trial 0), as well as a more profound effect of parameter differences on both the magnitude of the actor’s probability of choosing optimally, as well as the variance in that probability. The simulated behavior suggests that in the social condition this actor may progress much more rapidly toward certainty, with much stronger certainty, eventually, but also that there is the potential to be wildly wrong about the long-run optimal choice during some of the run.

Age, puberty, and learning. Relations between measurements of development and both observed behavior as well as model parameters are explored below. The presentation of results focuses first on scatter-plots of the data and GAM trend lines, which are then interpreted with the aid of standardized coefficients form LME models. Note that age for the adolescent samples is calculated using the session date and the participant’s birthday, while the young adult sample only provided their age, in years, as an integer.

The reliability of the proportion of optimal decisions and the individual-level model-estimated parameters is important to consider when interpreting the correlations with other individual differences

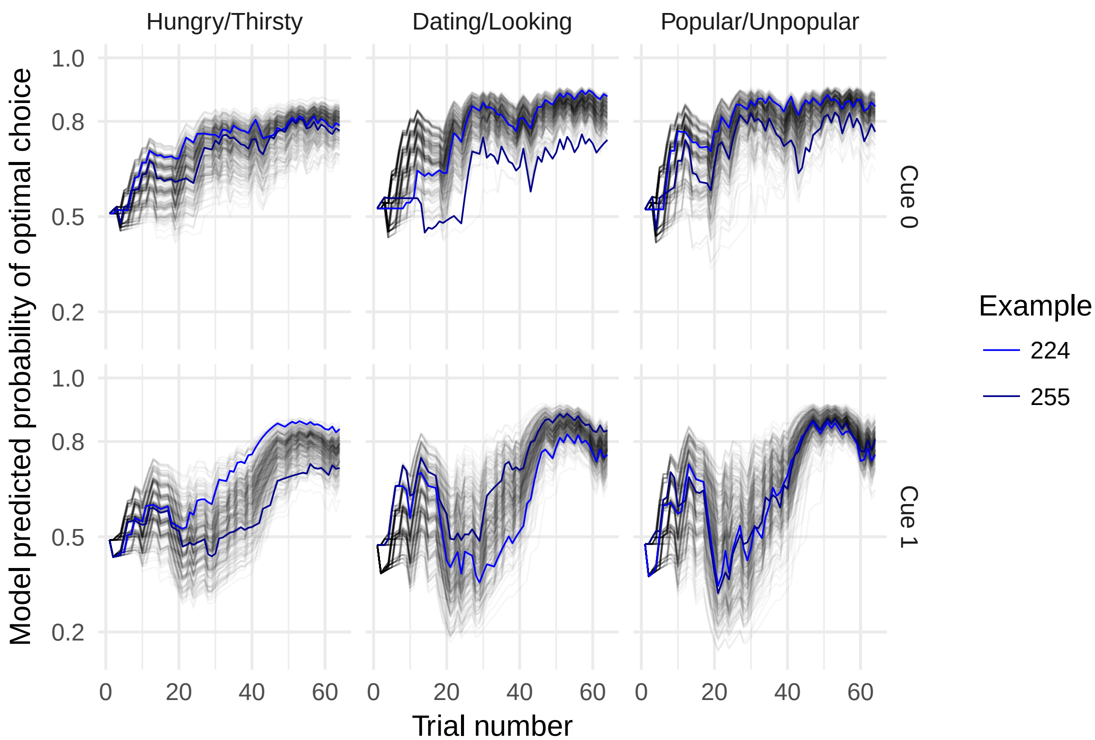

Figure 17. Model predicted behavior. Each gray line follows a simulated agent's probability of choosing the optimal descriptor over one of 313 identical stimulus presentations. The stimulus presentation is randomly generated for each of the two cues and repeated for each condition, but the agent's learning parameters, ε, ρ, ξ, and b are set at the estimated population means for each condition (left-arrow press is optimal for cue 0, right-arrow press is optimal for cue 1). Two simulated runs were chosen randomly to highlight unique trajectories through the task.

measures. Proportions of variance accounted for by participant-level random effects was calculated from a generalized linear mixed effects model implemented in the lme4 package. The model included population and participant-level effect terms for each condition (with Hungry/Thirsty as the reference condition), the trial number, and an interaction between trail number and condition to account, crudely, for learning. The proportion of variance denominator was calculated by summing the variance of all participant-level effects with the residual variance for binomial logistic regression equal to $(15/16)^2\pi^2/3$. The proportion of total variance accounted for by individual differences in optimal presses (averaged across the run) in the Hungry/Thirsty condition was .25. Individual differences in the difference in optimal responding to the two social conditions accounted for and additional .13 (Dating/Looking) and .14 (Popular/Unpopular). The proportion of variance accounted for by the linear effect of trial number in all three conditions was very small (.02 for each of the three conditions).

Reliability of individual differences in the model for reinforcement learning was calculated for each parameter by dividing the variance in the means of the participant-level posterior distributions by the total variance across all the samples in the posterior distributions. In other words, we take the variance in the means of the individual parameters and divide it by the total uncertainty around those means (which also includes the variance of the means themselves). The contrasted parameters of primary interest. Individual-level mean posterior differences between the learning rate $\epsilon_{\text{Dating/Looking}}$ and $\epsilon_{\text{Hungry/Thirsty}}$ account for .31 of the total variance in the posterior distributions for this contrast. For $\epsilon_{\text{Popular/Unpopular}} - \epsilon_{\text{Hungry/Thirsty}}$, the proportion is .33. For the reward modifier, ρ, the proportions for the contrasts are Dating/Looking (DL) = .23 and Popular/Unpopular (PU) = .25; for the noise parameter, ξ, DL = .22, PU = .26; and for right-arrow bias, b, DL = .43, PU = .42.

Mean optimal presses. First, we examine the relation between the proportion of optimal choices during the last half of the run for each condition relative to age and mean scores on the pubertal development scale (PDS). Older participants appear to make more optimal choices up until about age 18 years, at which point this positive relation seems to flatten out (Figure 18). Scores on the PDS show a very similar pattern (Figure 19). To quantify the magnitude of these trends, we calculated standardized regression coefficients from the fixed effects of a LME model separately for the adolescent and college samples. With respect to age, sample is almost totally confounded, and so the magnitude of relation in the adolescent sample stands in for the relation during early to late adolescence, while the magnitude of relation

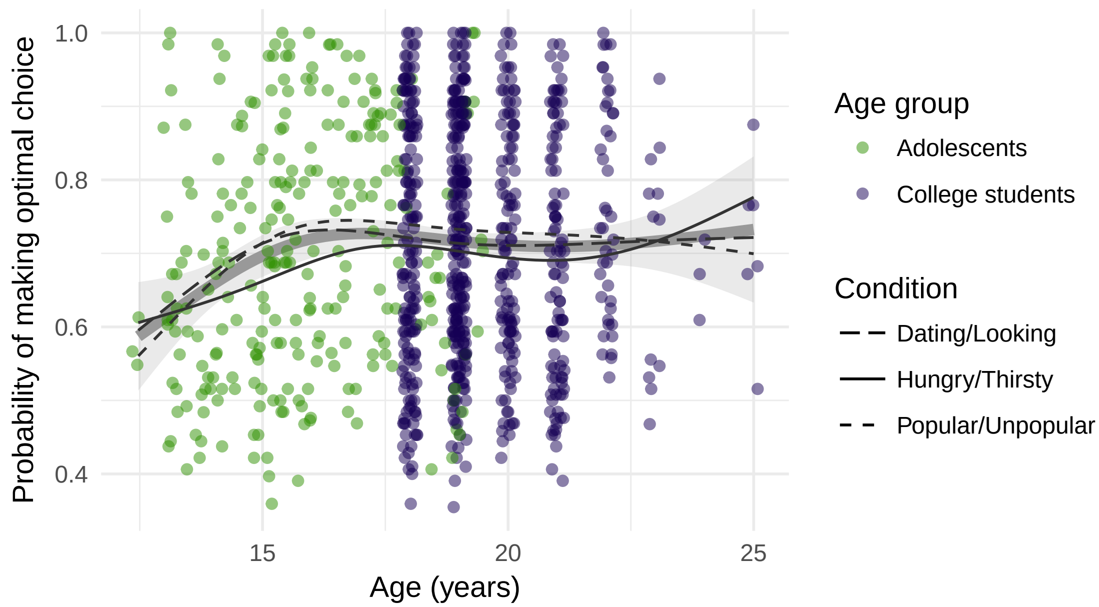

Figure 18. Probability of optimal response versus age. Optimal choice probability is calculated as the proportion of optimal choices in the last half of the run. All best-fit lines are from generalized additive models. Grey line is fit to all data.

in the college sample stands in for the relation during late adolescence and early adulthood. With respect to the PDS, although there is considerably less variability in the college sample as compared to the adolescent sample, there is actually *some* variance, and so standardized coefficients are presented for the PDS for each subset as well. The magnitude of these trends is positive, though fairly small, with both age and PDS showing the largest standardized regression coefficients associations in the adolescent sample (Table 7). In the college sample, the relation between optimal choice proportion and age is estimated to be essentially 0, while PDS shows a positive trend (with a smaller magnitude than in the adolescent sample).

Examining the relation between optimal choice improvements in social conditions and age reveals fairly flat associations (Figure 21). Consistent with the data plot, there is a very small but positive coefficient between age and the optimal choices contrast in the Adolescent group, and an even smaller coefficient in the college group (Table 7; note that the trend in the plot appears to be slightly negative – the trends here are so close to zero that slightly negative is not so different from slightly positive). The plots of optimal choice improvement in social conditions versus the PDS also show relatively flat trends, with much more uncertainty with regard to the relation at low ends of the PDS, where the data are extremely sparse (Figure 21). This is echoed in the magnitude of the standardized coefficients which are both very small (Table 7).

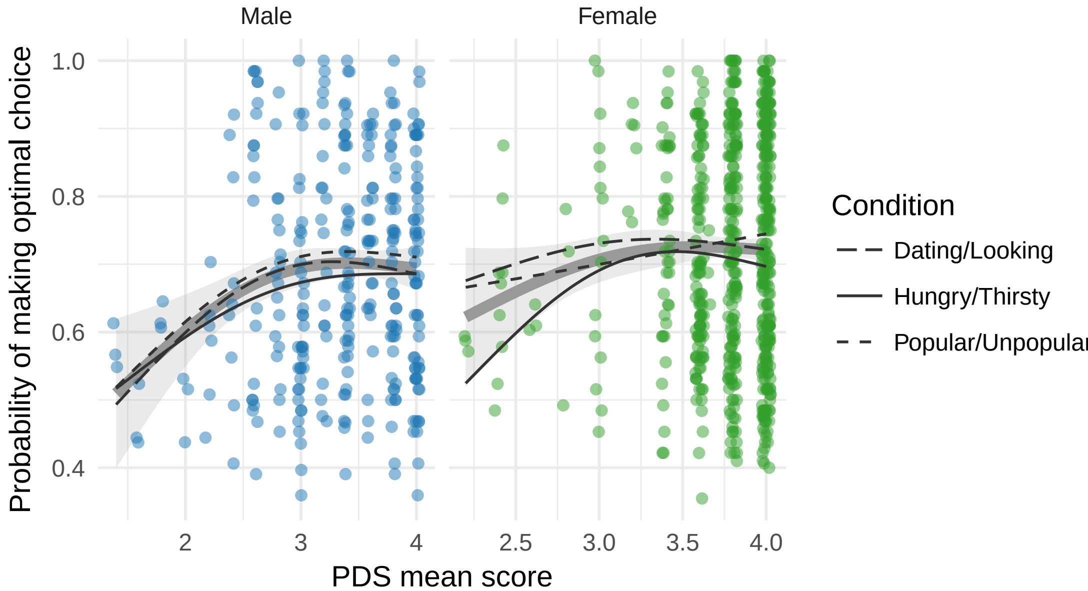

Figure 19. Probability of optimal response versus puberty. Optimal choice probability is calculated as the proportion of optimal choices in the last half of the run. All best-fit lines are from generalized additive models. Grey line is fit to all data. PDS = Pubertal Development Scale.

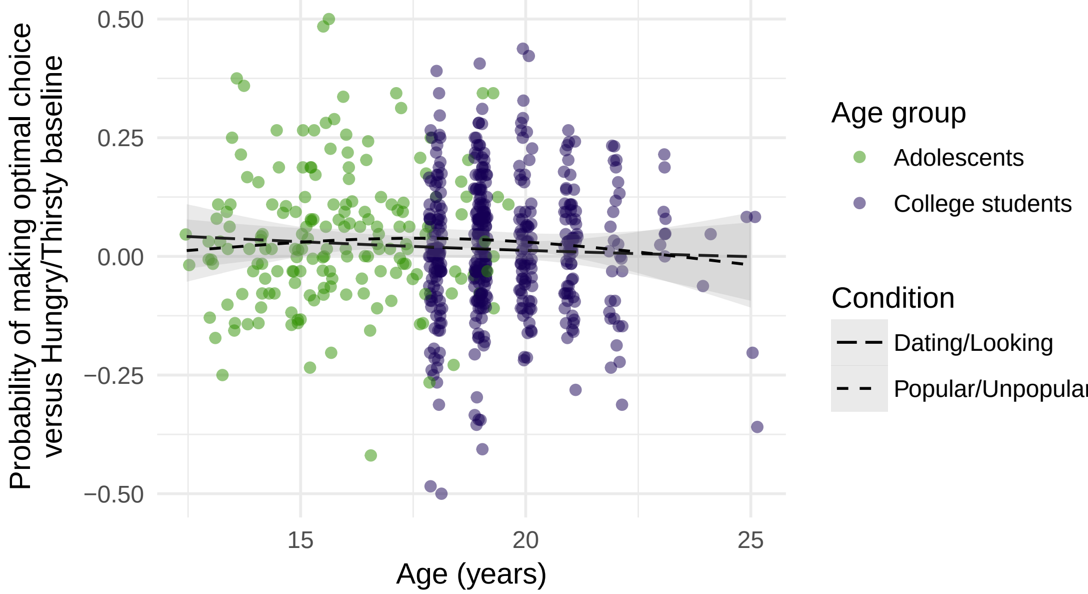

Figure 20. Condition differences in probability of optimal response versus age. Optimal choice probability is calculated as the proportion of optimal choices in the last half of the run. Difference scores are calculated versus the baseline, Hungry/Thirsty, condition. All best-fit lines are from generalized additive models. Grey line is fit to all data.

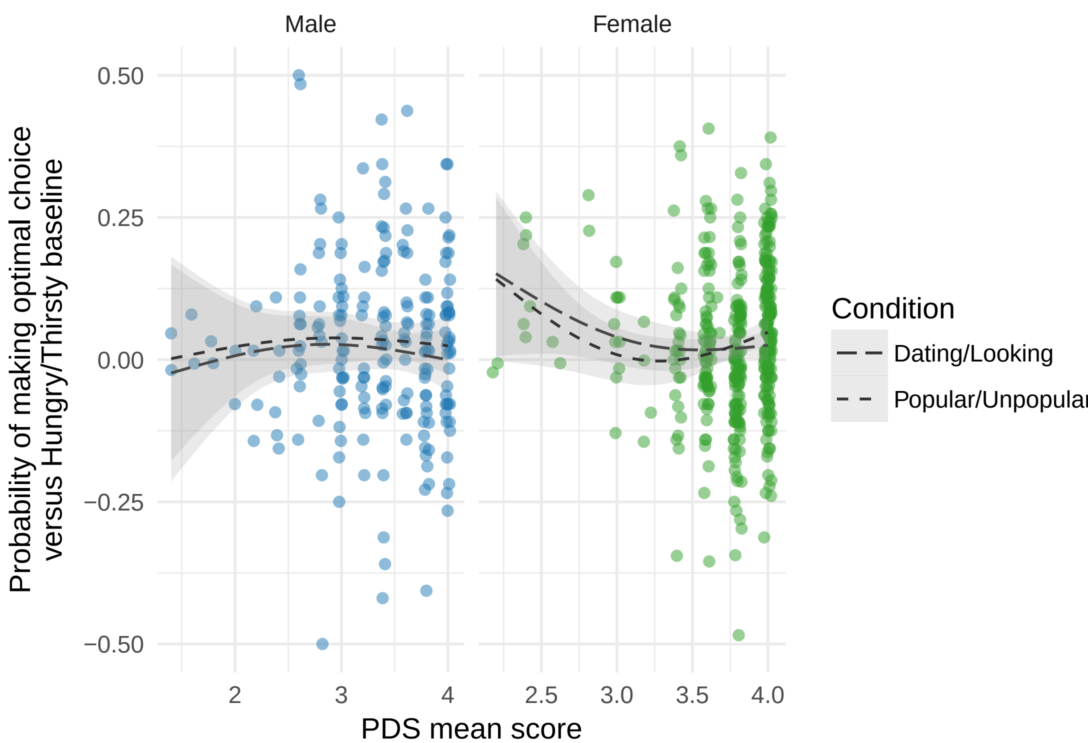

Figure 21. Condition differences in probability of optimal response versus puberty. Optimal choice probability is calculated as the proportion of optimal choices in the last half of the run. Difference scores are calculated versus the baseline, Hungry/Thirsty, condition. All best-fit lines are from generalized additive models. Grey line is fit to all data. PDS = Pubertal Development Scale.

Table 7. Associations of optimal performance with developmental differences

Variables	Age-group	β	SE_{β}	t
Optimal choices ∼ Age	Adolescent	0.17	0.09	1.87
Optimal choices ∼ Age	College	0.00	0.06	0.02
Optimal choices ∼ Age	All	0.07	0.05	1.48
Optimal choices ∼ Age	Adolescent	0.11	0.08	1.37
Optimal choices ∼ Age	College	0.03	0.05	0.66
Optimal choices ∼ Age	All	0.09	0.04	2.11
Optimal choices ∼ PDS	Adolescent	0.17	0.09	1.85
Optimal choices ∼ PDS	College	0.09	0.06	1.45
Optimal choices ∼ PDS	All	0.12	0.05	2.49
Optimal choices ∼ PDS	Adolescent	0.16	0.08	1.95
Optimal choices ∼ PDS	College	0.09	0.05	1.81
Optimal choices ∼ PDS	All	0.13	0.04	2.99
Optimal choices contrasts ∼ Age	Adolescent	0.08	0.10	0.82
Optimal choices contrasts ∼ Age	College	0.01	0.06	0.15
Optimal choices contrasts ∼ Age	All	-0.03	0.05	-0.67
Optimal choices contrasts ∼ Age	Adolescent	-0.01	0.09	-0.07
Optimal choices contrasts ∼ Age	College	-0.03	0.06	-0.54
Optimal choices contrasts ∼ Age	All	0.00	0.05	0.07
Optimal choices contrasts ∼ PDS	Adolescent	0.03	0.10	0.29
Optimal choices contrasts ∼ PDS	College	0.02	0.06	0.27
Optimal choices contrasts ∼ PDS	All	-0.01	0.05	-0.22
Optimal choices contrasts ∼ PDS	Adolescent	0.03	0.09	0.31
Optimal choices contrasts ∼ PDS	College	0.00	0.06	-0.06
Optimal choices contrasts ∼ PDS	All	0.02	0.05	0.38

Optimal choice outcome was calculated for each condition. The optimal choice contrasts outcome includes just the contrasts between the proportions for the social conditions versus the proportions during the Hungry/Thirsty condition.

Age and parameters. While the relation of developmental differences to overall performance can provide a broad view of possible differences that may be related to maturation, it is possible that the latent causes (as operationalized by model parameters in these analyses) of behavior change in different ways. Using a similar approach as above, each parameter is plotted as a function of age or PDS, overlayed with GAM best fit lines for the average of all conditions, and separately for each condition. Inferences from these descriptive plots are then aided by the use of standardized regression coefficients of the average linear relation between the developmental variable and parameter posterior mean (or mean contrast) from a LME model that accounts for the non-independence of observations. Again, this coefficient is estimated separately for each sample.

Both the reward magnitude, ρ, and irreducible noise, ξ, parameters show age relations that are consistent with what we see above regarding the relations of both parameters and age to performance. Specifically, ρ is bigger and ξ is smaller for older participants, at least up until about age 18 years (Figure 22). This corresponds to small standardized coefficients (positive and negative for ρ and ξ, respectively) in the adolescent age group, and much smaller coefficients in the college age group (Table 8). The relation between age and both b and ε is much more flat, although the learning rate ε seems to show slightly more variability over the age span (Figure 22). The almost negligible standardized coefficients for ε indicate a very small negative relation in the adolescent group, and a very small positive relation in the college group, while b shows a very small positive relation with age in the adolescent group (Table 8).

The difference between the parameters in the social conditions versus the minimally social reference condition across this developmental span is one of the primary observations that bears on the question about whether motive effects on learning reflect what would be expected based on the theories discussed in the introduction. Observations consistent with those theories would be age-increasing differences in the contrast between parameters for learning during the social versus reference conditions. These slopes of the contrasts with respect to age are very flat for all parameters, with a very slight increase for ε, and decrease for ρ (Figure 23). The standardized coefficients, nearly all very small, mirror the best fit lines, although the estimate for ε in the adolescent sample is small and negative (Table 9).

Puberty and parameters. The results for the relation between the PDS and parameter estimates follow roughly the same pattern as above (which is to be expected given the high correlation between pubertal development and age). The parameter ξ shows a negative association with PDS score, while ρ is

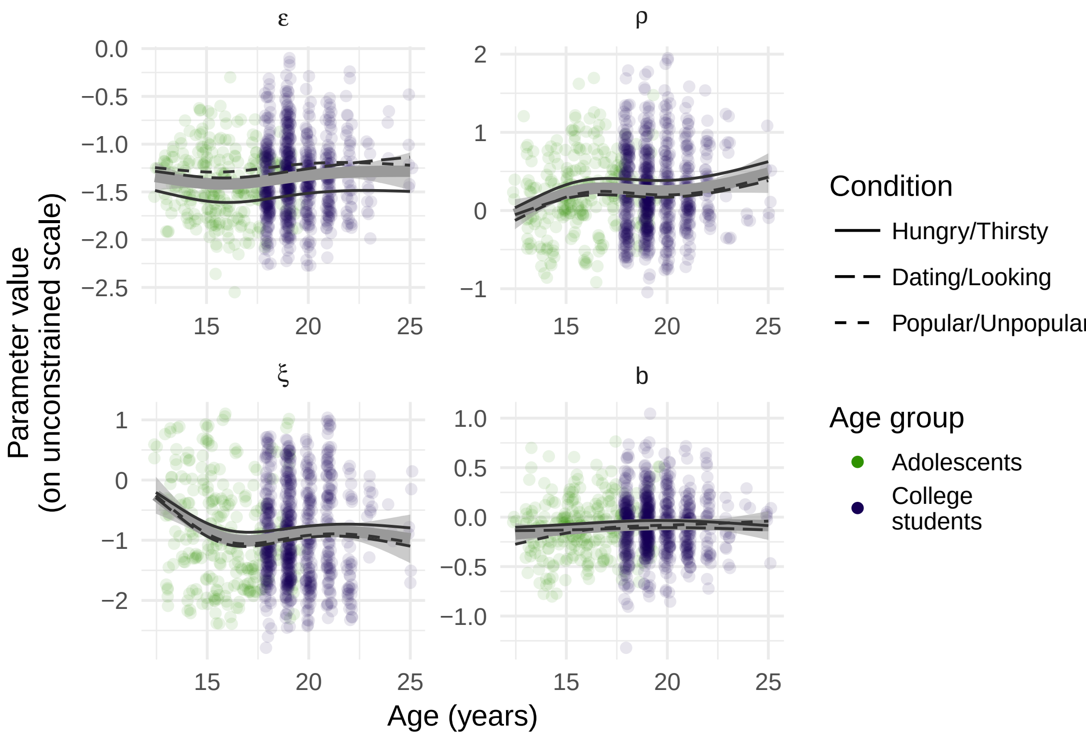

Figure 22. Reinforcement learning model parameters versus age. All best-fit lines are from generalized additive models. Grey line is fit to all data. Parameters are plotted on the unconstrained scales with possible values in $(-\infty, \infty)$. ε: learning rate; ρ: reward modifier; ξ: noise; b: right-arrow bias.

Table 8. Age and parameter estimates

Parameter	Age-group	β	SE_β	t
ε	Adolescent	-0.08	0.09	-0.90
ε	College	0.05	0.06	0.89
ε	All	0.08	0.05	1.69
ρ	Adolescent	0.14	0.10	1.35
ρ	College	0.04	0.06	0.60
ρ	All	0.08	0.05	1.44
ξ	Adolescent	-0.19	0.10	-1.83
ξ	College	0.02	0.06	0.35
ξ	All	-0.05	0.05	-0.92
b	Adolescent	0.10	0.09	1.21
b	College	0.01	0.06	0.26
b	All	0.06	0.05	1.35

Models were estimated using posterior mean parameters on unconstrained scales. ε: learning rate; ρ: reward modifier; ξ: noise; b: right-arrow bias.

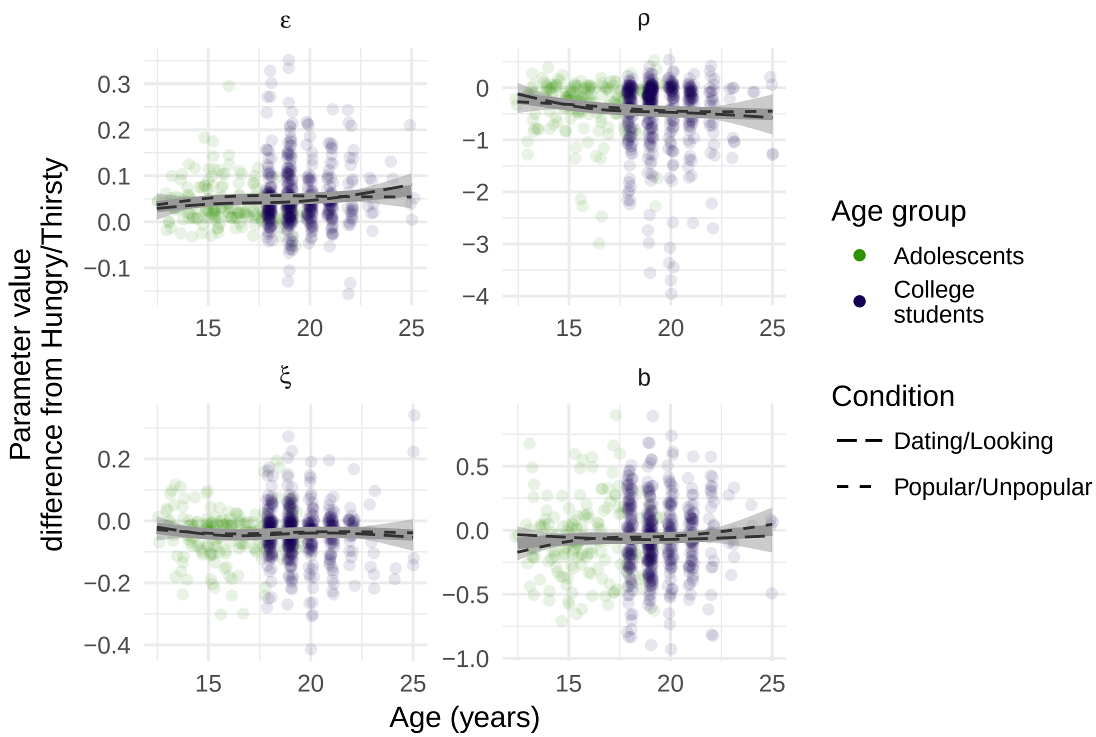

Figure 23. Reinforcement learning model parameter condition contrasts versus age. All best-fit lines are from generalized additive models. Grey line is fit to all data. Differences are calculated using parameters transformed to be on the model scale. ε: learning rate; ρ: reward modifier; ξ: noise; b: right-arrow bias.

Table 9. Age and parameter estimate contrasts

Parameter	Age-group	β	SE_β	t
ε	Adolescent	-0.13	0.08	-1.59
ε	College	0.01	0.06	0.15
ε	All	0.06	0.05	1.26
ρ	Adolescent	-0.06	0.10	-0.59
ρ	College	-0.02	0.07	-0.27
ρ	All	-0.09	0.05	-1.60
ξ	Adolescent	-0.01	0.10	-0.05
ξ	College	-0.03	0.06	-0.45
ξ	All	0.00	0.05	0.05
b	Adolescent	0.03	0.09	0.33
b	College	0.03	0.06	0.50
b	All	0.04	0.05	0.71

Parameter contrasts were computed by subtracting the mean posterior parameter (on the constrained scales) in the Hungry/Thirsty condition from parameter estimates in the two social conditions. ε: learning rate; ρ: reward modifier; ξ: noise; b: right-arrow bias.

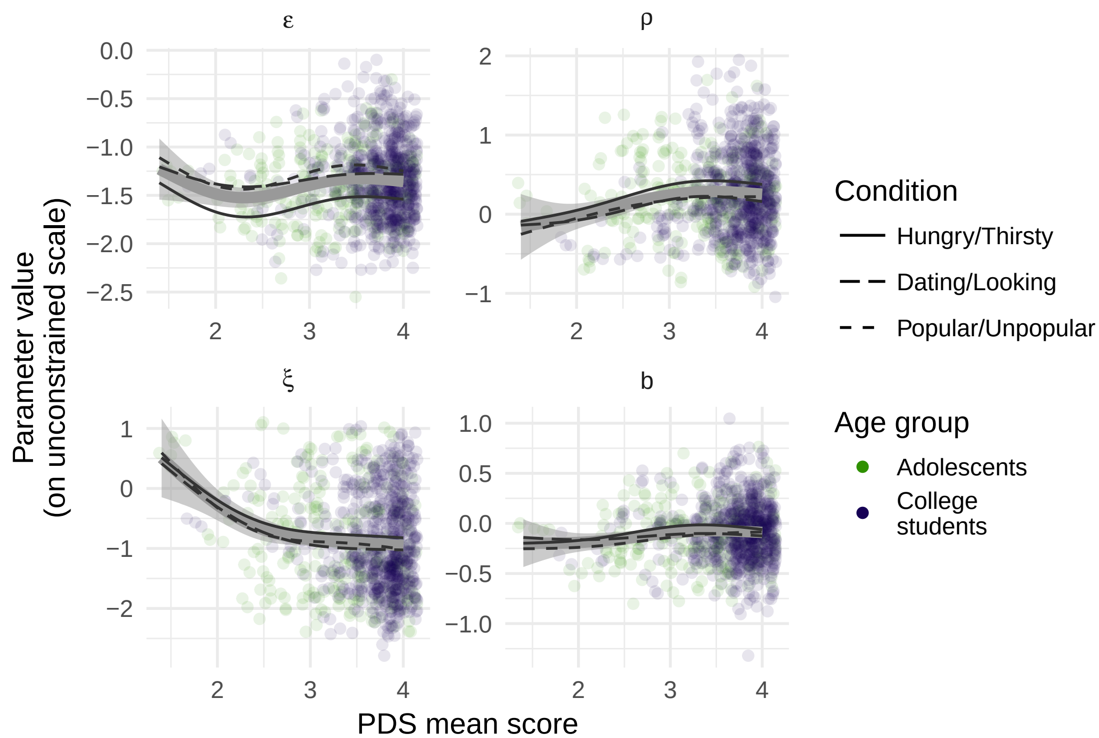

Figure 24. Reinforcement learning model parameters versus puberty. All best-fit lines are from generalized additive models. Grey line is fit to all data. Parameters are plotted on the unconstrained scales with possible values in $(-\infty, \infty)$. PDS = Pubertal Development Scale. ε: learning rate; ρ: reward modifier; ξ: noise; b: right-arrow bias.

fairly flat or positive, ε shows a somewhat variable relation, and b shows a fairly flat relation (Figure 24). The standardized coefficients for the relation between PDS and parameter posterior means are negative for ξ, and otherwise consistent with the plots (Table 10).

Contrasts between learning parameters for the social versus reference conditions are all quite flat in their relation to PDS (Figure 25). Standardized coefficients were all negligible or very small (Table 10).

Discussion, Aim 1

Overall, the best-fitting model for learning captures the behavior of participants on the task well. We see that the model simulations generate data that covers the full range of possible behavior, and that simulated data from the fitted parameters reproduce the average learning trajectory in each condition. The relation between the parameters and both performance, and confidence in learning conform to our expectations based on model construction. Specifically, generally higher learning rates, greater inverse temperature, and lower noise all relate positively to performance and confidence. Generally, these results

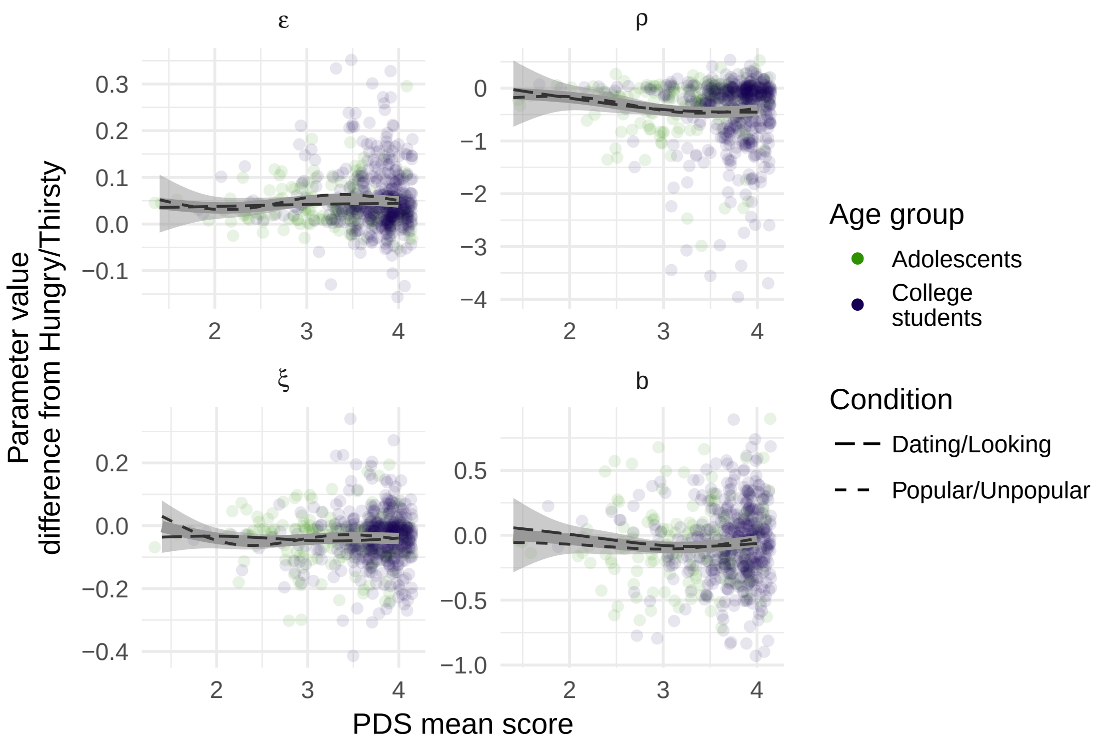

Figure 25. Reinforcement learning model parameter condition contrasts versus puberty. All best-fit lines are from generalized additive models. Grey line is fit to all data. Differences are calculated using parameters transformed to be on the model scale. PDS = Pubertal Development Scale. ε: learning rate; ρ: reward modifier; ξ: noise; b: right-arrow bias.

Table 10. Pubertal development and parameter estimates

Parameter	Age-group	β	SE_β	t
ε	Adolescent	0.09	0.09	0.95
ε	College	0.00	0.06	0.00
ε	All	0.07	0.05	1.46
ρ	Adolescent	0.03	0.10	0.32
ρ	College	0.10	0.06	1.56
ρ	All	0.08	0.05	1.50
ξ	Adolescent	-0.21	0.10	-2.10
ξ	College	-0.10	0.07	-1.50
ξ	All	-0.14	0.06	-2.55
b	Adolescent	0.11	0.09	1.26
b	College	-0.01	0.06	-0.10
b	All	0.05	0.05	1.12

Models were estimated using posterior mean parameters on unconstrained scales. ε: learning rate; ρ: reward modifier; ξ: noise; b: right-arrow bias.

Table 11. Pubertal development and parameter estimate contrasts

Parameter	Age-group	β	SE_{β}	t
ε	Adolescent	0.00	0.09	-0.04
ε	College	-0.03	0.06	-0.57
ε	All	0.02	0.05	0.48
ρ	Adolescent	0.01	0.10	0.09
ρ	College	-0.04	0.07	-0.54
ρ	All	-0.06	0.06	-1.01
ξ	Adolescent	-0.07	0.10	-0.74
ξ	College	0.04	0.06	0.72
ξ	All	0.01	0.05	0.12
b	Adolescent	0.04	0.09	0.39
b	College	0.01	0.06	0.18
b	All	0.03	0.05	0.50

Parameter contrasts were computed by subtracting the mean posterior parameter (on the constrained scales) in the Hungry/Thirsty condition from parameter estimates in the two social conditions. ε: learning rate; ρ: reward modifier; ξ: noise; b: right-arrow bias.

can give us confidence that the model adequately capture a plausible cognitive process that gives rise to the response data.

Performance on the learning task, indexed by proportions of optimal presses, follows expected age and pubertal-development trends, with very early adolescent participants performing less well on the task. The cognitive demands of this task are likely high as is reflected in the protracted period of learning, and so an increase on overall performance with increasing age cohort is not unexpected. Age and development trends in parameter estimates follow this pattern as well.

Differences in performance, as well as model parameters, indicate that the two social-motive conditions enhance learning. This is consistent with the idea that humans are motivated, especially, to learn social information. However, estimates of these differences between conditions were fairly consistent across the entire age range, which is not what would be expected if differences between conditions in learning reflect motivations that are thought to be developing during this age range.

Limitations. Before connecting the results of this analysis to the broader literature on reinforcement learning and social motivation, it is important to set out the limitations that must constrain interpretations of these findings. First, learning may be obviously irrelevant to actual social goals participants have. That is, learning about the popularity of a computer-generated face may not be enhanced by a person's status motives because it is clear that this information is not instrumentally valuable for satisfying that motive. However, we did not ask participants whether they felt like they learned something that could help them in their daily lives (that is, we did not assess whether they thought there was some true, learned association between facial characteristics and either dating or social status). So it is possible that either the irrelevance should not matter, or that many participants were motivated to learn about the faces because they thought it might be relevant information for planning future behavior.

Second, it is well known that binary choice outcomes are less reliable indicators of latent constructs than continuously measured behavior. The proportion of variance in the parameter estimates (for learning parameters), or performance estimates, versus the variance due to uncertainty in those estimates was generally low. This would diminish the sensitivity of this measure to detect differences resulting from motive development at different ages.

Third, a minimal stimulus set only examines a small slice of the population of descriptors that may be relevant for mate-seeking and status motives. Depending on the idiosyncratic characteristics of the chosen stimuli, this could produce both mean differences between conditions, as well as age correlations, or it could mask true differences and correlations. The generalizability of any results is limited by the extent to which there is variation in any relations due to the choice of particular descriptors. Note that randomizing the associated faces eliminates this constrained stimulus set as a possible confound, though the use of the same six faces may limit generalizability to some extent as well.

Fourth, the face-valid relevance of the descriptors to the motivational domain is not the only dimension along which the conditions differed. Unlike the minimally social condition, the social descriptors convey information not just about the target, but about that target's relationship to other individuals. From the perspective of social network analysis, the descriptors "Dating" and "Looking" convey information about at least two individuals, as well as the tie(s) between them (both "Dating" and "Looking" require at least one unnamed conspecific). The descriptors "Popular" and "Unpopular" convey information about a person's position in a broader hierarchy, and thus about the structure of ties among a much larger community. That participants may have been sensitive to variation along this dimension is speculative, but this is at least one influence that is not related to specific social motivations that could result in salience differences between descriptor pairs and so explain the behavior of participants on this task. Another possible dimension along which the stimuli vary may be the extent to which people believe that one can tell from facial characteristics whether someone is likely to match a descriptor. If participants believe that there is a signal in the facial characteristics for the two social conditions, but not the minimally social condition, they may pay more attention during those trials and learn faster. The stimuli also vary along a trait-state dimension, with Hungry/Thirsty descriptors indicating states that fluctuate on a daily basis, and the social-motive descriptors describing individual differences that fluctuate more slowly. If this affected participant expectations, we would expect to see a higher learning rate in the conditions where the descriptor is more state-like, reflecting an expectation that one should update their beliefs more readily in the face of conflicting information. Contrary to this pattern, we estimated higher learning rates (on average) in the social motive conditions.

Fifth, although these analyses compare a suite of models that build on the basic updating rule proposed by Rescorla and Wagner (1972), there are other learning models that may be relevant. For example, although the task was designed with prediction-error updating as the guiding principle, it might be possible to examine these data using instance based learning (Gonzalez, Lerch, & Lebiere, 2003; Lejarraga, Dutt, & Gonzalez, 2012), or a model that unifies the Rescorla-Wagner, temporal difference, and Bayesian Kalman filter approaches (Gershman, 2015). The distinction between model-free and model-based learning strategies is also important, with evidence for developmental trends in the transition between strategies (Decker, Otto, Daw, & Hartley, 2016), but this task is not designed to examine these distinctions.

Finally, although this is a large sample that covered a broad age range, the data are cross-sectional and so it would be inappropriate to extrapolate differences across age to developmental effects within individuals (Fisher, Medaglia, & Jeronimus, 2018). One stark example of this problem in the present research is that to infer a developmental effect in these data, one must assume that participants are exchangeable across age; that is, one must assume that the sample of young participants is more or less the same as the older participants, but just younger. This is clearly not the case in this sample when one considers the fact that all of the older participants are attending a four-year university, whereas it is almost certainly not the case that all of the community and foster-care-involved adolescents will follow this educational path. There is no way to gauge how much this impacts the results of these analyses. However, it is reassuring, for example, that there does not seem to be a discontinuity between the samples in performance or parameter estimates. Cross-sectional age differences may be a confounding source of variance for developmental effects that again may produce apparent associations or obscure true associations.

Differences in reinforcement learning, generally, across adolescence. How do the results from this study inform our understanding of, first, reinforcement learning across adolescence and, second, social-context effects on reinforcement learning? Almost all work on reinforcement learning in adolescence is cross-sectional, with small samples partitioned by age into groups with fewer than about 30 participants. A few investigations exceed these sample sizes, somewhat (Davidow, Foerde, Galván, & Shohamy, 2016, adults = 31, adolescents = 41; Jones et al., 2014, adults = 37, adolescents = 45, children = 38; McCormick & Telzer, 2017, adolescents = 77). In one notable exception, Peters and Crone (2017) report on a longitudinal sample with 299 participants age 8-25 years over 3 waves yielding 736 observations. They report results that are consistent with improvements in performance on reinforcement learning tasks across adolescence.

Performance differences across development. The most consistent finding from this literature, and which is consistent with the present study, is that adults, or young adults, perform better than adolescents and children (van Duijvenvoorde, Zanolie, Rombouts, Raijmakers, & Crone, 2008; Decker, Lourenco, Doll, & Hartley, 2015; van den Bos, Güroğlu, Van Den Bulk, Rombouts, & Crone, 2009; Cohen et al., 2010; Christakou et al., 2013; Palminteri, Kilford, Coricelli, & Blakemore, 2016; Rosenblau, Korn, & Pelphrey, 2017; Peters & Crone, 2017; McCormick & Telzer, 2017). In one study, however, adolescents performed better than adults, which was interpreted as indicating that heightened reward sensitivity may lead to better learning during adolescence (Davidow et al., 2016). The task used was a probabilistic reinforcement learning task, just as in much of the literature indicating more optimal learning in adults compared to adolescents. Notably, adults showed higher learning rate parameters than adolescents, but performed worse as a result. These results are anomalous in the context of the considerable evidence for greater adult performance on standard reinforcement learning paradigms.

This consistent age finding holds across different operationalizations of learning. Both probabilistic reinforcement (van den Bos et al., 2009; Cohen et al., 2010; Palminteri et al., 2016; Decker et al., 2015), and deterministic reinforcement (van Duijvenvoorde et al., 2008; Peters & Crone, 2017) show the increase in performance with age. Also, tasks which are more often described as risk-taking, but which involve a reinforcement learning component like the Iowa gambling task (Christakou et al., 2013), and balloon analogue risk task (McCormick & Telzer, 2017), also show this age trend.

Several studies did not include a measure of performance, but focus only on learning rate. Notably, Learning rates are not linearly associated with performance when the feedback is probabilistic, but show optimal performance when a balance is struck between sensitivity to feedback and the slow accumulation of evidence about optimal behavior in a relatively stable environment. In studies where both learning rate and performance were measured, two studies show an association with age in the same direction (with adults performing better than younger participants; Rosenblau et al., 2017; McCormick & Telzer, 2017). The only study in which this relation is reversed (as mentioned above) was that reported by Davidow et al. (2016), though another study showed differently signed relations between learning rates for positive versus negative feedback (Christakou et al., 2013).

Some authors have suggested that adolescent reward sensitivity makes the developmental period of adolescence advantageous for reward learning (Davidow et al., 2016; McCormick & Telzer, 2017). However, if there is indeed a peak in reward sensitivity during adolescence, it does not result in a particular advantage in most of the standard reward learning tasks reviewed above. The evidence from the present study also supports a gradual increase in performance across the age range (roughly 12-18 years), and which is not consistent with an adolescent specific advantage driven by reward sensitivity. In fact, adolescent reward sensitivity should be apparent in either overall performance, or in the parameter ρ which modulates the magnitude of rewards. Overall, this study represents one of the largest samples to date, second only to the cohort-sequential design of Peters and Crone (2017), and presents results that greatly strengthen the evidence for an age-related increase in performance on reinforcement learning tasks.

Effects of social manipulations. Very few studies examine social effects on reinforcement learning; those that do have relied on two approaches. In the first approach, studies use stimuli that contain social content that is the target of learning (this is the approach taken above in the SPLT; Jones et al., 2014; Rosenblau et al., 2017). The second approach is to situate a standard, abstract stimulus, reinforcement learning task in a social context (Decker et al., 2015; Lockwood, Apps, Valton, Viding, & Roiser, 2016) (and U/P).

The two studies that use social stimuli in the reinforcement learning task do so in a way that makes straightforward interpretation of the learning results difficult. In one, the authors use a Pavlovian conditioning paradigm wherein the unconditioned stimulus is social in nature (Jones et al., 2014). Specifically, on each trial, the participant sees a picture of one of three faces, which winks with the left or right eye. The participant then indicates via button press which eye winked. Following the response, they see a screen that indicates whether she or he receives a virtual note from the virtual peer in the picture, or whether another peer received the note. The three faces vary by how often their wink is followed by the positive social reinforcement of receiving the note. Reaction time for the response to the wink cue was used as the outcome for the reinforcement learning model, and accuracy (e.g., choosing ‘left’ when the left eye winked) was also examined. In a sample of adults (N = 37), adolescents (N = 45), and children (N = 38), they found that wink responses were more accurate for the high-reinforcement faces (but they did not find a significant interaction with age, or report the estimated interaction coefficient), and that learning rates vary by age (with younger participants having higher learning rates). This study indicates that social rewards by themselves may be sufficient to induce accuracy differences for conditioned cues.

However, there are several aspects of the analysis that make the association between age and learning rates difficult to interpret. The primary issue is that the learning model is described only in broad strokes, and so the link between the expected value parameter and reaction time is not stated. Inspecting the reported scatter plot of learning rate versus age (Figure 2, p. 691, Jones et al., 2014) reveals that many of the estimated learning rate parameters were 0, indicating that for a substantial portion of the participants, no conditioning occurred (assuming the model is well specified). Also, with regard to the substantive conclusions assuming the data interpretation is correct, it is not possible, without a comparison with non-social-reinforcement, to interpret age differences in learning rates, or accuracy differences by reinforcement probability, as being specific to *social* reinforcers. Notably, the results from this task, consistent with those from the SPLT, show that the impact of differential social reinforcement is not significantly different across ages. At the very least, this null finding corroborates the findings from the SPLT that the salience, or value, added by social content does not vary systematically with respect to participant age.

Another study incorporates meaningful social information directly into the stimuli used in an instrumental learning paradigm (specifically preference ratings of objects in three broad domains: activities, fashion, and food; Rosenblau et al., 2017). These authors use preference ratings from either three adolescents or three adults as learning targets for the adolescent (N = 24) or adult (N = 21) participants, respectively. On each trial, the participant guessed how much the target likes the activity, fashion, or food related item, and then received feedback about the target's true score. The authors define prediction error as the the difference between the participant's guess and the target's true preference, using this as the primary outcome of performance. They find that across both samples, prediction error decreases during the course of the run, with less prediction error overall for the adult group relative to the adolescent group. They also formally model the learning process and find that the best fitting model incorporates reinforcement learning as well as information about the participant's own preferences. The primary findings relevant to this discussion are that adults perform better than adolescents (that is, they accumulate less prediction error over the course of all trials), and this is mirrored by a higher average learning rate for adults relative to adolescents.

Although the paradigm used by Rosenblau et al. (2017) is set up superficially like more abstract instrumental learning tasks, there are some features that may preclude clear interpretations of the results. Each item is seen only once by the participants, meaning that any learning that occurs requires generalization from one item to other items. In other words, learning on the task requires that there is an inherent structure to the targets' preferences that can be learned. This is, of course, trivially true in that each target has a mean rating that could be learned. Though the authors present several non-significant *p*-values to support the equivalence of the adolescent and adult preference profiles, the sample size is quite under-powered to detect differences, and it is not clear what size differences might matter with respect to learning and prediction error differences across age-groups. It *is* clear that more consistent ratings by the target (whatever the assumed structure of relation among item preferences) would lead to better performance. It is also clear that the similarity of a target's and a participant's rating would influence the prediction error and estimated learning parameters. The effect of similarity is included in the learning model, and does improve the model fit, indicating that participants may have been using their own preferences when guessing target preferences. Conclusions about age-group differences in instrumental learning of social information are obscured by the idiosyncrasies in the task design mentioned above. However, the finding of better adult performance is in line with the wider literature both when the content of learning is social, or non-social.

Two studies take the second approach of examining the effect of social context or cues on learning, and while they are less directly relevant to the present research, they are worth considering because they demonstrate that different social cues can affect learning. Lockwood et al. (2016) use a probabilistic reinforcement learning paradigm, manipulating whether the participants (N = 31 males, age 19-32 years) gain rewards for themselves, a friend, or no-one (that is, the points are displayed but do not accrue to anyone). They find that performance and learning rate is higher in the self condition than in the other two conditions. Decker et al. (2015) use a standard probabilistic reinforcement task, but add a cue set for which participants are given information about what the best choice is. They find that adults (N = 26) are more sensitive to this information than adolescents (N = 31). When the instruction is erroneous, this incurs a penalty for adult performance relative to adolescents.

Across social reinforcement learning studies, the tentative conclusion is that social information or reinforcement results in similar developmental patterns in performance as does non-social reinforcement. It does seem possible, though the evidence is limited, that there could be an adolescent-specific advantage with regard to performance in the face of social-misdirection. However, there seems to be scant evidence for any particular adolescent sensitivity to social versus non-social stimulus content, or for advantages (or disadvantages) driven specifically by developmental sensitivity to rewards.

Conclusions. The results from the present study add weight to the evidence for better overall performance on reinforcement learning tasks for older participants in the age range from early to late adolescence and young adulthood. We also find, consistent with prior literature, a positive correlation between performance and learning rate, though there is some indication in this sample that there is a non-linearity such that at some point, higher learning rates impede performance. This is the expected behavior of the learning model with respect to probabilistic feedback, and future work should ensure that this possibility is accounted for, or else risk misinterpreting linear correlations between learning rates and other measures. We also see a positive correlation between the inverse temperature parameter, ρ, that governs the magnitude of rewards (versus no-reward, in this case) and performance. In the mate-seeking and status conditions where performance is best, the learning rate, ε, is indeed higher than in the minimally social condition. Unsurprisingly, the noise parameter, ξ, which estimates the degree of random responding across the entire run, is also monotonically related to performance. The increased performance in the two social conditions is also reflected (though much less robustly) in this parameter being somewhat lower than in the reference condition.

There is no relation of age or PDS with the social condition boost in learning performance, or with the learning parameters for those conditions relative to the reference condition. In other words, the enhancement of learning due to the increased salience of the social descriptors does not vary consistently with age or pubertal status. While this is consistent with a small number of prior studies, the present study demonstrates this phenomenon using a straightforward design, and a much larger sample size than in previous studies. In light of the evidence upon which the social reorientation hypothesis is based, there are several possible explanations. In the early part of this age range (12 years), information about dating and social status may already be clearly important, and so salience differences may already be at ceiling with respect to task difficulty. As mentioned in the limitations, above, it may also be the case that although the social conditions are somewhat more interesting, salience is not affected by motivation in the relevant motive domains and so is not sensitive to developmental changes. Of course, it is possible that this null effect truly does reflect an underlying, age-stable motivational orientation toward information about dating and social status that is not affected by the developmental-stage-specific neurobiological changes in motivational systems, as proposed by the social reorientation hypothesis. If this is the case, it would be somewhat inconsistent with the hypothesis that underlying motivational changes, rather than changes in the social milieux, are predominantly responsible for behavioral changes during adolescence.

CHAPTER IV

AIM 2: LEARNING DIFFERENCES, SELF-REPORT MOTIVES, AND BEHAVIORAL CORRELATES

Background

Associations between variables derived from behavior on the Social Probabilistic Learning Task and both self-report trait measures, and self-report motive-relevant behavior were examined in order to test whether individual differences in task response were, as expected, driven by differences in social motives. For example, if an individual is relatively more driven by mate-seeking concerns (for example, is especially interested in meeting potential romantic partners), the expectation of this paradigm is that this person should also attend relatively more to information in the environment related to the romantic availability of others. This might result in increased task performance through either decreased random responding (due to overall increased attentiveness) as reflected in the parameter ξ, increased sensitivity to each piece of reward feedback (ρ), or an increased rate of incorporation of new information (ε). Of course, given that both ρ and ε have non-monotonic relations to performance, a linear relation between mate-seeking motive and either of these parameters could also result in non-monotonic relations to performance. For this reason, self-report trait measures and motive-relevant (health-related) behaviors will be examined in relation to performance both in the first and last half of the task, as well as with learning model parameters. There may be relations between overall task ability and social motives as well as other traits, so the minimally social Hungry/Thirsty condition will serve as a baseline by which overall ability on the task is controlled for each individual.

Self-report measures of social motives.

Fundamental Social Motives Inventory (FSMI). The FSMI was developed to capture individual differences in social motives that relate to an evolutionary-psychology framework that has as its foundation the idea that motivational systems have adapted to a set of social problems with direct relevance to survival and reproductive fitness (Neel et al., 2015). The inventory contains seven subscales, of which two, the Mate-seeking and Status scales, are relevant to this investigation of the SPLT. In the paper which introduced this scale, Neel et al. (2015) showed that measurements on both of these scales was associated with other motive-relevant self-report instruments, and motive-relevant behaviors. For example, they found that Mate-seeking was correlated with having asked someone out on a date and Status was correlated with being in a position of leadership at work. Also, as predicted by the Fundamental Social Motives theoretical framework, Neel et al. (2015) found that current relationship status strongly influenced scores on the Mate-seeking questionnaire. This scale was only administered to the two college samples.

Kids' Social Reward Questionnaire (K-SRQ). The original SRQ was developed to measure individual differences in what kinds of social interactions people find rewarding (Foulkes, Viding, McCrory, & Neumann, 2014). Understanding what is rewarding and valuable to people is directly related to what people find motivating, with greater reward derived from a particular domain indicating greater motivation in that domain. In the work that established this measure, validity was investigated using other existing self-report trait scales, with favorable results. The Sexual Relationships subscale, administered to all samples as a measure of mate-seeking motive, correlated strongly with sociosexual orientation measures. Admiration, administered to all samples as a measure of status motive, correlated with dominance, friendliness, hostility, extraversion and narcissism. Two other subscales, though not focal, are reported in this investigation as well: Passivity, which measures enjoyment of giving control to others, was negatively correlated with dominance; Sociability, which measures enjoyment of group interactions, showed a pattern of correlations similar to Admiration, but with a correlation to hostility close to zero. The Kids' SRQ was modified for this study to be more appropriate for adolescents by replacing questions explicitly about sex (e.g., "I enjoy having an active sex life"), with questions that are closely related to sexual and romantic experience (i.e., "I like kissing," "I like having a crush on someone," and "I like flirting").

Dominance and Presige (D&P) Scale. The Dominance and Prestige scale was developed to measure distinct and non-mutually exclusive strategies for attaining social status (Cheng, Tracy, Foulsham, Kingstone, & Henrich, 2013). Dominance refers to the use of physical or social force to exert control over others such that social rank is increased. Prestige refers to the strategy of using expertise and performance to garner respect from peers. These measures were administered to the college sample in part to help reaffirm the validity of the FSMI Status scale, as well as to examine differential associations between SPLT task behavior and particular strategies. For example, it is plausible that those employing dominance strategies to achieve status need to attend more closely to the social rank, or popularity, of others in order to avoid attempting to exert force upon those more powerful. For those employing prestige strategies, where social rank is in part a by-product of excellence in other domains, peer popularity may be less important.

Urgency, Premeditation, Perseverance, and Sensation Seeking (UPPS-P). The constructs of impulsiveness and sensation seeking figure prominently in research on adolescence (e.g, Collado, Felton, MacPherson, & Lejuez, 2014; Harden & Tucker-Drob, 2011; McCabe, Louie, & King, 2015; Roberti, 2004; Romer, 2010; Shulman, Harden, Chein, & Steinberg, 2015; Shulman et al., 2016; Steinberg et al., 2008). These constructs are also necessarily linked to motivation – behaviors must be selected from those already having some motivational weight. For this reason, the UPPS-P (Whiteside, Lynam, Miller, & Reynolds, 2005), a popular measure of these constructs, was included, with subscales used as covariates against which the above measures could be compared. This instrument contains scales for both Positive and Negative Urgency, which are supposed to measure impulsivity related to both positive and negative affective states. It also contains the Sensation Seeking subscale, a measure of enjoyment of activities that pose some amount danger, or uncertainty (i.e., risk). The Premeditation and Perseverance scales, measuring the tendency, respectively, to think before acting and to persevere in the face of difficulty or tedium, were also included.

Behavioral outcomes. Examining the relation between the above constructs and sexual behavior is done primarily to help establish concurrent validity. Adolescence is a time when both hormonal and social changes increase propensity to engage in sexual behavior (Udry, 1988). These outcomes are developmentally normative, and as such these analyses are not intended as investigations into the relation of mate-seeking motives to adolescent (or early adult) *risk-taking* behavior. Of course, these behaviors are health-*relevant* not just because having sex confers risk of sexually transmitted infection, or pregnancy, but also because successful navigation of romantic and sexual relationships is *developmentally healthy* (Harden, 2014). In short, sexual and romantic desire, and behavior, should correlate, as well as show developmentally expected increases during the transition from childhood to sexual maturity and adulthood.

Questions about sexual behavior were derived from the the Sexual Experience Survey used by Smith, Leve, and Chamberlain (2006), which was itself derived from questions administered by Capaldi et al. (2002). They ask about the number of recent sexual partners, and sexual encounters. Health-protecting sexual behavior is assessed by an item asking the participant to report on the frequency of a variety of safe sex practices. In these analyses, this is treated as a health-*risking* behavior given that these behaviors are the primary means for adults and adolescents alike to ameliorate the risks of developmentally normative sexual behavior.

Alcohol use is another common target outcome for research on adolescent motivation and decisions making, and so provides another touchstone by which we might measure the effectiveness of our instruments, as well as provide additional data on the relations between facets of adolescent motivation and health-relevant behaviors. Alcohol consumption has been shown to be related to popularity in adolescence (Allen et al., 2005), and there is some evidence that alcohol consumption can itself be a strategy for gaining peers' admiration (Mallett, Lee, Turrisi, & Larimer, 2009). Social gatherings in adulthood and in adolescence are also, more simply, occasions both for navigating social hierarchies, and for drinking alcohol.

As such, in both the college student and full sample, we expect alcohol consumption to be related to measures of motivation to gain status. We take two outcomes from the Youth Risk Behavior Survey measuring number of days the participant recently had any amount of alcohol, and the number of days where they drank heavily (Centers for Disease Control and Prevention, 2015). Given that desire for socialization as well as status may motivate attendance at occasions where alcohol is a central feature, associations with the measure of socialization motivation are also examined.

Method

The primary goals of these analyses are to establish the scales as measures of the constructs of interest, examine whether they provide any convergent evidence that individual differences in the Social Probabilistic Learning Task (SPLT) are related to motives, explore their association with adolescent health-related behavior, and, finally, to explore the associations between this behavior and learning during the SPLT. Most interpretive decisions will be made on the basis of effect size, without thresholding based on test statistic parametric probabilities. The primary exception to this is during the psychometric testing phase, as described below. Analyses of self-report scales will all employ latent variable models to protect against most types of differential item functioning (Gregorich, 2006), and to properly account for reliability of measurement (Westfall & Yarkoni, 2016). All analyses, unless specified otherwise, were performed using the lavaan package (version 0.6.1.1193; Rosseel, 2012) in R (version 3.4.4; R Core Team, 2018). Data wrangling and plotting was performed with help of the tidyverse packages (version 1.2.1; Wickham, 2017), and the psych package (version 1.7.8; Revelle, 2017).

Scale validation. Quality of measurement is a crucial precursor to drawing inferences from statistical tests, and so guidelines for testing measurement invariance are followed for each scale below (Gregorich, 2006). Invariance between groups based on self-reported gender, sample (when possible), and long-term relationship status (in the college samples) is tested for each self-report scale that is based on a reflective latent variable model (which does not include self-report sexual or substance use behaviors). Although comparisons between reported gender groups, or samples, are not the focus of this investigation, both groupings appear in models in order to examine the sensitivity of results to group characteristics.

For testing metric invariance, which is required to make inferences based on latent covariance structures, we use the McDonalad Non-centrality Fit Index (MFI) because this has been shown to be robust to differences in sample size and magnitude of item loadings (Kang, McNeish, & Hancock, 2016). In simulations performed by Kang et al. (2016), ΔMFI $>$ .01 was seen in less than 1% of invariance tests when the null of measurement invariance between groups is true, though the authors caution against rules-of-thumb like this more generally. In other simulations, ΔMFI $>$ .012 was roughly the cut-off for the 1% level (Flournoy, 2018a). The tests below follow these guidelines roughly: when fit differences in the MFI between invariant, or partially-invariant, scales is more than .012, constraints on particular items (suggested by modification indices) are relaxed until the fit difference is no longer substantially concerning. However, because these items not longer contribute to the model in the same way, and because content coverage of the scales is another important concern for validity (Borsboom, 2006), the final partially invariant scale is not embraced credulously. When substantial modification is necessary, sensitivity analysis is performed to compare estimates between the fully constrained scales and the partially constrained (i.e., partially invariant) measurement models. If results are not too different, then we can be assured that in this case, neither measurement nor content differences are strongly influential on the range of inferences we might make. If results do appear markedly different, then it is difficult to adjudicate which is the more interpretable.

Group equality constraints on the covariance between latent variables will be examined next, with the final model chosen for more extensive interpretation being selected on the basis of parsimonious model fit. Smaller values of the Akaike Information Criterion (AIC), an estimate of out-of-sample predictive accuracy that incorporates model parsimony (Akaike, 1998; Gelman et al., 2014), will be used to determine whether emphasis should be given to models that constrain substantive estimates across groups (e.g., samples). When appropriate, even if the AIC indicates the constraint may provide better predictive accuracy, we can examine differences in estimates between different groups to ensure that this is a reasonable conclusion.

In order to characterize the cross-sectional age associations for each scale, the measurement models will be used to regress each latent variable on a quadratic function of age, with the gender factor interacting with age. These structural models are intended to provide information on how well these associations converge with expectations based on the social reorientation theory (in the case of status and mate-seeking related constructs), as well as expectations based on past research on impulsivity and sensation seeking during adolescence as described above. Specifically, the hope, with regard to validating the construct validity of these scales, is to see higher levels of interest in status-related, and mate-seeking-related constructs in older and more pubertally advanced participants (at least up until physical maturity), and higher levels of both sensation seeking and impulse control (or rather, lower levels of urgency) in older participants. An important caveat to the examination of this descriptive evidence regarding construct validity or more substantive claims is that in cross-sectional data, both convergence and divergence with expectations may be due to factors other than age.

To further examine the case for the validity of these scales as measurements of the constructs they purport to measure, we examine the relations between them. In order to do so, all latent variable measurement models will be combined to examine inter-scale correlations, with attention to convergence between ostensibly related constructs. The college sample was administered all questionnaires, allowing the full inter-scale correlation matrix to be estimated in this sample only. Correlations among latent variables will also be examined across all samples for the K-SRQ and UPPS-P. Special attention will be given to the FSMI Mate-seeking and Status scales, the K-SRQ Sexual Relationships and Admiration scales, and the UPPS-P Sensation Seeking scale.

Self report relation to SPLT. Four parameters from the reinforcement learning model of SPLT behavior, ϵ, ρ, ξ, b, and two proportions of optimal responses during the first and second half of the run, were correlated with all latent variables in the measurement models established in the step above. The variables for the SPLT comprise a value for every parameter, for every condition (Hungry/Thirsty, Popular/Unpopular, Dating/Looking), as well as for the contrasts between the two social conditions and the minimally social condition for a total of 30. The contrast variables for model parameters are calculated on the model scale, such that $0 \leq \varepsilon, \xi \leq 1$ and $0 \leq \rho$, and analyzed without further transformation (differences in proportions are treated similarly). The parameters from the baseline condition are transformed as they were during model estimation, such that $\varepsilon' = \varphi^{-1}(\varepsilon)$, $\xi' = \varphi^{-1}(\xi)$, $\rho' = ln(\rho)$, where φ^{-1} is the quantile function for the unit normal density. The condition-specific (but not contrast) values for proportion of optimal responses were also transformed using the quantile function.

Only the contrasts between the social motive and baseline conditions, and the baseline condition itself are of interest. In particular, the correlations between the SPLT contrast variables and self-report measures are expected to indicate the degree to which motive effects on learning are related to individual differences in self-report motives. The relation between SPLT variables in the baseline condition are relevant for interpreting the possible explanations for individual differences in reinforcement learning more generally. Interpretation of the results will focus on the size and sign of the correlation with regard to the precision as suggested by the 95% confidence intervals.

Self report relation to outcomes. First, the motive-related scales may be further evaluated for construct validity on the basis of their relation to self-reported behaviors which would be expected to result from those motives. For example, in general, those who are in long-term relationships should report lower levels of interest in seeking out new romantic or sexual partners (i.e., lower scores on the FSMI Mate-Seeking scale). Among all participants, higher levels on the K-SRQ Sexual Relationships scale should be related to higher numbers of sexual encounters (specifically, "how many times, in the last six months, have you had sex/sexual relations," with sexual relations defines as "oral, vaginal or anal, and can be with male or female partners"). Among those who are not in long-term relationships, higher FSMI Mate-Seeking and K-SRQ Sexual Relationships should be related to having had more, recent, sexual partners (specifically, "with how many different people in the last 6 months"). Based on the literature relating popularity and status to alcohol consumption, higher scores on FSMI Status and K-SRQ Admiration may be expected to relate to greater number of days having had at least one drink ("In the past 30 days, on how many days did you have at least one drink of alcohol").

The more substantively interesting questions involve the relation between trait scales and both safe-sex behavior and heavy alcohol use. The outcomes of interest are:

- "In the last 6 months, what percentage of the time when you've had sexual relations (oral, vaginal or anal) did you USE safe sex practices to prevent sexually transmitted diseases - condoms, blood tests, dental dams, etc.)"
- "In the past 30 days, on how many days did you have 5 or more drinks of alcohol in a row, that is, within a couple hours?"

The data generating processes for the behavioral outcomes are not Gaussian, nor do they necessarily result in data where a unit difference between responses has the same meaning. To account for this, certain analysis and data transformation decisions were made. For frequency of sex, number of partners, number of past 30 days with alcohol, and number of past 30 days with five or more alcoholic drinks, responses were analyzed as ordinal outcomes. This required binning responses for some variables. Participants were free to input any number for the frequency of sex question, so to ensure that ordered categories contained at least 30 participants, responses were binned into groups with 0, 1-9, 10-19, 20-39, and 40 or more sexual encounters in the past six months. Responses to the number of partners in the last six months was also binned into groups with 0, 1, 2, and 3 or more, with at least 40 participants in each bin. The response options for both alcohol use questions were constrained, with response options, in days, set to 0, 1-2, 3-5, 6-9, 10-19, 20-29, and "All 30". For both responses, very few or no participants selected 20 or more days, and only 6 participants selected 10-19 days for the question asking about having had five or more drinks in one session. Bins on the high end were combined resulting in five bins for the question asking about days having any number of drinks (0, 1-2, 3-5, 6-9, 10 or more), and in four bins for the question about days having five or more drinks (0, 1-2, 3-5, 6 or more). The question regarding safe sex practice was analyzed continuously because the response options asked about continuous frequency. Analysis of categorical outcomes used ordinal regression using a diagonally weighted least squares estimator with robust standard errors and a mean and variance adjusted test statistic (corresponding to the WLSMV estimator in the lavaan package). This estimation method requires listwise deletion for any cases without complete responses on the questionnaires. This results in lower samples sizes for analyses comparing the UPPS-P scales as described below.

Gender and (if the adolescent participants could be included) age were added as covariates to all models. Associations between target motive variables and outcomes were estimated both with and without UPPS-P Sensation Seeking and both Urgency variables as covariates to explore sensitivity of the outcomes to variance shared between these measurements. Also, in the case of K-SRQ variables, sensitivity to the inclusion of highly correlated subscales was also tested. For example, if the Sexual Relationships variable was included as the focal independent variable, the Admiration and Sociability variables were added as covariates.

Interpretation of associations (both with self-report scales, and when considering the SPLT variables) with self report behavior is based on the sign of the focal motive variable coefficient, and on the χ^2 difference test ($\alpha = .005$) between a model including the focal motive variable and covariates, and the same model with the motive variable coefficient fixed to 0. To examine sensitivity to inclusion of the UPPS-P scales, comparison was between the model with all variables (motive, UPPS-P Sensation Seeking, and Positive and Negative Urgency), and the same model with the motive coefficient fixed to 0. That is, the test examined fit decrease due to excluding the relation between the outcome and motive variable. Sensitivity to the addition of K-SRQ subscales tested models with the zero-coefficient constraint set on the added variable, with examination of size and sign change of the focal motive variable coefficient.

Finally, the subset of these analyses examining associations with number of sexual partners are estimated using only the participants who report not being in a long-term relationship. For completeness, the same analysis is repeated without regard for long-term relationship status and note is made of any substantial differences.

SPLT relation to outcomes. The same behavioral outcomes are also regressed on the relevant contrasts of all SPLT variables. Specifically, self report sexual behavior is regressed on the Dating/Looking - Hungry/Thirsty contrast for each parameter, $\varepsilon, \rho, \xi, b$, and the same contrast for the proportion of optimal decisions made during both the first and last half of the run. Self-report drinking is regressed on the Popular/Unpopular - Hungry/Thirsty contrasts of these variables. An initial model is fit to a limited sample that includes just unpartnered college students (sexual behavior outcomes), or just college students (alcohol consumption outcomes), and another model is fit using both college and both adolescent samples. Gender and age (when the both college and adolescent samples are combined) are included as covariates. The test of association between the SPLT variable and the outcome is, as above, an adjusted $\Delta\chi^2$ test, with decision cut-off of $\alpha = .005$ (note, however, that this procedure was not established *a priori*).

Table 12. Descriptive statistics for self-report scales, college-samples

Scale name	Sample	α	λ-6	Mean	SD
(F) Mate-seeking	CSYA	0.92	0.92	3.94	1.65
	CSYA-O	0.85	0.85	3.86	1.46
(F) Status	CSYA	0.67	0.68	4.71	0.85
	CSYA-O	0.71	0.69	4.52	0.87
(D) Dominance	CSYA	0.78	0.79	3.11	0.96
	CSYA-O	0.84	0.84	3.62	1.03
(D) Presitige	CSYA	0.83	0.86	5.21	0.82
	CSYA-O	0.78	0.81	5.02	0.70

FCA: foster-care-involved adolescents; CA: community adolescents; CSYA: college students; CSYA-O: college students, online. SD = Standard Deviation;F = Fundamental Social Motives Inventory; D = Dominance and Prestige Scale.

Table 13. Descriptive statistics for K-SRQ

Scale name	Sample	α	λ-6	Mean	SD
(K) Admiration	FCA	0.89	0.88	5.59	1.38
	CA	0.74	0.68	5.88	0.84
	CSYA	0.84	0.81	5.46	1.03
	CSYA-O	0.87	0.84	5.40	1.04
(K) Passivity	FCA	0.72	0.66	3.07	1.22
	CA	0.81	0.74	2.92	1.38
	CSYA	0.75	0.70	3.27	1.27
	CSYA-O	0.80	0.73	3.20	1.27
(K) Sexual Rel.	FCA	0.76	0.76	5.22	1.27
	CA	0.80	0.74	5.25	1.33
	CSYA	0.77	0.72	5.62	1.10
	CSYA-O	0.73	0.66	5.52	1.17
(K) Sociability	FCA	0.57	0.49	4.96	1.18
	CA	0.51	0.47	5.06	1.16
	CSYA	0.51	0.42	5.51	1.01
	CSYA-O	0.69	0.62	5.24	1.22

FCA: foster-care-involved adolescents; CA: community adolescents; CSYA: college students; CSYA-O: college students, online. SD = Standard Deviation; K = Kids' Social Reward Questionnaire.

Results

Evaluating scales. Descriptive statistics including mean, standard deviation, Cronbach's α, and Guttman's λ-6 coefficient are provided for each computed scale, for each sample. These statistics were computed as means of all items responded to, including cases with missing values, and with no imputation, using the alpha function in the psych package (Revelle, 2017). Though all scales will be evaluated using latent variable measurement models, these statistics may provide a helpful starting point of reference. Statistics for scales administered to the college samples, only, appear in Table 12, and for the K-SRQ and the UPPS-P, administered to both adolescent and both college samples, can be found in Tables 13 and 14.

Table 14. Descriptive statistics for UPPS-P

Scale name	Sample	α	λ-6	Mean	SD
(U) -Urgency	FCA	0.87	0.90	2.86	0.61
	CA	0.89	0.91	2.91	0.63
	CSYA	0.87	0.89	2.76	0.56
	CSYA-O	0.88	0.89	2.60	0.56
(U) Premeditation	FCA	0.78	0.86	3.04	0.45
	CA	0.83	0.87	3.04	0.49
	CSYA	0.71	0.81	3.15	0.37
	CSYA-O	0.82	0.85	3.09	0.43
(U) Perseverance	FCA	0.84	0.88	2.79	0.55
	CA	0.80	0.84	3.00	0.47
	CSYA	0.84	0.89	3.14	0.48
	CSYA-O	0.79	0.81	3.04	0.42
(U) Sensation S.	FCA	0.77	0.86	2.27	0.50
	CA	0.80	0.87	2.24	0.57
	CSYA	0.80	0.87	2.07	0.49
	CSYA-O	0.83	0.86	2.25	0.56
(U) +Urgency	FCA	0.86	0.94	3.17	0.46
	CA	0.88	0.92	3.16	0.51
	CSYA	0.84	0.90	3.06	0.43
	CSYA-O	0.87	0.90	2.93	0.51

FCA: foster-care-involved adolescents; CA: community adolescents; CSYA: college students; CSYA-O: college students, online. SD = Standard Deviation; U = Urgency, Premeditation, Perseverance, Sensation Seeking Scale.

Fundamental Social Motives Inventory. The Fundamental Social Motives Inventory (FSMI) was administered to the college sample only (both because the questions are not appropriate for younger adolescents, and because time did not allow for this additional inventory). For the specific purposes of this study, the analyses below focus on the Mate-seeking and Status factors. Invariance of loadings of items on factors (metric invariance) between participants reporting male versus female gender on the pubertal development scale (PDS) was tested and found acceptable, with a slight increase in fit after constraining factor loadings ($\Delta\text{MFI} = 0.005$). The final confirmatory factor analysis (CFA) for mate-seeking and status factors fit well (N_{total} = 224, RMSEA = 0.059, MFI = 0.90, $\hat{\gamma}$ = 0.97). Adding equality constraints on the latent covariance structure resulted in better parsimony-adjusted model fit as measured by the differences in the Akaike information criterion (AIC; $\Delta\text{AIC} = -5$), with the two factors showing a very small, positive, non-significant correlation ($r = 0.06, Z = 0.69$). Note that item 54, “I do not worry very much about losing status,” has a particularly low loading on the status factor (β = -0.29). Associations with age and pubertal development were not examined given the narrow range of each of these measurements in this sample.

Dominance and Prestige. The Dominance and Prestige scale (D&P) was also administered to the college sample only (because time did not allow for it). Invariance of loadings between males and females was tested and found to decrease fit ($\Delta\text{MFI} = -0.018$). Modification indices indicated that freeing the constraint of equality between groups for the item, “Some people are afraid of me”, would improve fit, and doing so resulted in an acceptable decrease in fit between the (partially) metric invariant and unconstrained models (ΔMFI $= -0.011$). Further improvements were not sought as relaxing constraints on this single item both broadly protects the scale’s content coverage (Borsboom, 2006), while providing fairly acceptable fit differences. This partial-metric invariance model is the model used in all analyses below, and sensitivity of results to the relaxation of this constraint is not tested. The final, partially metric invariant CFA for Dominance and Prestige factors fit somewhat poorly (N_{total} = 224, RMSEA = 0.096, MFI = 0.56, $\hat{\gamma}$ = 0.881). Further constraining the latent covariance and variance to be equal between both females and males resulted in slightly better model fit (ΔAIC $= -4$), with the two factors showing a small, positive, correlation ($r = 0.23, Z = 2.77$). Associations with age and pubertal development were not examined given the narrow range of each of these measurements in this sample.

Kids’ Social Reward Questionnaire. The Kids’ Social Reward Questionnaire (K-SRQ), and remaining self-report instruments, were administered to each college and adolescent samples. The Negative Social Potency and Prosocial Interactions factor indicators were almost all very highly skewed, so these subscales were dropped before any further analysis was performed. The four factors included below are: Admiration, Passivity, Sexual Relationships, and Sociability.

A CFA for the four K-SRQ factors was tested for metric invariance across samples by comparing a model with loadings constrained to equality to one with loadings free to vary across the four groups. This resulted in considerably worse model fit (ΔMFI $= -0.036$). Loading constraints were removed from four items (“I like it if others looks up to me”, “I like being a member of a group/club”, “I like kissing”, “I like flirting”) that load onto Admiration, Sociability, and Sexual Relationships, respectively. This resulted in adequately small fit differences between the unconstrained and partial-metric invariance models (ΔRMSEA = -0.00083, ΔMFI = -0.009, $\Delta\hat{\gamma}$ = -0.003). The final model showed somewhat poor fit (N_{total} = 319, RMSEA = 0.09, MFI = 0.77, $\hat{\gamma}$ = 0.924).

Metric invariance was also evaluated between reported gender groups. The partial-metric invariance model from above was compared to the model without loadings constrained. The fit decrease between these two models was not too big (ΔMFI $= -0.01$).

Next, models with and without between-sample equality constraints on latent variances and covariances were compared, resulting in better fit ($\Delta \text{AIC} = -11$) for the constrained model. Loadings without constraints across groups (in the partial-metric invariance model) do not contribute to this comparison, so the magnitude of correlations was compared across the partial-metric invariant and fully constrained models (both with the constraint of covariance across groups). The differences in correlations were in the range r = [-0.02, 0.02], indicating very low sensitivity the constraints on item loadings. To maintain content coverage, the fully constrained model, with equality of loadings across all samples, is used, though sensitivity to constraints on the four non-invariant loadings is assessed again in subsequent analyses.

Correlations among the Admiration, Sociability, and Sexual Relationships subscales were all high ($r = [0.66, 0.84]$). Correlations of Passivity with these three subscales were generally small, but positive ($r = [0.07, 0.21]$; see Tables 19 and 20 for all latent variable correlations). This pattern of correlation is similar to what was seen in the scale construction sample (Foulkes et al., 2014).

All items were forward-coded, so one possible explanation for the high inter-scale correlations is acquiescence bias. To explore this possibility, we examined the association between attention-check items and the K-SRQ subscales for those who had data for both (N = 138). Spearman rank-order correlations were calculated between the computed scale scores and two attention-check items: “This is a control question, please select ‘Strongly disagree’ and continue,” and “Most people find puppies cute.” Correlations between these two questions and K-SRQ subscales were somewhat high for Admiration ($r_s = -.23, .34$, respectively), Sexual Relationships ($r_s = -.16, .27$), and Sociability ($r_s = -.17, .35$), but smaller for Passivity ($r_s = .07, 0$).

Age

Variation in latent factors was described by a quadratic age regression, with age effects allowed to vary by gender. The partial metric invariance model was used to test whether constraining regression coefficients across samples resulted in a better fitting model as assessed by the AIC. Constraining the regression coefficients across sample resulted in reduced AIC ($\Delta \text{AIC} = -67$), indicating better model fit. Standardized coefficients were extremely similar between the model with partial metric invariance and the model with all loadings constrained to equality across samples ($\Delta\beta = [-0.04, 0.02]$), so the fully constrained model is interpreted. Model-expected trend lines show somewhat higher scores in

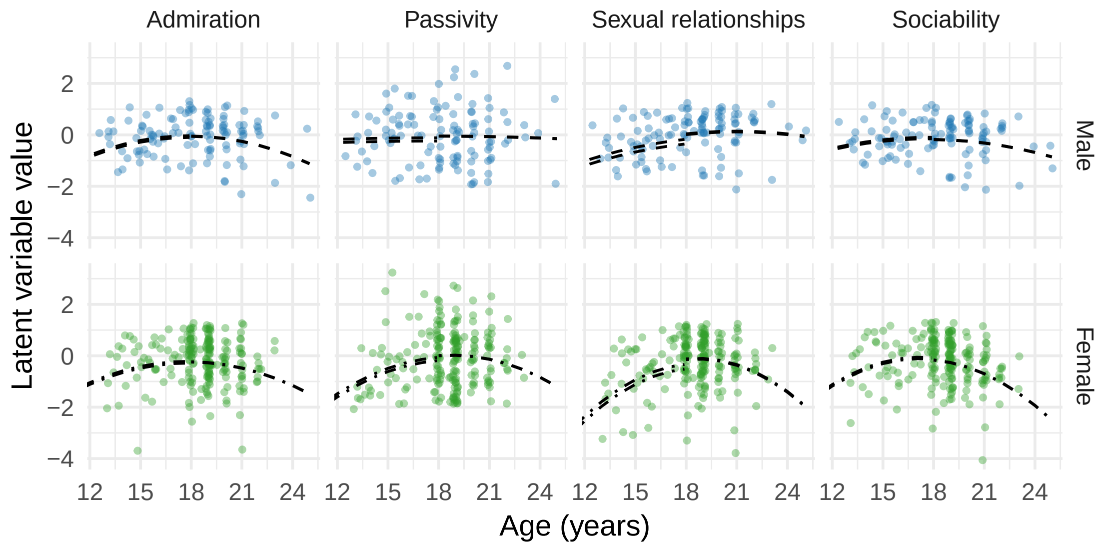

Figure 26. K-SRQ model-predicted age associations. Best fit lines are expected latent variable values drawn from the model with all loadings, latent variances and covariances, and regression coefficients constrained to be the same across samples. Best fit lines are plotted separately for each sample, but because regression coefficients are identical, and the expected trajectories are otherwise very similar, they are not separately labeled. Points are individual, model predicted, latent variable values.

late adolescent participants (Figure 26), with the largest quadratic effects (averaged across gender) for Admiration, Sociability, and Sexual Relationships (Table 15). Gender differences in the linear model coefficients were fairly small.

Puberty

We also described variation in latent factors by a quadratic PDS regression, with effects allowed to vary by gender. The partial metric invariance model with constrained regression coefficients across samples resulted in a better fit (ΔAIC $= -58$). Again, standardized coefficients were extremely similar between the model with partial metric invariance and the model with all loadings constrained to equality across samples ($\Delta\beta = [-0.05, 0.05]$), so the fully constrained model is interpreted. Model-expected trend lines show somewhat lower scores associated with lower PDS, primarily for female participants (Figure 27), with the largest linear effects (averaged across gender) for Sociability, and Sexual Relationships (Table 16). Though a model including the quadratic effect of PDS was estimated, this was done to allow deceleration in the linear trend to approximate the observed asymptotic relation.

Table 15. K-SRQ factors regressed on age and gender

K-SRQ factor	Sample	Age	Age × Gender	Age^2	Age^2 × Gender
Admiration	FCA	-0.01 [-0.17, 0.15]	-0.01 [-0.14, 0.13]	-0.17 [-0.31, -0.03]	-0.01 [-0.19, 0.16]
	CA	-0.01 [-0.15, 0.13]	-0.01 [-0.17, 0.16]	-0.19 [-0.34, -0.04]	-0.02 [-0.25, 0.21]
	CSYA	-0.01 [-0.15, 0.13]	-0.01 [-0.14, 0.13]	-0.23 [-0.41, -0.05]	-0.02 [-0.23, 0.20]
	CSYA-O	-0.01 [-0.12, 0.11]	-0.00 [-0.11, 0.10]	-0.13 [-0.23, -0.02]	-0.01 [-0.14, 0.12]
Passivity	FCA	0.03 [-0.11, 0.17]	0.05 [-0.08, 0.17]	-0.09 [-0.21, 0.03]	-0.10 [-0.25, 0.05]
	CA	0.03 [-0.09, 0.15]	0.06 [-0.09, 0.21]	-0.10 [-0.24, 0.03]	-0.13 [-0.33, 0.06]
	CSYA	0.03 [-0.10, 0.16]	0.05 [-0.08, 0.17]	-0.13 [-0.30, 0.04]	-0.13 [-0.31, 0.06]
	CSYA-O	0.02 [-0.08, 0.13]	0.04 [-0.06, 0.14]	-0.07 [-0.16, 0.02]	-0.08 [-0.19, 0.03]
Sexual rel.	FCA	0.12 [-0.04, 0.27]	-0.01 [-0.15, 0.13]	-0.20 [-0.32, -0.08]	-0.15 [-0.30, 0.00]
	CA	0.10 [-0.03, 0.23]	-0.01 [-0.18, 0.16]	-0.21 [-0.34, -0.09]	-0.19 [-0.37, -0.00]
	CSYA	0.11 [-0.04, 0.25]	-0.01 [-0.15, 0.13]	-0.28 [-0.45, -0.10]	-0.19 [-0.38, 0.00]
	CSYA-O	0.09 [-0.03, 0.21]	-0.01 [-0.12, 0.11]	-0.15 [-0.25, -0.05]	-0.11 [-0.23, 0.00]
Sociability	FCA	-0.06 [-0.25, 0.12]	-0.07 [-0.21, 0.08]	-0.19 [-0.34, -0.04]	-0.12 [-0.30, 0.06]
	CA	-0.06 [-0.22, 0.11]	-0.09 [-0.27, 0.10]	-0.21 [-0.38, -0.05]	-0.16 [-0.39, 0.07]
	CSYA	-0.06 [-0.22, 0.10]	-0.07 [-0.21, 0.08]	-0.26 [-0.45, -0.06]	-0.15 [-0.36, 0.07]
	CSYA-O	-0.05 [-0.18, 0.09]	-0.05 [-0.17, 0.06]	-0.14 [-0.25, -0.03]	-0.09 [-0.22, 0.04]

FCA: foster-care-involved adolescents; CA: community adolescents; CSYA: college students; CSYA-O: college students, online. Values are standardized regression coefficients from the model with all loadings, latent variances and covariances, and regression coefficients constrained to be the same across samples, with 95% confidence intervals in brackets.

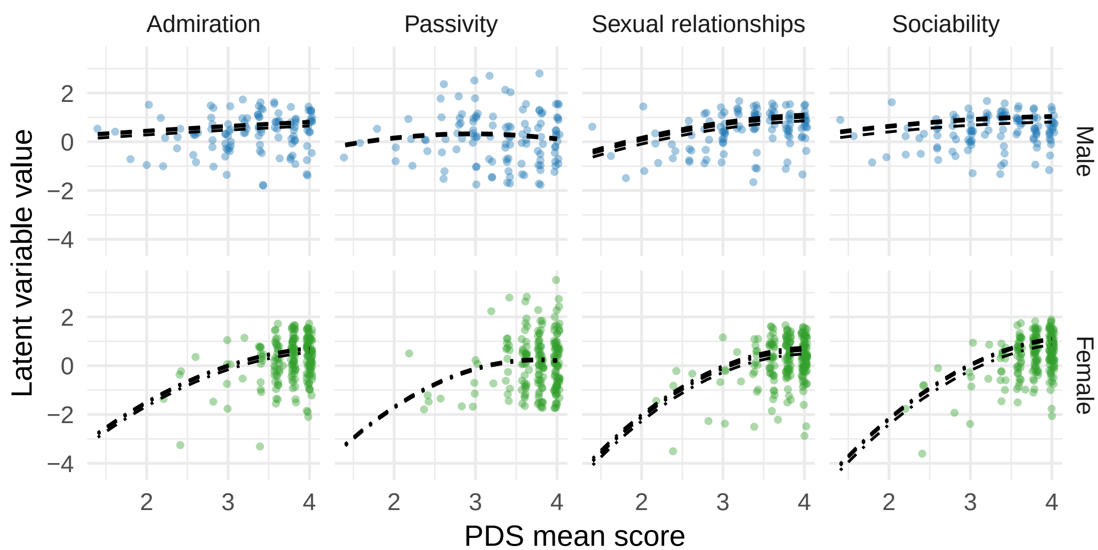

Figure 27. K-SRQ model-predicted PDS associations. Best fit lines are expected latent variable values drawn from the model with all loadings, latent variances and covariances, and regression coefficients constrained to be the same across samples. Best fit lines are plotted separately for each sample, but because regression coefficients are identical, and the expected trajectories are otherwise very similar, they are not separately labeled. Points are individual, model predicted, latent variable values.

Table 16. K-SRQ factors regressed on age and gender

K-SRQ factor	Sample	PDS	PDS × Gender	PDS^2	PDS^2 × Gender
Admiration	FCA	0.44 [0.12, 0.75]	0.28 [0.02, 0.54]	-0.10 [-0.32, 0.12]	-0.14 [-0.45, 0.17]
	CA	0.39 [0.13, 0.65]	0.26 [0.03, 0.48]	-0.09 [-0.28, 0.10]	-0.11 [-0.36, 0.14]
	CSYA	0.19 [0.05, 0.32]	0.38 [0.02, 0.75]	-0.07 [-0.23, 0.08]	-0.14 [-0.46, 0.18]
	CSYA-O	0.27 [0.08, 0.45]	0.35 [0.03, 0.67]	-0.07 [-0.22, 0.08]	-0.13 [-0.44, 0.17]
Passivity	FCA	0.24 [-0.06, 0.54]	0.22 [-0.04, 0.48]	-0.15 [-0.38, 0.07]	-0.12 [-0.43, 0.20]
	CA	0.22 [-0.05, 0.49]	0.21 [-0.03, 0.44]	-0.14 [-0.34, 0.07]	-0.10 [-0.37, 0.17]
	CSYA	0.10 [-0.03, 0.23]	0.30 [-0.05, 0.66]	-0.11 [-0.28, 0.05]	-0.12 [-0.45, 0.21]
	CSYA-O	0.15 [-0.04, 0.33]	0.28 [-0.04, 0.60]	-0.11 [-0.27, 0.05]	-0.12 [-0.43, 0.20]
Sexual rel.	FCA	0.59 [0.23, 0.96]	0.26 [-0.04, 0.56]	-0.19 [-0.43, 0.06]	-0.16 [-0.52, 0.20]
	CA	0.53 [0.24, 0.82]	0.23 [-0.01, 0.48]	-0.16 [-0.37, 0.04]	-0.13 [-0.41, 0.16]
	CSYA	0.26 [0.09, 0.42]	0.36 [-0.05, 0.77]	-0.14 [-0.32, 0.04]	-0.16 [-0.53, 0.21]
	CSYA-O	0.37 [0.15, 0.58]	0.32 [-0.03, 0.68]	-0.13 [-0.30, 0.04]	-0.15 [-0.50, 0.19]
Sociability	FCA	0.66 [0.28, 1.03]	0.46 [0.13, 0.78]	-0.17 [-0.43, 0.09]	-0.19 [-0.55, 0.17]
	CA	0.55 [0.29, 0.81]	0.39 [0.16, 0.63]	-0.14 [-0.35, 0.06]	-0.15 [-0.42, 0.12]
	CSYA	0.28 [0.12, 0.45]	0.64 [0.19, 1.08]	-0.13 [-0.32, 0.07]	-0.20 [-0.57, 0.17]
	CSYA-O	0.39 [0.19, 0.60]	0.56 [0.19, 0.92]	-0.12 [-0.29, 0.06]	-0.18 [-0.52, 0.15]

FCA: foster-care-involved adolescents; CA: community adolescents; CSYA: college students; CSYA-O: college students, online.
Values are standardized regression coefficients from the model with all loadings, latent variances and covariances, and regression coefficients constrained to be the same across samples, with 95% confidence intervals in brackets.

Urgency, Premedidation, Perseverence, Sensation Seeking and Positive Urgency. The Urgency, Premeditation, Perseverance, Sensation Seeking and Positive Urgency (UPPS-P) scale was also administered to all four samples. This instrument uses a large number of indicators for each latent factor. When doing an analysis with a separate group for each sample, this results in a number of variables greater than the number of observations. For this reason, measurement is assessed between combined adolescent and combined college-student samples.

A CFA for the five UPPS-P factors was tested for metric invariance across age groups by comparing a model with loadings constrained to equality to one with loadings free to vary across the four groups. This resulted in adequately small fit differences between the unconstrained and invariant models (ΔRMSEA = 0.00010, ΔMFI = -0.001, $\Delta\hat{\gamma}$ = -0.003). Metric invariance was also evaluated between reported gender groups, with practically no fit decrease ($\Delta\text{MFI} = -0.0005$).

The fit statistics show mixed results, with acceptable RMSEA but also low MFI and $\hat{\gamma}$ (N_{total} = 301, RMSEA = 0.07, MFI = 0.01, $\hat{\gamma}$ = 0.766). Other work has reported similarly ambiguous fit statistics for the UPPS-P in confirmatory factor analysis (Cyders, 2013). This may be the reason for relatively similar degrees of fit between the unconstrained and metric invariance models. We proceed with these analyses with the caveat that we cannot be confident latent variable covariance is not contaminated by differential item use.

Table 17. UPPS-P factors regressed on age and gender

UPPS-P factor	Sample	Age	Age × Gender	Age^2	Age^2 × Gender
Neg Urgency	Adolescent	0.01 [-0.14, 0.16]	-0.13 [-0.28, 0.01]	-0.07 [-0.21, 0.07]	-0.17 [-0.37, 0.02]
	College	0.01 [-0.11, 0.13]	-0.10 [-0.21, 0.01]	-0.07 [-0.20, 0.06]	-0.14 [-0.29, 0.02]
Perseverance	Adolescent	-0.07 [-0.22, 0.08]	0.04 [-0.10, 0.18]	0.08 [-0.05, 0.21]	-0.21 [-0.37, -0.04]
	College	-0.06 [-0.19, 0.07]	0.03 [-0.08, 0.14]	0.08 [-0.05, 0.20]	-0.17 [-0.30, -0.03]
Pos Urgency	Adolescent	-0.03 [-0.17, 0.10]	-0.01 [-0.16, 0.13]	-0.05 [-0.18, 0.09]	-0.11 [-0.30, 0.08]
	College	-0.03 [-0.14, 0.09]	-0.01 [-0.12, 0.10]	-0.04 [-0.17, 0.09]	-0.09 [-0.24, 0.07]
Premeditation	Adolescent	0.12 [-0.05, 0.28]	0.04 [-0.13, 0.20]	0.02 [-0.14, 0.18]	-0.07 [-0.28, 0.14]
	College	0.10 [-0.04, 0.23]	0.03 [-0.10, 0.15]	0.02 [-0.13, 0.17]	-0.05 [-0.22, 0.11]
Sensation Seeking	Adolescent	0.00 [-0.14, 0.14]	-0.11 [-0.27, 0.05]	0.02 [-0.11, 0.14]	-0.19 [-0.37, -0.01]
	College	0.00 [-0.11, 0.11]	-0.08 [-0.19, 0.03]	0.02 [-0.10, 0.13]	-0.15 [-0.28, -0.01]

Values are standardized regression coefficients from the model with all loadings, latent variances and covariances, and regression coefficients constrained to be the same across samples, with 95% confidence intervals in brackets.

Models with and without constraints on the equality of latent variances and covariances between age-groups were compared, resulting in lower AIC ($\Delta AIC = -13$). Absolute magnitude of correlations among all scales were between small and large (r = [0.13, 0.48]), with a very large correlation between Negative and Positive Urgency ($r = 0.78$; see Tables 19 and 20 for all latent variable correlations). Scales keyed in the positively valenced direction (Premeditation and Perseverance) were negatively correlated with negatively valenced scales (Negative and Positive Urgency, and Sensation Seeking).

Age

As with the K-SRQ, the association between UPPS-P factors and age was examined by regressing the latent variables on the interaction between a quadratic age model and gender. Again, constraining the regression coefficients across sample results in reduced AIC ($\Delta AIC = -22$), indicating better model fit. The model expectations reveal that the association between age and factor score may be conditional on reported gender. Plots for Negative and Positive Urgency, Perseverance, and Sensation seeking show diverging quadratic trends for females and males (though note that the expected correlation with age is fairly flat for both genders for Premeditation; Figure 28). The standardized effects indicate that, averaged across gender, linear and quadratic associations were very small, but that for Perseverance, and Sensation seeking, the gender difference in the quadratic term is especially robust (Table 17).

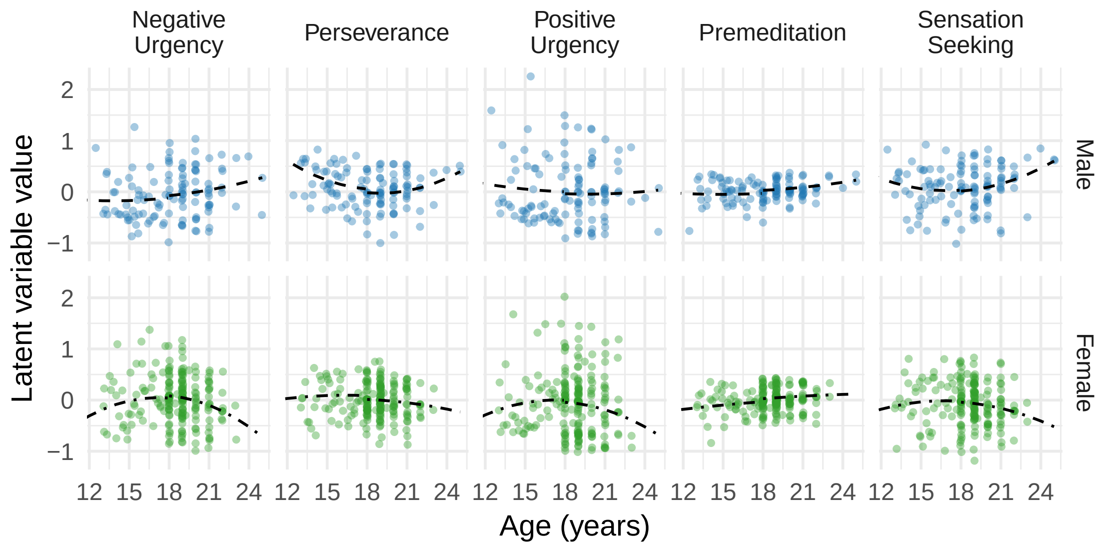

Figure 28. UPPS-P model-predicted age associations. Best fit lines are expected latent variable values drawn from the model with all loadings, latent variances and covariances, and regression coefficients constrained to be the same across samples. Best fit lines are plotted separately for each sample, but because regression coefficients are identical, and the expected trajectories are otherwise very similar, they are not separately labeled. Points are individual, model predicted, latent variable values.

Puberty

The association between the UPPS-P latent factors and PDS was described using the same procedures as above. Constraining the regression coefficients across sample results in better fit (ΔAIC $= -17$). Model expected trend lines show very flat relations between PDS and latent factor scores with the exception of an accelerating quadratic fit and negative linear slope (tangent at PDS = 3) for Positive Urgency (Figure 29). Standardized regression coefficients echo the plots, all having small magnitude, with the exception of the larger negative slope for Positive Urgency (Table 18).

All scale correlations. All measurement models were combined in order to estimate correlations among the latent factors. Though each of the measurement models show acceptable fit, this combine model shows very poor fit (college samples: $N = 224$, RMSEA $= 0.19$, MFI $= 0$; combined samples: $N = 328$, RMSEA $= 0.23$, MFI $= 0$), indicating that this set of latent variables does not reproduce well the observed covariance among the indicators. The possible reasons for this are discussed, below, but not further explored analytically. Correlations are described within the full college sample for all questionnaires, with a combined model with the K-SRQ and the UPPS-P across all four samples.

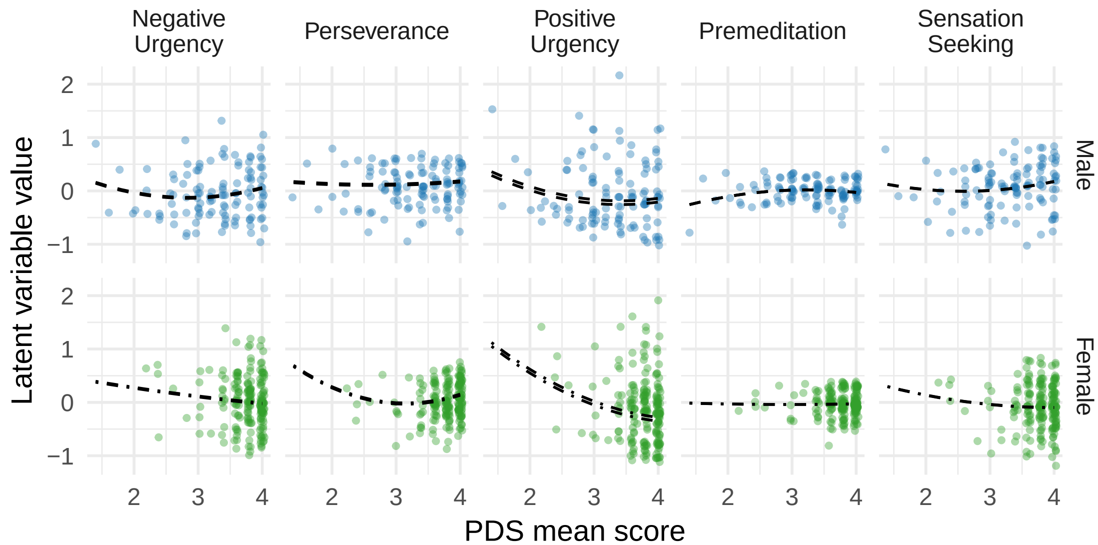

Figure 29. UPPS-P model-predicted PDS associations. Best fit lines are expected latent variable values drawn from the model with all loadings, latent variances and covariances, and regression coefficients constrained to be the same across samples. Best fit lines are plotted separately for each sample, but because regression coefficients are identical, and the expected trajectories are otherwise very similar, they are not separately labeled. Points are individual, model predicted, latent variable values.

Table 18. K-SRQ factors regressed on age and gender

UPPS-P factor	Sample	PDS	PDS × Gender	PDS^2	PDS^2 × Gender
Neg Urgency	Adolescent	-0.06 [-0.31, 0.19]	-0.10 [-0.32, 0.12]	0.07 [-0.13, 0.27]	-0.07 [-0.34, 0.19]
	College	-0.04 [-0.18, 0.11]	-0.14 [-0.43, 0.16]	0.05 [-0.10, 0.20]	-0.08 [-0.38, 0.22]
Perseverance	Adolescent	-0.04 [-0.29, 0.22]	-0.07 [-0.29, 0.16]	0.16 [-0.05, 0.36]	0.16 [-0.12, 0.43]
	College	-0.02 [-0.17, 0.13]	-0.09 [-0.39, 0.21]	0.12 [-0.04, 0.27]	0.17 [-0.13, 0.47]
Pos Urgency	Adolescent	-0.26 [-0.49, -0.02]	-0.14 [-0.35, 0.07]	0.10 [-0.10, 0.29]	0.01 [-0.25, 0.26]
	College	-0.15 [-0.29, -0.01]	-0.19 [-0.48, 0.10]	0.07 [-0.07, 0.22]	0.01 [-0.28, 0.30]
Premeditation	Adolescent	0.05 [-0.20, 0.30]	-0.04 [-0.26, 0.18]	-0.07 [-0.27, 0.13]	0.12 [-0.15, 0.38]
	College	0.03 [-0.12, 0.17]	-0.06 [-0.36, 0.24]	-0.05 [-0.20, 0.10]	0.13 [-0.17, 0.43]
Sensation Seeking	Adolescent	-0.03 [-0.28, 0.23]	-0.11 [-0.34, 0.11]	0.07 [-0.13, 0.28]	-0.02 [-0.29, 0.25]
	College	-0.01 [-0.16, 0.13]	-0.15 [-0.45, 0.15]	0.05 [-0.10, 0.20]	-0.02 [-0.32, 0.27]

Values are standardized regression coefficients from the model with all loadings, latent variances and covariances, and regression coefficients constrained to be the same across samples, with 95% confidence intervals in brackets.

College samples

As was done above, the model constrained all loadings (except those freed above) and the latent covariance structure to be equal across gender groups, which fit better than allowing covariance to vary between genders (ΔAIC $= -56$). Further constraining the model to have equal loadings for all items resulted in only trivial differences in the latent variable correlation magnitudes (maximum absolute $\Delta r =$ 0.01). This same pattern held when examining differences across partnered and unpartnered status (ΔAIC $=$ -46; maximum absolute $\Delta r = 0.04$).

The primary variables of interest are FMSI Mate-seeking, FSMI Status, K-SRQ Admiration, K-SRQ Sexual Relationships, and UPPS-P Sensation seeking. These scales were, *a priori*, intended to measure mate-seeking and status motivation. All latent variable correlations for the comprehensive measurement battery administered to the college samples can be found in Table 19. The FSMI Mate-seeking variable shows correlations in the small-to-moderate range with most of the other variables, including both K-SQR Sexual Relationships and UPPS-P Sensation Seeking. The FSMI Status variable showed more robust correlations with closely related variables. Correlations with Dominance and Prestige status-strategies were large and very large, respectively, while the correlation with admiration was also very large. Interestingly, the association between FSMI Status and K-SRQ Sexual Relationships was somewhat large. The correlation between FSMI Status and K-SRQ Sociability is also large. The strength of association between K-SRQ Admiration and Sexual Relationships is extremely large, as is the correlation between these variables and the K-SRQ Sociability scale. Sensation Seeking shows medium correlations with the FMSI Status variable, both Dominance and Prestige variables, as well as K-SRQ Sexual Relationships and Sociability scales, although the size of these correlations is of roughly the same magnitude as the correlations between Sensation Seeking and other UPPS-P variables.

College and adolescent samples

As was done above, the model constrained all loadings (except those freed above) and the latent covariance structure to be equal across subsamples, which fit better than allowing covariance to vary between them (ΔAIC $= -70$). Further constraining the model to have equal loadings for all items resulted in only trivial differences in the latent variable correlation magnitudes (maximum absolute $\Delta r = 0.02$).

Table 19. Latent variable correlations, college sample

Latent variable	1	2	3	4	5	6	7
1. Mate-Seeking[F]		[-.09, .23]	[-.14, .15]	[-.17, .13]	[.14, .43]	[-.25, .05]	[-.32,-.02]
2. Status[F]	.07		[.35, .63]	[.62, .83]	[.01, .35]	[.12, .45]	[.26, .57]
3. Dominance[D]	.01	.49		[.07, .37]	[.36, .62]	[-.31, .01]	[-.11, .22]
4. Prestige[D]	-.02	.72	.22		[-.25, .07]	[.23, .52]	[.45, .69]
5. Negative Urgency[U]	.29	.18	.49	-.09		[-.52,-.25]	[-.49,-.20]
6. Premeditation[U]	-.10	.28	-.15	.38	-.39		[.38, .64]
7. Perseverance[U]	-.17	.41	.05	.57	-.35	.51	
8. Sensation Seeking[U]	.15	.31	.35	.24	.24	-.23	.23
9. Positive Urgency[U]	.28	.14	.39	-.11	.77	-.40	-.29
10. Admiration[K]	.12	.65	.17	.49	.06	.16	.29
11. Passivity[K]	.22	-.10	-.15	-.17	.20	.00	-.33
12. Sexual Relationships[K]	.22	.43	.15	.41	.19	-.03	.20
13. Sociability[K]	.22	.50	-.06	.41	.09	-.03	.17

Latent variable	8	9	10	11	12	13
1. Mate-Seeking[F]	[-.00, .30]	[.14, .42]	[-.02, .27]	[.08, .37]	[.07, .38]	[.06, .38]
2. Status[F]	[.14, .48]	[-.03, .31]	[.53, .77]	[-.27, .07]	[.27, .59]	[.35, .66]
3. Dominance[D]	[.20, .50]	[.25, .52]	[.02, .32]	[-.31, .01]	[-.01, .32]	[-.23, .12]
4. Prestige[D]	[.09, .40]	[-.26, .05]	[.36, .61]	[-.32,-.01]	[.26, .56]	[.26, .56]
5. Negative Urgency[U]	[.08, .39]	[.70, .84]	[-.09, .22]	[.04, .36]	[.02, .35]	[-.08, .26]
6. Premeditation[U]	[-.38,-.08]	[-.53,-.26]	[.00, .32]	[-.16, .17]	[-.20, .15]	[-.21, .14]
7. Perseverance[U]	[.07, .38]	[-.43,-.14]	[.14, .44]	[-.48,-.18]	[.03, .37]	[-.00, .34]
8. Sensation Seeking[U]		[.12, .41]	[-.05, .28]	[-.26, .07]	[.15, .48]	[.08, .43]
9. Positive Urgency[U]	.27		[-.23, .08]	[.02, .33]	[-.19, .15]	[-.29, .05]
10. Admiration[K]	.12	-.07		[.01, .31]	[.59, .80]	[.70, .88]
11. Passivity[K]	-.10	.18	.16		[-.06, .28]	[.06, .40]
12. Sexual Relationships[K]	.32	-.02	.70	.11		[.72, .94]
13. Sociability[K]	.25	-.12	.79	.23	.83	

F = Fundamental Social Motives Inventory; D = Dominance & Prestige scale; U = Urgency, Premeditation, Perseverance, Sensation-seeking, & Positive Urgency scale; K = Kids' Social Reward Questionnaire.

Correlations between latent variables appear in the lower triangle; 95% confidence intervals appear in the upper triangle. Source model constrains correlations and loadings for all items to be equal across gender groups.

Of the variables measured in all four samples, those of primary interest are K-SRQ Admiration, K-SRQ Sexual Relationships, and UPPS-P Sensation seeking. The first two scales were, *a priori*, intended to measure mate-seeking and status motivation, while the UPPS-P Sensation Seeking scale is intended as a reference point for other research on adolescence. All correlations across the full sample can be found in Table 20. Across all samples, as with just the college sample, the correlations among the K-SRQ Admiration, Sexual Relationships, and Sociability scales are extremely large. Sensation Seeking shows slightly stronger correlations with K-SRQ Admiration, Sexual Relationships and Sociability scales than in the college sample alone. However, as with the college sample, above, the size of these correlations are similar to, or smaller than, the UPPS-P scale inter-questionnaire correlations.

Even though model fit was improved by constraining correlations to be the same across samples, examining the pattern of estimated within-sample correlations may help increase our confidence that this is reasonable. It is fairly clear in Figure 30 that, with some variability, the same high level of correlation among three of the four K-SRQ variables is present in all samples, though that between Admiration and Sexual Relationships may be somewhat attenuated for the in-lab college students (CSYA) and the community adolescents (CA). The relation between UPPS-P Sensation Seeking and K-SRQ Admiration, Sexual Relationships, and Sociability seems quite consistent. The most striking differences are tangential to the scope of this manuscript, but may be worth pointing out: the correlation between KSR-Q Passivity and both Sexual Relationships and Admiration is large and negative within the foster-care-involved adolescent sample (FCA), whereas for the other samples it is usually closer to zero (with community adolescents [CA] showing a large positive correlation between Passivity and Sexual Relationships).

Self report relation to SPLT. Four parameters from the reinforcement learning model of the SPLT behavior, and two summary statistics of optimal responding during the first and second half of the run, were correlated with all latent variables. The variables for the SPLT have a value for every condition (Hungry/Thirsty, Popular/Unpopular, Dating/Looking), as well as for the contrasts between the two social conditions and the minimally social condition. Including all SPLT variables in the combined measurement model caused problems with convergence that were not solved by changing the optimizer or adjusting optimization control options. For this reason, latent variable correlations were estimated for each SPLT-variable parameter subset separately.

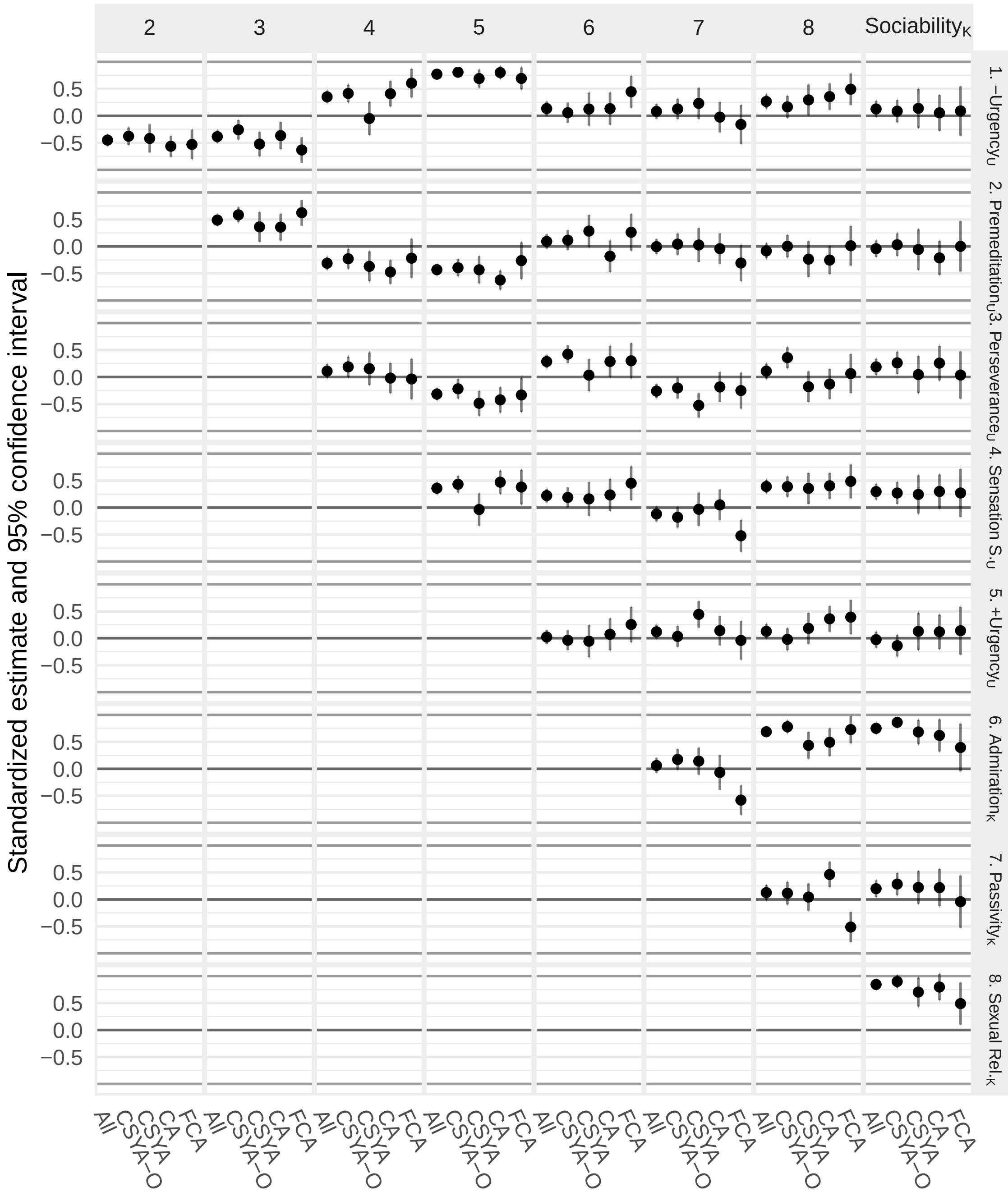

Figure 30. Latent variable correlations by subsample. The correlation for all subsamples is taken from the model with all loadings and latent (co)variances constrained to be equal across subsamples. Subsample correlations are taken from the model with partial metric invariance. U = Urgency, Premeditation, Perseverance, Sensation-seeking, & Positive Urgency scale; K = Kids' Social Reward Questionnaire; - = Negative; + = Positive. FCA: foster-care-involved adolescents; CA: community adolescents; CSYA: college students; CSYA-O: college students, online.

Table 20. Latent variable correlations, all samples

Latent variable	1	2	3	4	5
1. Negative Urgency^{U}		[-.55,-.34]	[-.50,-.27]	[.23, .47]	[.71, .83]
2. Premeditation^{U}	-.45		[.38, .59]	[-.43,-.19]	[-.54,-.33]
3. Perseverance^{U}	-.38	.49		[-.02, .24]	[-.43,-.20]
4. Sensation Seeking^{U}	.35	-.31	.11		[.25, .47]
5. Positive Urgency^{U}	.77	-.43	-.32	.36	
6. Admiration^{K}	.14	.09	.29	.22	.02
7. Passivity^{K}	.08	-.00	-.26	-.12	.12
8. Sexual Relationships^{K}	.27	-.08	.11	.39	.13
9. Sociability^{K}	.13	-.04	.19	.30	-.03

Latent variable	6	7	8	9
1. Negative Urgency^{U}	[.00, .27]	[-.05, .22]	[.14, .39]	[-.02, .28]
2. Premeditation^{U}	[-.04, .23]	[-.14, .13]	[-.22, .05]	[-.20, .11]
3. Perseverance^{U}	[.16, .41]	[-.39,-.13]	[-.03, .25]	[.04, .34]
4. Sensation Seeking^{U}	[.09, .35]	[-.26, .02]	[.27, .51]	[.15, .44]
5. Positive Urgency^{U}	[-.11, .15]	[-.01, .25]	[-.01, .26]	[-.18, .12]
6. Admiration^{K}		[-.07, .20]	[.60, .77]	[.64, .86]
7. Passivity^{K}	.06		[-.01, .26]	[.05, .35]
8. Sexual Relationships^{K}	.69	.13		[.75, .94]
9. Sociability^{K}	.75	.20	.84	

U = Urgency, Premeditation, Perseverance, Sensation-seeking, & Positive Urgency scale; K = Kids' Social Reward Questionnaire.
Correlations between latent variables appear in the lower triangle; 95% confidence intervals appear in the upper triangle. Source model constrains correlations and loadings for all items to be equal across gender groups.

College sample. Correlations were first examined in the college sample (N = 225) between the target SPLT parameters and the relevant subscales from the FSMI, K-SRQ, D&P, and UPPS-P questionnaires (adolescents were not administered the FSMI or D&P). All point estimates for these correlations were between r = [-0.16, 0.18], with most confidence intervals not exceeding r = [-0.20, 0.20] (Figure 31). There is also no pattern with regard to the contrast and the scale: for example, the association between FSMI Mate-seeking and either $\varepsilon_{\text{dating}}$ or the proportion of optimal responses dating contrasts (First or Last $1/2_{\text{dating}}$) is not noticeably stronger than the relation between these contrasts and FSMI Status. Moreover, though the correlation between FSMI Mate-seeking and the First $1/2_{\text{dating}}$ contrast stands out, it's in the wrong direction. This same absence of pattern is also apparent for the Popular/Unpopular - Hungry/Thirsty contrast.

Latent variable correlations with baseline (Hungry/Thirsty condition) performance also resulted in generally small estimates (r = [-0.23, 0.17]), though some sensible patterns emerge (Figure 32). Specifically, overall performance in the last half of the run shows a small positive correlation with D&P Prestige, and UPPS-P Sensation Seeking and Perseverance, and small negative correlations with D&P Dominance, and UPPS-P Positive and Negative Urgency. Examining correlations for the learning

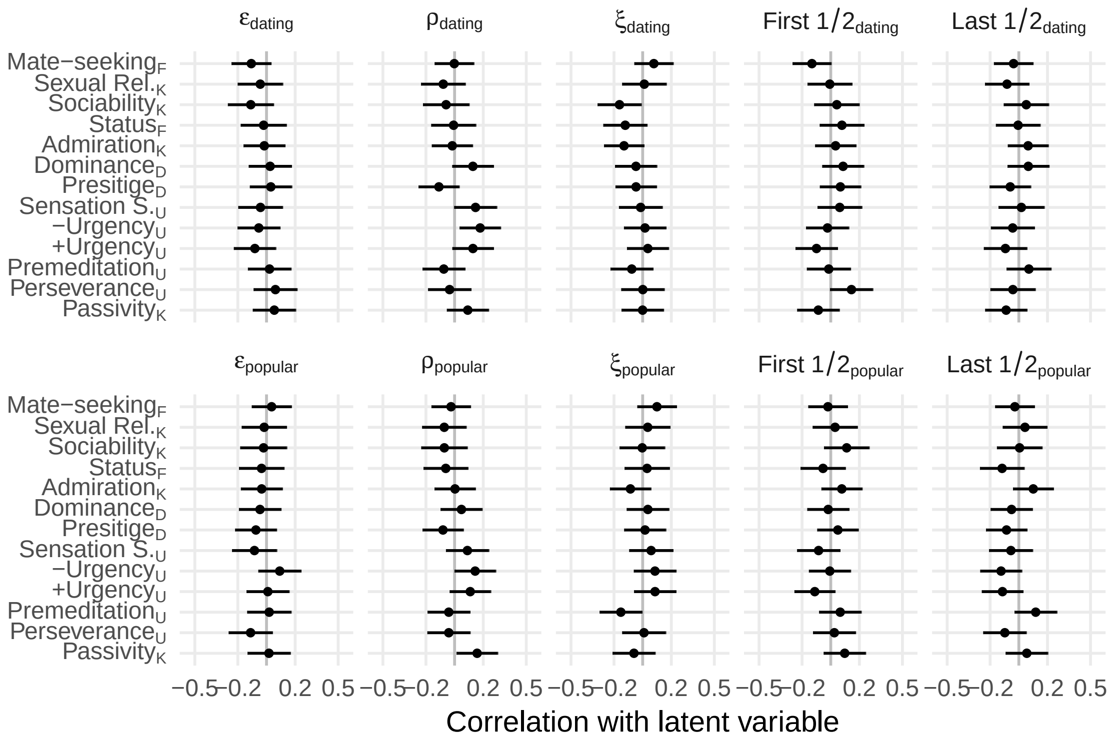

Figure 31. SPLT parameter contrast correlations with latent variables, college sample. Subscripts for each parameter indicate the contrast condition versus baseline (Hungry/Thirsty). “First 1/2” and “Last 1/2” refer to contrasts in proportions of optimal responses in the first and last half of the run, respectively. U = Urgency, Premeditation, Perseverance, Sensation-seeking, & Positive Urgency scale; K = Kids’ Social Reward Questionnaire; - = Negative; + = Positive.

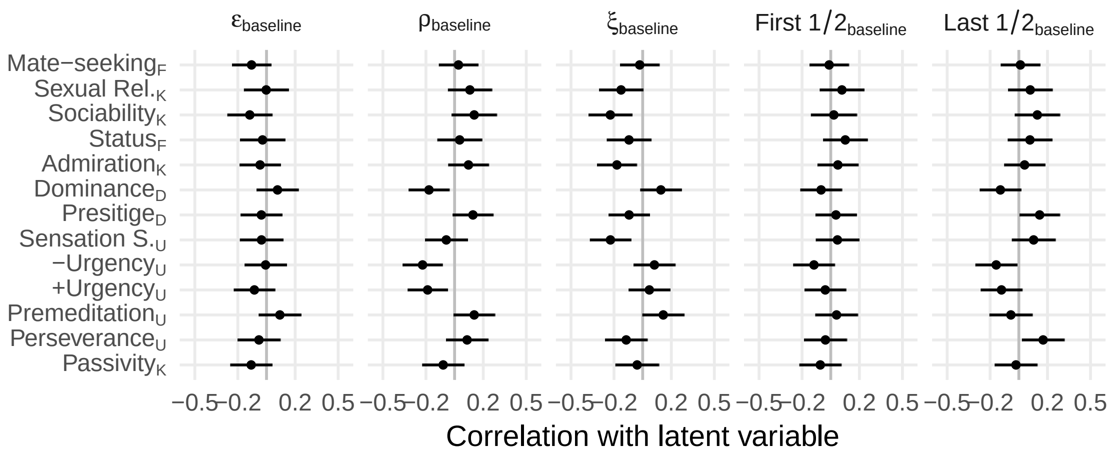

Figure 32. SPLT parameter baseline correlations with latent variables, college sample. "First 1/2" and "Last 1/2" refer to contrasts in proportions of optimal responses in the first and last half of the run, respectively. U = Urgency, Premeditation, Perseverance, Sensation-seeking, & Positive Urgency scale; K = Kids' Social Reward Questionnaire; - = Negative; + = Positive.

parameters, higher performance may be due to lower ξ for those higher in Sensation Seeking, higher ρ for those higher in Prestige, either or both lower ξ and higher ρ for those higher in Perseverance. Lower performance may be the result of either or both lower ρ and higher ξ for those high in Dominance, and lower ρ for those higher in Positive and Negative Urgency.

Combined college and adolescent samples. Next, using just the questionnaires administered to all four samples (N = 333, College students - online = 145, College students = 84, Community adolescents = 65, Foster-care involved adolescents = 39), correlations were estimated between the target SPLT parameters and the relevant subscales from the K-SRQ, and UPPS-P. The models that constrained correlations to be equal across samples fit better than those with correlations freed to differ between samples ($\Delta\text{AIC}_{\text{max}} = -47$; models for the *b* parameter did not converge). All point estimates for these correlations remained small in the combined sample (r = [-0.16, 0.18]), again, with most confidence intervals not exceeding r = [-0.20, 0.20] (Figure 33). None of the expected patterns appeared between (SPLT learning rate or performance) Dating contrasts and mate-seeking-relevant self-report scales, or Popular contrasts and status-relevant self-report. However, there was an interesting consistency in positive correlations between ρ for both contrasts and UPPS-P Sensation Seeking, and Positive and Negative Urgency, in line with the results from the college sample, above. Although one might be tempted to link these constructs to reward sensitivity, it is clear in the baseline results below that this likely arises because of *negative* correlation between these subscales and ρ_{baseline}.

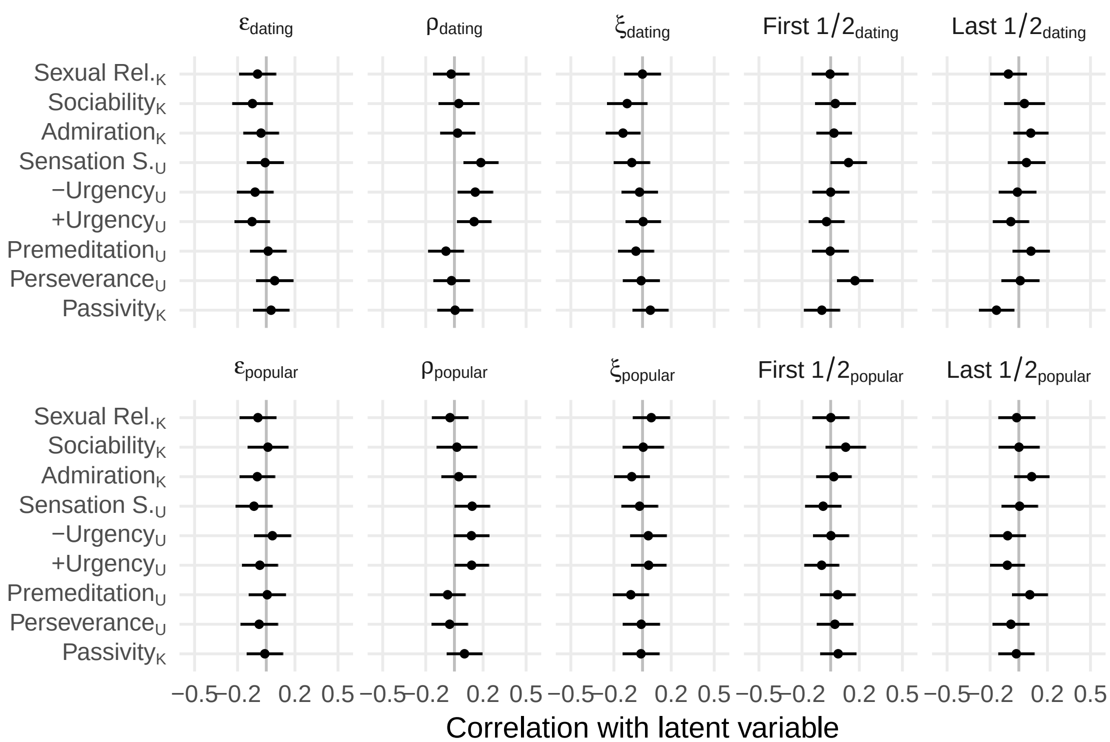

Figure 33. SPLT parameter contrast correlations with latent variables, all samples. Subscripts for each parameter indicate the contrast condition versus baseline (Hungry/Thirsty). "First 1/2" and "Last 1/2" refer to contrasts in proportions of optimal responses in the first and last half of the run, respectively. U = Urgency, Premeditation, Perseverance, Sensation-seeking, & Positive Urgency scale; K = Kids' Social Reward Questionnaire; - = Negative; + = Positive.

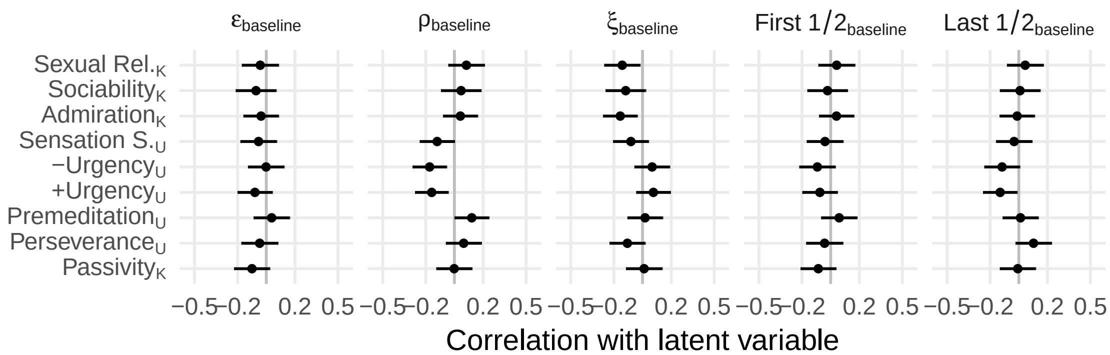

Figure 34. SPLT parameter baseline correlations with latent variables, all samples. "First 1/2" and "Last 1/2" refer to contrasts in proportions of optimal responses in the first and last half of the run, respectively. U = Urgency, Premeditation, Perseverance, Sensation-seeking, & Positive Urgency scale; K = Kids' Social Reward Questionnaire; - = Negative; + = Positive.

The latent variable correlations with baseline (Hungry/Thirsty condition) performance also resulted in generally small estimates (r = [-0.17, 0.12]). Both overall performance in the last half of the baseline condition, and the reward modifier, ρ, show negative correlations with UPPS-P Sensation Seeking, and Positive and Negative Urgency, as was the case in the college sample (Figure 34). The UPPS-P Perseverance variable continued to show somewhat attenuated correlations with ξ (negative) and performance (positive). Finally, though this does not seem to result in a concomitant correlation with performance, the three highly-correlated K-SRQ variables show small negative correlations with the noise parameter, ξ.

As a final note, it should be clear from the plots above that if one were to have set an *a priori* equivalence-testing threshold at $|r| = 0.20, \alpha = .05$, no correlation examined would be unambiguously different from the region of equivalence, and in many cases one would accept the null hypothesis that there is no correlation of any importance (Lakens, 2017).

Self-report variable associations with behavioral outcomes. In order to check the concurrent validity of self-report measures of motives (and impulsivity/sensation seeking), associations with self-report behavior (and relationship status) were assessed. Descriptive plots of responses to the following questions can be examined in the accompanying figures:

- "How many times, in the last six months, have you had sex/sexual relations," with sexual relations defines as "oral, vaginal or anal, and can be with male or female partners" (Figure 35)

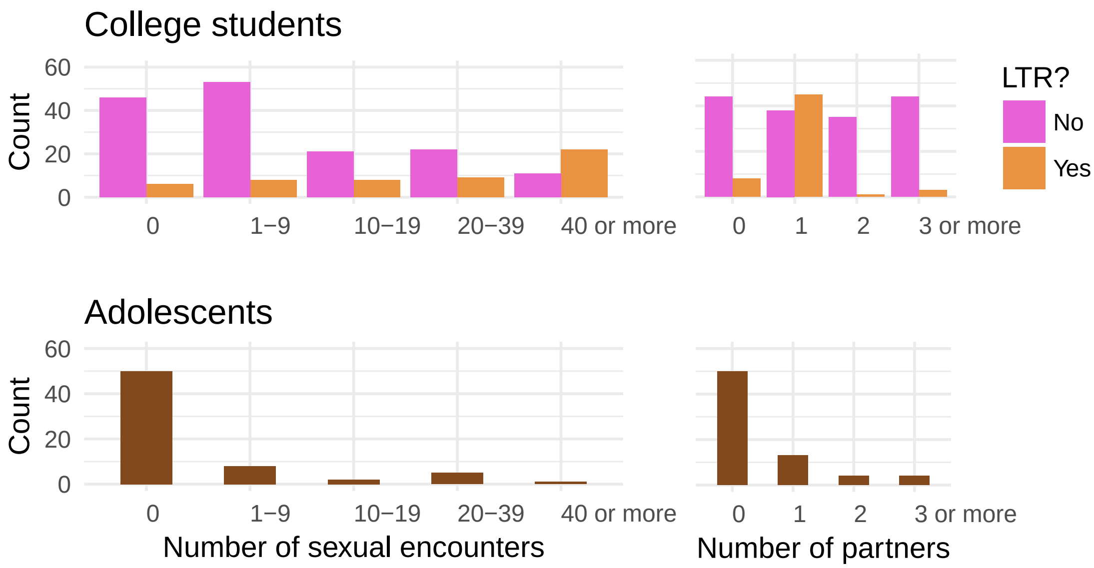

Figure 35. Frequency of sex and number of partners, last six months. Count is number of participants endorsing a particular response. LTR = Long term relationship, indicating that the participant reported they were in a such a relationship.

– "With how many different people in the last 6 months" (Figure 35)

– "In the last 6 months, what percentage of the time when you've had sexual relations (oral, vaginal or anal) did you USE safe sex practices to prevent sexually transmitted diseases - condoms, blood tests, dental dams, etc.)." (Figure 36)

– "In the past 30 days, on how many days did you have at least one drink of alcohol" (Figure 37)

– "In the past 30 days, on how many days did you have 5 or more drinks of alcohol in a row, that is, within a couple hours?" (Figure 37)

Overall, as would be expected, adolescent participants reported engaging in less sexual and drinking behavior than the college student participants. For those who were sexually active, the great majority also tended to report using safe sex practices. College participants in long term relationships (adolescent participants were not asked) tended to have only one partner, and also tended to report higher numbers of sexual encounters. Unpartnered college students were in the majority, tended to report a wider range of number of partners, and tended to more often endorse no, or lower numbers of sexual encounters. Safe sex practices show similar distributions for partnered and unpartnered college students, with heavier use of the poles of the scale. Most college students reported drinking at least one drink at least once in the

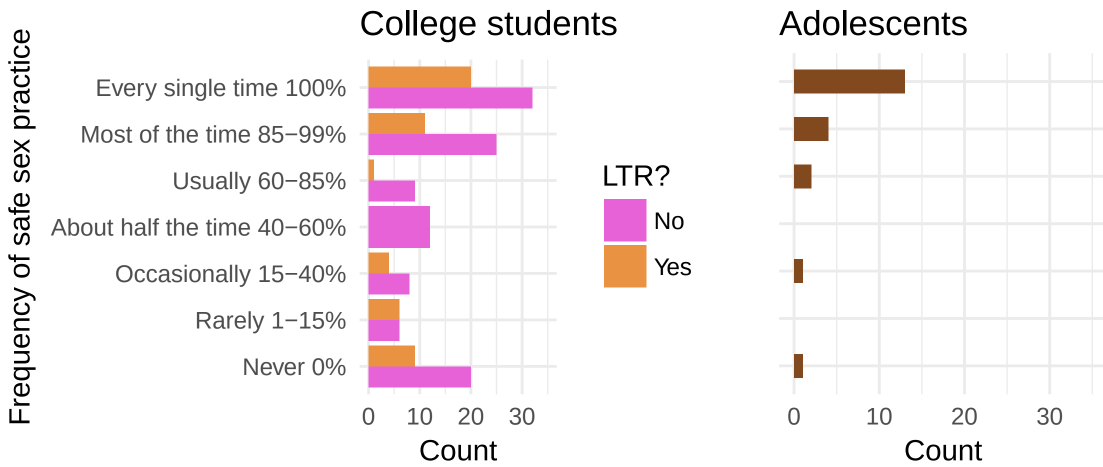

Figure 36. Frequency of safe sex practice, last six months. Count is number of participants endorsing a particular response. LTR = Long term relationship, indicating that the participant reported they were in a such a relationship. Overlapping percentage ranges appear in the original questionnaire.

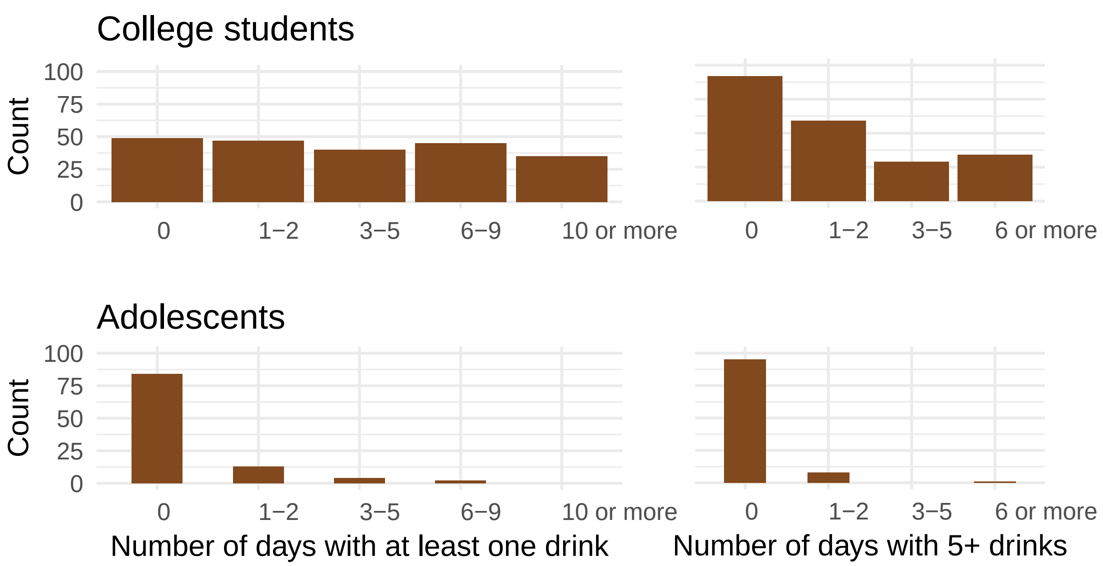

Figure 37. Frequency of drinking, last 30 days. Count is number of participants endorsing a particular response.

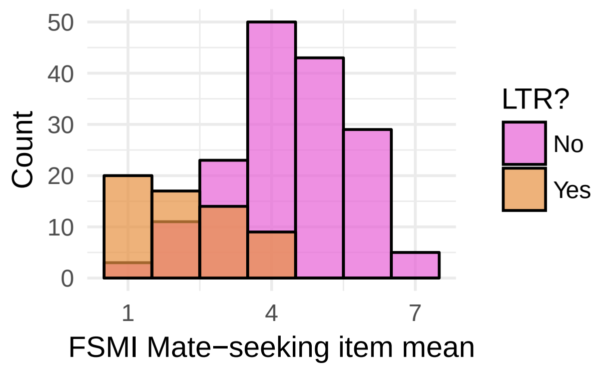

Figure 38. FSMI Mate-seeking scale and long-term relationship status. LTR = Long term relationship, indicating that the participant reported they were in a such a relationship. Values on X axis are binned means of six items.

last month, with a considerable number also reporting at least one heavy drinking day in the last month (five or more drinks in a session). Given the low variability on these questions in the adolescent sample, associations with all measures proceeds by first examining the relation in the college samples, and then adding in the adolescent samples to assess sensitivity.

Finally, the difference in the distribution of the FSMI Mate-seeking items for those college students in long-term relationships is markedly different from those who are not (Figure 38). The item content for the FSMI Mate-seeking scale asks primarily about intentions to find new romantic and sexual partners. As such (and given social norms favoring monogamy in committed relationships), it is encouraging evidence for the validity of this measurement to see that those who are in long-term relationships report very low means on this scale, while those who are unpartnered have computed scores that span the full range.

Number of partners. The analyses in the next several sections test whether sexual behavior is associated with the FSMI Mate-seeking, or K-SRQ Sexual Relationships variables. Initial models include both the focal motive variable and gender, which are compared to models in which the coefficient for the motive variable is fixed to zero. Significant decrease in fit is taken as evidence for an association between the motive variable and the behavioral outcome.

In the second step, the initial model is augmented by adding UPPS-P Sensation Seeking, Positive Urgency, and Negative Urgency, with the same constraint of the focal motive variable tested. Significant difference in fit here indicates that the focal motive is associated with the outcome even when conditioning on (or perhaps because of this conditioning) the UPPS-P impulsivity and sensation-seeking variables.

For the K-SRQ, an additional test is made by including the Admiration variable in the regression. Significant decrease in fit when the coefficient for this variable is fixed to zero indicates that it has an association with the behavioral outcome over and above the variance it shares with the Sexual Relationships variable. The sign of the coefficients between the unconstrained and constrained model will also be examined.

FSMI Mate-seeking

Constraining the coefficient of the FSMI Mate-seeking latent variable to zero in a model that regressed number of sexual partners on this variable and gender did not result in significantly worse fit when examining data from unpartnered college students ($N = 153$, $\Delta\chi^2[1] = 4.7$, $p = 0.030$). The coefficient for the effect of Mate-seeking on number of partners was, however, positive.

Next, we added to the above model the three UPPS-P variables (Sensation Seeking, Positive and Negative Urgency) and then compared this to a model that fixed the Mate-seeking coefficient to 0. This did not result in a significantly worse fit ($N = 122$, $\Delta\chi^2[1] = 1.3$, $p = 0.26$). The coefficient for the Mate-seeking variable, when included, was positive.

K-SRQ Sexual Relationships

In a model with the K-SRQ Sexual Relationships (SR) and gender predicting number of partners, constraining the coefficient for SR to zero resulted in significantly worse fit both in the unpartnered college student sample ($N = 136$, $\Delta\chi^2[1] = 11.2$, $p = 0.0008$), and in the the full (college and adolescent) sample (which also includes age as a covariate; $N = 253$, $\Delta\chi^2[1] = 24.1$, $p = 0.000001$). In both cases, the coefficient for SR is positive, indicating that those participants with higher levels on SR also report more sexual partners.

Comparing models with and without the K-SRQ Admiration variable included in the regression did not result in significantly worse fit in the unpartnered college sample ($N = 136$, $\Delta\chi^2[1] = 2.1$, $p = 0.14$), but did significantly reduce fit in the full samples ($N = 253$, $\Delta\chi^2[1] = 17.7$, $p = 0.00003$). In both cases, the coefficient for SR was positive and for Admiration was negative. This indicates that for a given level on SR, a participant with a higher level of Admiration would be expected to have fewer partners. Models that included K-SRQ Sociability did not converge, probably because of the very high correlation with the SR factor.

Removing SR from a regression that also included the three UPPS-P variables did not result in worse fit in the unpartnered college student sample ($N = 119$, $\Delta\chi^2[1] = 7.7$, $p = 0.006$), but did significantly reduce fit in the combined sample ($N = 225$, $\Delta\chi^2[1] = 22.9$, $p = 0.000002$). The coefficient for SR in these models was also positive.

Number of sexual encounters.

FSMI Mate-seeking

Comparing models with and without the FSMI Mate-seeking regression coefficient constrained to zero showed non-significant mis-fit in the unpartnered college sample ($N = 145$, $\Delta\chi^2[1] = 1.7$, $p = 0.19$). In the unconstrained model, the coefficient was negative and very close to zero. Adding UPPS-P covariates to the model, and again testing whether constraining the Mate-seeking coefficient to zero also resulted in non-significant reduction in fit ($N = 116$, $\Delta\chi^2[1] = 5.6$, $p = 0.02$), and in the unconstrained model the Mate-seeking coefficient remained negative.

K-SRQ Sexual Relationships

The test of K-SRQ SR showed non-significant reduction in fit when constraining its association with number of sexual encounters to zero in the unpartnered college sample ($N = 130$, $\Delta\chi^2[1] = 7.4$, $p = 0.007$), but significant reduction in fit in the full sample ($N = 239$, $\Delta\chi^2[1] = 20.5$, $p = 0.000006$). The coefficient for the SR variable in both cases was positive, indicating that those higher on SR tend to report having more sexual encounters (note that the coefficients for partnered and unpartnered college students were also both positive when considered separately).

Considering the models with and without the Admiration covariate showed non-significant difference in fit for the unpartnered college student sample ($N = 130$, $\Delta\chi^2[1] = 1.4$, $p = 0.24$), as well as the full samples ($N = 239$, $\Delta\chi^2[1] = 6.9$, $p = 0.009$). As was the case with number of partners, the coefficients for the Admiration variable were negative (while SR remained positive), indicating that, given the same level of SR, those higher in Admiration tended to report fewer sexual encounters.

Not accounting for the association between number of sexual encounters and SR in a model that also included the three UPPS-P variables did not result in significant difference in fit in the unpartnered college sample ($N = 113$, $\Delta\chi^2[1] = 1.9$, $p = 0.16$), but did result in a significant reduction in fit in the full sample ($N = 212$, $\Delta\chi^2[1] = 10.6$, $p = 0.001$). In both cases the coefficient for SR was positive.

Safe sex practices.

FSMI Mate-seeking

The fit difference tests for the effect of Mate-seeking on frequency of safe sex practices was not significant for either the unpartnered college sample ($N = 164$, $\Delta\chi^2[1] = 0.0$, $p = 0.97$) or for the full college sample ($N = 225$, $\Delta\chi^2[1] = 0.0$, $p = 0.90$). The coefficients in both cases were very close to zero. Including the three UPPS-P variables as covariates also resulted in no significant fit decrease when constraining the effect of Mate-seeking to be zero in the unpartnered ($N = 164$, $\Delta\chi^2[1] = 0.4$, $p = 0.51$) and combined college student sample ($N = 225$, $\Delta\chi^2[1] = 0.7$, $p = 0.40$).

K-SRQ Sexual Relationships

Neither the unpartnered college student sample ($N = 164$, $\Delta\chi^2[1] = 1.6$, $p = 0.20$), nor the full sample ($N = 333$, $\Delta\chi^2[1] = 1.3$, $p = 0.25$) showed significant decrease in fit when constraining the effect of the SR variable on frequency of safe sex practice to zero. The coefficient in both cases was positive, indicating that those higher on SR also tend to report more frequent practice of safe sex.

In a model that also included Admiration as a covariate, removing this variable from the regression did not reduce fit significantly for ether unpartnered college students ($N = 164$, $\Delta\chi^2[1] = 0.2$, $p = 0.63$) or the full sample ($N = 333$, $\Delta\chi^2[1] = 0.2$, $p = 0.66$). The coefficients were positive for both SR and Admiration in these models, though close to zero. Models with the three UPPS-P variables included demonstrated the same pattern for the unpartnered college students ($N = 164$, $\Delta\chi^2[1] = 0.2$, $p = 0.69$) and combined sample ($N = 333$, $\Delta\chi^2[1] = 0.6$, $p = 0.43$).

Number of drinking days. The following models test whether drinking behavior is associated with the FSMI Status, or K-SRQ Admiration variables. As above, initial models include both the focal motive variable and gender, which are compared to models in which the coefficient for the motive variable is fixed to zero. Significant decrease in fit is taken as evidence for an association between the motive variable and the behavioral outcome.

In the next step, the initial model is augmented by adding UPPS-P Sensation Seeking, Positive Urgency, and Negative Urgency, with the same constraint of the focal motive variable tested. Significant difference in fit here indicates that the focal motive is associated with the outcome even when conditioning on (or perhaps because of this conditioning) the UPPS-P impulsivity and sensation-seeking variables.

For the K-SRQ, an additional test is made by including the Sociability variable in the regression. Significant decrease in fit when the coefficient for this variable is fixed to zero indicates that it has an association with the behavioral outcome over and above the variance it shares with the Admiration variable. The coefficients between the unconstrained and constrained model will also be examined.

FSMI Status

In the college sample, constraining FSMI Status to have no association with number of drinking days did not significantly reduce model fit ($N = 207$, $\Delta\chi^2[1] = 0.3$, $p = 0.56$), with a coefficient for the effect, when it was freed, close to zero. This remained the case whether or not the UPPS-P variables were included in the model ($N = 169$, $\Delta\chi^2[1] = 0.5$, $p = 0.47$).

K-SRQ Admiration

Constraining the Admiration variable to zero in the regression of number of drinking days in the college sample resulted in a significant reduction in fit ($N = 185$, $\Delta\chi^2[1] = 13.0$, $p = 0.0003$). This was also the case for the full sample ($N = 279$, $\Delta\chi^2[1] = 15.3$, $p = 0.00009$). In both cases, the freed coefficient was positive, indicating that those with higher values on Admiration also reported more days where they had at least one drink (in the past 30 days).

The exclusion of the Sociability variable also led to significant decreases in fit for both the college sample ($N = 185$, $\Delta\chi^2[1] = 12.6$, $p = 0.0004$), as well as the full sample ($N = 279$, $\Delta\chi^2[1] = 21.1$,

$p = 0.000004$). In both cases, the freed coefficient for Sociability was positive, while the coefficient for Admiration became negative (with increased standard error). This indicates that participants reporting higher sociability also tend to report more drinking days, and that for a given level of Sociability, higher Admiration may be associated with fewer drinking days.

When tested in a model that included the three UPPS-P variables, constraining the Admiration effect on drinking days to zero did not reduce fit in the college sample ($N = 164$, $\Delta\chi^2[1] = 2.4$, $p = 0.12$) or the combined sample ($N = 247$, $\Delta\chi^2[1] = 1.7$, $p = 0.19$). The coefficient for the association between Admiration and drinking days remained positive in these models.

Number of heavy-drinking days.

FSMI Status

We did not reject the null of no association between FSMI Status and number of heavy drinking days in either the initial model ($N = 205$, $\Delta\chi^2[1] = 1.8$, $p = 0.18$) or the model including the UPPS-P variables ($N = 167$, $\Delta\chi^2[1] = 0.0$, $p = 0.93$). The coefficient of the association was positive in the first case, and close to zero (very slightly negative) in the UPPS-P model.

K-SRQ Admiration

The results for the the K-SRQ Admiration association with number of *heavy* drinking days were much the same as those for number of days having consumed any alcohol at all. There was a significant decrease in fit when the coefficient for Admiration was set to zero for both the college ($N = 182$, $\Delta\chi^2[1] = 15.0$, $p = 0.0001$) and full ($N = 277$, $\Delta\chi^2[1] = 16.4$, $p = 0.00005$) samples, with positive coefficients for the freed parameters. In models including the association with Sociability, significant fit differences were seen when fixing that association to zero for both the college ($N = 182$, $\Delta\chi^2[1] = 8.2$, $p = 0.004$) and full sample ($N = 277$, $\Delta\chi^2[1] = 20.6$, $p = 0.000006$). In both cases, freeing the Sociability coefficient resulted in a negative Admiration coefficient with smaller magnitude than the models excluding Sociability. In addition, in models including the UPPS-P variables, the coefficient for Admiration could be constrained to zero without significant fit difference in the college ($N = 162$, $\Delta\chi^2[1] = 2.4$, $p = 0.12$) and combined sample ($N = 245$, $\Delta\chi^2[1] = 1.8$, $p = 0.19$).

Table 21. SES and YRBS behavior reports regressed on SPLT variables

Behavior variable	SPLT variable	Coef sign$_{ltd}$	$\Delta\chi^2_{ltd}$	p_{ltd}	Coef sign$_{full}$	$\Delta\chi^2_{full}$	p_{full}
Number of partners	ε	$-$	5.0	.03	$-$	3.6	.06
SPLT contrast:	ρ	$+$	0.2	.69	$+$	0.9	.33
Dating/Looking	ξ	$+$	0.8	.37	$+$	1.5	.22
	b	$+$	0.2	.69	$+$	0.1	.80
	First 1/2 p_{opt}	$-$	2.5	.11	$-$	0.8	.38
	Last 1/2 p_{opt}	$-$	2.0	.16	$-$	4.9	.03
Number sexual encounters	ε	$-$	3.3	.07	$-$	0.6	.45
SPLT contrast:	ρ	$-$	1.3	.26	$+$	0.3	.58
Dating/Looking	ξ	$+$	0.1	.72	$+$	0.0	.90
	b	$+$	0.3	.56	$+$	0.2	.62
	First 1/2 p_{opt}	$-$	0.2	.68	$+$	2.0	.16
	Last 1/2 p_{opt}	$-$	0.9	.35	$-$	0.9	.33
Safe sex practice	ε	$+$	0.8	.36	$+$	1.2	.27
SPLT contrast:	ρ	$+$	1.5	.22	$+$	0.7	.40
Dating/Looking	ξ	$-$	0.2	.63	$-$	0.9	.35
	First 1/2 p_{opt}	$+$	0.4	.55	$+$	3.9	.05
	Last 1/2 p_{opt}	$+$	0.2	.66	$+$	3.4	.07
Days with 1+ drinks	ε	$-$	1.2	.27	$-$	0.1	.78
SPLT contrast:	ρ	$-$	0.7	.41	$-$	0.4	.54
Popular/Unpopular	ξ	$+$	0.2	.67	$+$	2.8	.09
	b	$+$	0.5	.50	$+$	0.7	.39
	First 1/2 p_{opt}	$-$	0.3	.60	$-$	1.2	.28
	Last 1/2 p_{opt}	$+$	0.7	.42	$-$	0.2	.69
Days with 5+ drinks	ε	$-$	6.4	.01	$-$	2.8	.10
SPLT contrast:	ρ	$-$	0.0	.92	$-$	0.0	.92
Popular/Unpopular	ξ	$+$	2.2	.14	$+$	4.0	.04
	b	$-$	0.4	.55	$-$	0.0	.89
	First 1/2 p_{opt}	$-$	0.4	.55	$-$	0.9	.33
	Last 1/2 p_{opt}	$-$	0.6	.44	$-$	2.0	.16

First and Last 1/2 p_{opt} refer to the proportion of optimal responses made during the first and last half of the SPLT run, respectively. SPLT variable contrasts are between the indicated condition and the minimally social Hungry/Thirsty condition. The "ltd" and "full" subscripts refer to the samples used in model fitting. For the sexual behavior questions, the limited (ltd) sample includes just unpartnered participants from the two college samples, and for the alcohol use questions, the ltd sample includes all participants from the two college samples. The full sample includes both college and both adolescent samples. "Coef sign" indicates the sign of the coefficient of the association between the behavior variable and SPLT contrast variable.

SPLT relation to outcomes. Some models regressing the frequency of safe sex practices on the SPLT parameter b did not converge, so comparisons involving these models were not tested. Since this parameter was not the focus of the investigation, these convergence issues were not resolved. No model comparisons met the $\alpha = .005$ threshold, and the direction of coefficients were often not in the expected direction – for example, the association between number of partners and the ε Dating/Looking contrast is negative (Table 21).

Differences in model performance between partnered and unpartnered college students were tested for each of the proportion of optimal response variables, as well as for each of the model parameters for the Dating/Looking - Hungry/Thirsty contrasts. Two-sample Yuen t-tests (computed using the PairedData package; Champely, 2017), which are robust to non-normality (Yuen, 1974), were used for these comparisons. No significant differences between these two groups were observed for any of the six tested variables.

Discussion, Aim 2

In these analyses, we observed very little evidence that behavior on the Social Probabilistic Learning Task (SPLT) is associated with either self-reported traits or health-related behavior. There was relatively adequate evidence that the self-report trait scales successfully measured the relevant constructs, eliminating this as a possible explanation for the absence of strong associations with SPLT behavior. Some measurement issues, specifically poor measurement-model fit and high correlations among K-SRQ scales, qualify these observations.

Self-report scale measurement and associations. All confirmatory factor analyses of the self-report measurement models showed questionable fit, with the exception of the FSMI Mate-seeking and Status model. Poor fit may be due to incorrect specification of the relation of the items to latent constructs, such as assuming the wrong number of factors, or an incorrect latent variable hierarchy. Factor structure was not tested as part of these analyses, but there is some evidence in the correlations between latent variables of considerable overlap that may warrant consideration of alternative structures. For example, Negative and Positive Urgency in the UPPS-P may not be separable factors, or may be lower-order factors of a higher-order impulsivity variable (Cyders, Littlefield, Coffey, & Karyadi, 2014). Poor fit and high subscale correlations for the K-SRQ scales, which only has forward keyed items, may also simply be the result of acquiescence bias. Some support for acquiescence bias was seen in the relatively high amount of shared variance between responses to two attention-check items and the K-SRQ subscales. Extreme responding driven either by social desirability or miscalibrated item content (e.g., almost no one dislikes being looked up to, or kissing) is also a concern with this scale, which shows skew in responses to almost every item. Indeed, two subscales were not included in these analyses because extreme responding in one direction was so pronounced. When measurement non-invariance was observed, sensitivity tests indicated without exception that substantive results were not sensitive to them.

Despite possible problems with measurement, correlations between latent variables, associations with age and puberty, and associations with self-report behavior, showed patterns that may increase our confidence in the validity of these self report instruments. Motivation to seek out new romantic and sexual partners, and enjoyment of romantic and sexual behaviors, should have a fairly direct influence on self-reported behaviors like number of sexual partners and frequency of sexual behavior. The number of sexual partners was consistently significantly associated with the K-SRQ Sexual Relationship measure of socio-sexual reward in both the set of college participants not in a long-term relationship (unpartnered) and in the full sample. This association was somewhat attenuated when including the UPPS-P variables in the unpartnered college student sample. Consistent but less robust positive associations were also seen between the FSMI Mate-seeking measure and number of sexual partners.

The K-SRQ Sexual Relationships measure also showed consistent positive associations with frequency of sexual encounters, whether or not UPPS-P variables were included, but significant reduction in model fit was only observed when looking at the full sample. Interestingly, the association between number of recent sexual encounters and mate-seeking, was close to zero or even slightly negative.

Neither the FSMI Mate-seeking nor K-SRQ Sexual Relationships scale showed a robust association with frequency of safe sex practice, regardless of the sample analyzed. The K-SRQ coefficients were positive, possibly indicating a slight tendency for those who seek social reward from sexual and romantic experiences to also be more well informed about best practices for reducing risk. The lack of association between Mate-seeking and safe sex practice seems to be inconsistent with findings during scale development that, among single individuals, Mate-seeking was positively correlated with condom use (Neel et al., 2015). However, this finding was also relatively weak, and was more focused on one particular aspect of safe sex practice, whereas the current investigation used a diffuse measure that incorporates more subjective information about a wider range of safe sex practice.

Several other findings also support the validity of the measures of mate-seeking motive. Helpfully for the validation of FSMI Mate-seeking, and as was first reported by Neel et al. (2015), this measure was directly impacted by current relationship status, with those in long-term-relationships scoring much lower. The K-SRQ Sexual Relationships scale also shows an expected positive association with both age and puberty, though this is somewhat more pronounced in the subset of female participants compared to male participants. The two sexual motivation subscales also show a positive correlation, though it is somewhat small. Overall, this pattern of associations between the measures of motives, self-report sexual behavior, and relationship status is consistent with our expectations of these constructs, and helps support their validity.

However, many of the correlations between K-SRQ Sexual Relationships and other measures are rather high in unexpected ways, possibly indicating less clarity in its interpretation. For example, it correlates more strongly with FSMI Status, and Prestige strategy than with FSMI Mate-seeking. As mentioned above, it correlates extremely highly with both K-SRQ Admiration and K-SRQ Sociability.

The several scales measuring aspects of status motivation, FSMI Status, Prestige, and K-SRQ Admiration all showed strong correlations. Dominance strategy correlated strongly with FSMI Status but not with either Prestige or K-SRQ Admiration. Given that dominance is a distinct strategy from prestige, the low correlation between these variables simply indicates that one might adopt either or both strategies in pursuit of achievement of social rank, consistent with the original report of Cheng et al. (2013). The high correlation of Prestige with K-SRQ Admiration (and high correlation of both with FSMI Status) indicates that the Admiration scale is perhaps primarily measuring individual differences in enjoyment of receipt of prestige-based status, in particular. This is also suggested both by the scale names and overlap in item content (e.g., Prestige: "I am held in high esteem by those I know"; Admiration: "I like achieving recognition from others"). The modest correlation between Admiration and UPPS-P Perseverance (a measure of prevailing when tasks are hard or boring) is consistent with this interpretation that Admiration is capturing prestige-based status motivation.

Status motives as measured in the college student sample using the FSMI Status scale did not show significant associations with number of drinking days or number of heavy drinking days, and did not show any consistent pattern in the sign of the coefficients. Admiration was significantly positively associated with both number of drinking days and heavy drinking days. However, Sociability also showed a significant relation even when controlling for Admiration, which showed a *negative* association with both drinking outcomes when Sociability was included. This is actually somewhat consistent with the investigation by Allen et al. (2005), which found that during adolescence, popularity was both associated with well-adjusted development but also increased exposure to peer groups where minor deviant behavior was socialized. Though alcohol consumption is almost certainly considered less (or not at all) deviant among the college participants, a similar principle may apply here. It is possible that sensitivity to admiration from peers is related to alcohol use just because enjoying admiration, to some extent, requires enjoying socializing as well. When taking into account motivation to socialize, specifically, concerns about receiving admiration from peers seem to be associated with more cautious consumption.

To use these self-report questionnaires to test whether individual differences in SPLT behavior arise from individual differences in related motives, it is essential that the self-report measures be valid instruments for those motives. It must also be the case that the way those motives are elicited in the task, specifically by presenting information about dating-pool availability or social status, is related to the way the constructs are conceptualized on the item level. In other words, does it make sense that someone who scores high on either the FSMI Mate-seeking or K-SRQ Sexual Relationships scale would care about whether another person was currently in a relationship or looking for someone to date? Would someone who is seeking social status, or who enjoys receiving recognition from others care about whether a peer is popular or unpopular? Of course, it is difficult to say for certain, but in the case of the mate-seeking motive questionnaires, the answer seems to be obviously yes.

With regard to the status-motive questionnaires, the answer is less clear, and may depend on whether a person's strategy for achieving status requires knowledge of their peers' social status. Just as there are at least two strategies for gaining status, there are also seemingly two sub-types of popular peers – those who are popular but not necessarily well-liked, and those who are popular *because* they are well-liked (de Bruyn & Cillessen, 2006). While there is some evidence that being friends with popular peers, per se, confers status (Meuwese et al., 2017), it is not clear whether this depends on the strategies of these particular peers. Based on the indications of construct validity reviewed above, and a plausible case for the relevance of the task manipulations to status and mate-seeking motives, these self-report scales are likely to function as adequate criteria for the relevance of individual differences in task behavior to motivation in these domains.

SPLT variable associations. Correlations between the self-report latent variables and SPLT variables were all very low. While the contrasted-SPLT variable correlation estimates did not show readily interpretable patterns, there was some indication that correlations with the baseline (Hungry/Thirsty condition) variables did show a sensible pattern in their sign. This provides some evidence that individual differences in learning behavior, overall, are associated with self report trait differences, but that differences between learning conditions are either not related to individuals' trait differences, or that the task measurement is not sensitive enough to detect these relations. Though the largest associations are discussed for the sake of completeness, note well that these effects are small, or very small, and are just a selection from a large number of estimates.

In the combined sample, baseline learning behavior seemed to be weakly associated with several subscales of the UPPS-P. Impulsivity, as measured by the Urgency scales, showed negative relations with performance during the last half of the run, while the Perseverance scale, a measure of ability to stay focused, was positively associated with performance. Interestingly, these same individual differences are also weakly associated with different model parameters, suggesting that if there is a relation between these traits and performance, they may arise through different psychological mechanisms. Specifically, poorer performance for those higher in Urgency may be due to lower sensitivity to any particular piece of feedback, as indexed by the parameter ρ. Perseverance, on the other hand, is also weakly positively associated with lower ξ, indicating that greater ability to focus may reduce the amount of random (or erroneous) responding, thereby increasing performance. Though these effects are very weak, and possibly false-positives, their sign and association with both performance and particular model parameters is coherent enough to be suggestive.

When the college sample was isolated, we observed similar patterns with the addition of both weak negative associations between performance and Dominance strategy, and weak positive associations between performance, and both Prestige and sensation seeking. Prestige and Dominance associations with performance seemed to be the result of opposite patterns of, respectively, higher and lower sensitivity to reward (ρ), and lower and higher random responding (ξ). On the other hand, the relations of Sensation Seeking to performance may be explained primarily by its negative correlation with ξ. The Prestige correlation is consistent with the view that such a strategy requires individuals to demonstrate competency, and though rather contrived, this task is an opportunity to perform well. Dominance-related popularity strategies have been associated with relational aggression and academic disengagement (de Bruyn & Cillessen, 2006; Kiefer & Ryan, 2008), which could explain this negative relation to performance, possibly through decreased focus. If there is a true relation between task performance and Sensation Seeking, it is much less straightforward – while some authors have suggested that purported heightened reward sensitivity during adolescence could lead to improved learning (Davidow et al., 2016; McCormick & Telzer, 2017), Sensation Seeking in these data correlate more strongly with the noise parameter, ξ, than with reward sensitivity, ρ. Generally, these patterns of correlations with baseline task behavior highlight one of the potential strengths of this method (which would be more easily verified with higher precision), which is that it allows decomposition of behavior into distinct psychological processes that may be difficult to disentangle using other means (van den Bos, Bruckner, Nassar, Mata, & Eppinger, 2017).

Although correlations between SPLT contrast variables and self-report scales are of a similar magnitude to the baseline correlations, they have a much less straightforward interpretation. For example, in the combined samples, and college-only samples, higher sensation seeking and impulsivity as measured by UPPS-P subscales are weakly positively correlated with the ρ parameter differences between both the Dating and Popular conditions versus baseline, but much more weakly, still, with overall performance. Only when the combined sample is considered does Passivity (which measures enjoyment of being told what to do by others) correlate negatively with final performance differences in the Dating condition, while Perseverance correlates positively with early performance in the dating condition. These results do not seem to be consistent with the expectations of this investigation or with literature about performance effects within these specific domains. Perhaps they may be of use to a future meta-researcher.

Finally, the SPLT variables also showed no coherent associations to any of the self-reported behavioral outcomes. Tests of group differences in learning in the Dating/Looking condition between partnered and unpartnered participants were unfortunately restricted to those in the college sample. It is possible that differences between college students and pre-college adolescents in the quality or meaning of long-term relationships would alter how they affect mate-seeking motives; this question must be left for future research. These analyses were included here primarily to further validate task behavior as a measure of social motives in the case that SPLT behavior and self-report scales showed correspondence, and to test whether SPLT behavior might be a *better* predictor of health-related behaviors than self-report trait measures. In the context of the results discussed above, the lack of significant associations between SPLT variables and behavioral outcomes serves as another indication that the task either does not capture individual differences in relevant social motives, or does so only very imprecisely.

Conclusions. Individual differences in learning performance and model-estimated variables derived from the SPLT do not clearly measure individual differences in social motives, nor other potentially related constructs. Self-report trait measures, on the other hand, were observed to relate coherently with each other, and with selected health-related behaviors, though there were also some indications that they could be further refined to hone their effectiveness.

CHAPTER V

DISCUSSION

The goals of this investigation were to examine how reinforcement learning is affected by adolescent-relevant social motives, and to look for evidence that individual differences in this effect are due to individual tendencies in the strength of these motives. It is clear that there were salience effects of the social-motive framing of stimuli that led to enhanced learning relative to the minimally social condition. However, it is also clear that individual differences in this effect are not strongly influenced by individual differences in social motives. Age, puberty, self-report trait, and health-related behaviors all showed negligibly small associations with individual differences in the social-motive effects on task performance and model parameters. This presents an interpretive conundrum.

One possibility is that individual differences in social motives do not result in differences in the salience of stimuli that are relevant to those motives. The rich history of work on attention-reward contingencies elicited by the interaction of a person with her environment makes this possibility seem very unlikely (Mackintosh, 1975; Grossberg, 1975; Kruschke, 2011). A more likely possibility, explored further below in the section on limitations, is that the stimuli and framing used in this task were not actually relevant, in the participants' perception, to the social motives we sought to explore. Instead, salience differences may be the result of incidental differences in the descriptor content along dimensions that are robust to the artificial laboratory setting.

Another, perhaps equally plausible, possibility is that information about dating and popularity is *too* relevant to almost everyone, regardless of their personal level of motivation. Perhaps it is the case that from even early adolescence, people understand that it is important to attend to and learn from possible signals in the social world about dating and popularity, regardless of one's own motives in relevant domains. In the domain of social status, this is in line with work that suggests status misperception can incur a social cost regardless of one's position (Srivastava & Anderson, 2011). Learning who is popular may be such a basic skill in forming accurate-enough impressions of social hierarchies that there is not room for status motives to affect learning. This same sort of ceiling effect is just as applicable to the mate-seeking condition.

Limitations

The possible explanations for null results discussed above have to do with substantive relations between motives and learning. There are also explanations for null results that have more to do with instrumentation and procedure than with anything theoretically very meaningful. For example, participants may accurately perceive that learning the face-word associations on this task is actually irrelevant to fulfilling their social motives. Learning might be enhanced in the social conditions just because it is slightly more interesting to imagine these computer generated faces being popular, or dating, than it is to imagine that they are hungry. The participants' own motives may never enter into the equation because it is obvious that whatever is learned during the task will not be instrumental in meeting new romantic partners or gaining the admiration of their peers.

The reliable variance in task behavior and model-derived variables may have been somewhat low, and thereby attenuated the observable correlation with the criteria measures. This is a problem more generally with measuring latent variables using only binary responses, even if accounting for task structure with the computational model does help somewhat. Compounding this problem is the possibility that the method itself is responsible for some proportion of the variance measured by both the task, and self-report scales. Correspondence between task-behavior and self-report measures of the same ostensible construct has been shown to be rather low across a wide range of commonly used adolescent risk-taking measures (Harden et al., 2017). Whether this is due to method variance, or problems with the construct validity of the tasks examined in that investigation, the SPLT task may be another example of this problem.

Some measurement problems, including skewed responding, extreme latent variable covariance, and unexpected patterns of latent variable correlation also may limit the degree of certainty in the construct validity of the self-report measures. This is not to suggest that these instruments are not responsive to some true signal, as is suggested by their association to self-reported health-related behaviors. Rather, these issues suggest that further work could be done to refine the conceptual space, and thereby produce more accurate instruments that measure more well-defined constructs that are structurally better situated in a more valid taxonomy. Such refinement might both raise the ceiling on the magnitude of observable relations, and also allow more straightforward interpretation of these relations.

There are also potential issues with generalizing from the effect of the present experimental manipulations to the population of possible manipulations related to the motivational constructs of interest. The results from the SPLT can really only tell us about people's responses to the particular manipulations that were used in this experiment. In order to infer back to the population of all status- and mate-seeking-related manipulations, we would need to account for the variance in responding to those various manipulations, either by estimating that variance from a small sample of those possible manipulations, or by incorporating a very large number of them into the design (Westfall, Nichols, & Yarkoni, 2017). Being able to generalize to this population of manipulations is also an important step in generalizing the effect of *a particular* status- or mate-seeking-relevant stimulus-manipulation to the effect of *motive-relevance* more generally. Indeed, given that only one manipulation for each motive domain was used, it is not even possible to estimate whether these represent very good, or very bad manipulations of motive-relevance.

Future directions

Although this task is unlikely to be a robust measure of individual differences in motives, the technique of manipulating the framing of stimuli does show promise for future work that would use it to estimate group-averaged, within-person differences in how particular stimulus characteristics affect learning. Future work could expand on this paradigm by expanding the content to test broader structural hypotheses regarding what kinds of social information is most salient. There are many other ways to expand on reinforcement learning tasks like the SPLT, as well. Since environmental stability is another factor that may influence the expression of these motives (Brumbach, Figueredo, & Ellis, 2009), one might manipulate the stability of the learning environment in different domains to examine advantageous or disadvantageous propensities to change one's mind about previously learned social information.

Allowing participants to have more control over the type of content that is learned by allowing both exploration and exploitation during the learning task may be a way to increase the diversity of stimuli, as well as examine how structural relationships between pieces of information influences what kinds of social information participants choose to learn about. At least one study, discussed in the previous chapter, took a step in this direction by allowing participants to learn about specific preferences, though there was not much attention given to how this set up could be more fully exploited (Rosenblau et al., 2017). One possibility would be to use information that is instrumental to goal pursuit as the reward feedback, rather than as content to be learned.

Examining different models for learning may also enhance future work. Reinforcement learning models like the one used in these analyses have a long history, and are very well understood. This model does, of course, make certain simplifying assumptions about latent cognitive process. However, without exploring multiple different models that could possibly characterize the learning process, it is not possible to assess which model features substantive results are sensitive to. Though this study does examine the influence of a few model features (i.e., inclusion of bias and reward modifier parameters), there are entire classes of learning models (e.g., instance based learning; Gonzalez et al., 2003; and model-based learning; Decker et al., 2016) that are not tested, or are not possible to test, in these data.

Conclusions

Reinforcement learning is enhanced by the social-motive framing manipulation, though individual difference correlations leave it ambiguous as to whether this is due to the words' relevance to social motives, or some other property. Generally, older participants performed more optimally during the learning task, which was reflected in better focus (smaller ξ), and possibly higher sensitivity to reward (higher ρ). Self-report trait questionnaires that measure specific motivations were sometimes helpful for accounting for health-related behaviors, and with regard to normative sexual behavior, this was the case even over and above more general measures of impulsivity and sensation seeking. Important future directions include further elaboration of learning paradigms, both to increase precision and expand content coverage, and refinement of motivational taxonomies and measurement instruments.

REFERENCES CITED

Ahn, W.-Y., Haines, N., & Zhang, L. (2017). Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package. *Computational Psychiatry*, *1*, 24–57. doi:10.1162/CPSY_a_00002

Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. In *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer Series in Statistics. Springer, New York, NY. doi:10.1007/978-1-4612-1694-0_15

Allen, J. P., Porter, M. R., McFarland, F. C., Marsh, P., & McElhaney, K. B. (2005). The Two Faces Of Adolescents' Success With Peers: Adolescent Popularity, Social Adaptation, and Deviant Behavior. *Child development*, *76*(3), 747–760. doi:10.1111/j.1467-8624.2005.00875.x

Anduiza, E. & Galais, C. (2016). Answering Without Reading: IMCs and Strong Satisficing in Online Surveys. *International Journal of Public Opinion Research*, edw007. doi:10.1093/ijpor/edw007

Bakan, D. (1966). *The duality of human existence: An essay on psychology and religion.* Chicago: Rand McNally.

Baltes, P. B. (1997). On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as foundation of developmental theory. *American Psychologist*, *52*(4), 366–380. doi:10.1037/0003-066x.52.4.366

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01

Blakemore, S.-J. & Mills, K. L. (2014). Is Adolescence a Sensitive Period for Sociocultural Processing? *Annual Review of Psychology*, *65*(1), 187–207. doi:10.1146/annurev-psych-010213-115202

Borsboom, D. (2006). When Does Measurement Invariance Matter? *Medical Care*, *44*, S176–S181. doi:10.1097/01.mlr.0000245143.08679.cc

van den Bos, W., Bruckner, R., Nassar, M. R., Mata, R., & Eppinger, B. (2017). Computational neuroscience across the lifespan: Promises and pitfalls. *Developmental Cognitive Neuroscience*. doi:10.1016/j.dcn.2017.09.008

van den Bos, W., Güroğlu, B., Van Den Bulk, B. G., Rombouts, S. A., & Crone, E. A. (2009). Better than expected or as bad as you thought? The neurocognitive development of probabilistic feedback processing. *Frontiers in human neuroscience*, *3*, 52. doi:10.3389/neuro.09.052.2009

de Bruin, W. B., Parker, A. M., & Fischhoff, B. (2007). Can Adolescents Predict Significant Life Events? *Journal of Adolescent Health*, *41*(2), 208–210. doi:10.1016/j.jadohealth.2007.03.014

Brumbach, B. H., Figueredo, A. J., & Ellis, B. J. (2009). Effects of Harsh and Unpredictable Environments in Adolescence on Development of Life History Strategies. *Human nature (Hawthorne, N.Y.) 20*(1), 25–51. doi:10.1007/s12110-009-9059-3

de Bruyn, E. H. & Cillessen, A. H. N. (2006). Popularity in Early Adolescence: Prosocial and Antisocial Subtypes. *Journal of Adolescent Research*, *21*(6), 607–627. doi:10.1177/0743558406293966

Buzwell, S. & Rosenthal, D. (1996). Constructing a Sexual Self: Adolescents' Sexual Self-Perceptions and Sexual Risk-Taking. *Journal of Research on Adolescence*, *6*(4), 489–513.

Cairns, R. B., Leung, M.-C., Buchanan, L., & Cairns, B. D. (1995). Friendships and Social Networks in Childhood and Adolescence: Fluidity, Reliability, and Interrelations. *Child Development*, *66*(5), 1330–1345. doi:10.1111/j.1467-8624.1995.tb00938.x

Capaldi, D. M., Stoolmiller, M., Clark, S., & Owen, L. D. (2002). Heterosexual risk behaviors in at-risk young men from early adolescence to young adulthood: Prevalence, prediction, and association with STD contraction. *Developmental Psychology*, *38*(3), 394–406. doi:10.1037/0012-1649.38.3.394

Caravita, S. C. S. & Cillessen, A. H. N. (2012). Agentic or Communal? Associations between Interpersonal Goals, Popularity, and Bullying in Middle Childhood and Early Adolescence: Goals, Popularity, and Bullying. *Social Development*, *21*(2), 376–395. doi:10.1111/j.1467-9507.2011.00632.x

Carlson, D. L., McNulty, T. L., Bellair, P. E., & Watts, S. (2014). Neighborhoods and Racial/Ethnic Disparities in Adolescent Sexual Risk Behavior. *Journal of Youth and Adolescence*, *43*(9), 1536–1549. doi:10.1007/s10964-013-0052-0

Casey, B. J., Galván, A., & Somerville, L. H. (2016). Beyond simple models of adolescence to an integrated circuit-based account: A commentary. *Developmental Cognitive Neuroscience*, *17*, 128–130. doi:10.1016/j.dcn.2015.12.006

Cavazos-Rehg, P. A., Krauss, M. J., Spitznagel, E. L., Schootman, M., Bucholz, K. K., Peipert, J. F., … Bierut, L. J. (2009). Age of sexual debut among US adolescents. *Contraception*, *80*(2), 158–162. doi:10.1016/j.contraception.2009.02.014

Centers for Disease Control and Prevention. (2003). Web-based Injury Statistics Query and Reporting System (WISQARS). Retrieved March 17, 2016, from www.cdc.gov/ncipc/wisqars

Centers for Disease Control and Prevention. (2015). *Youth Risk Behavior Survey Questionnaire*. Retrieved from www.cdc.gov/yrbs

Champely, S. (2017). *PairedData: Paired Data Analysis*. Retrieved from https://CRAN.R-project.org/package=PairedData

Chassin, L., Hussong, A., & Beltran, I. (2009). Adolescent Substance Use. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of Adolescent Psychology* (3rd, pp. 723–763). Hoboken, NJ: John Wiley & Sons Inc.

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, *104*(1), 103–125. doi:10.1037/a0030398

Christakou, A., Gershman, S. J., Niv, Y., Simmons, A., Brammer, M., & Rubia, K. (2013). Neural and Psychological Maturation of Decision-making in Adolescence and Young Adulthood. *Journal of Cognitive Neuroscience*, *25*(11), 1807–1823. doi:10.1162/jocn_a_00447

Cillessen, A. H. N. & Borch, C. (2006). Developmental trajectories of adolescent popularity: A growth curve modelling analysis. *Journal of Adolescence*, *29*(6), 935–959. doi:10.1016/j.adolescence.2006.05.005

Cillessen, A. H. N., Mayeux, L., Ha, T., de Bruyn, E. H., & LaFontana, K. M. (2014). Aggressive effects of prioritizing popularity in early adolescence: Prioritizing Popularity. *Aggressive Behavior*, *40*(3), 204–213. doi:10.1002/ab.21518

Cohen, J. R., Asarnow, R. F., Sabb, F. W., Bilder, R. M., Bookheimer, S. Y., Knowlton, B. J., & Poldrack, R. A. (2010). A unique adolescent response to reward prediction errors. *Nature Neuroscience*, *13*(6), 669–671. doi:10.1038/nn.2558

Collado, A., Felton, J. W., MacPherson, L., & Lejuez, C. (2014). Longitudinal trajectories of sensation seeking, risk taking propensity, and impulsivity across early to middle adolescence. *Addictive Behaviors*, *39*(11), 1580–1588. doi:10.1016/j.addbeh.2014.01.024

Collins, A. G. E. & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis: Working memory in reinforcement learning. *European Journal of Neuroscience*, *35*(7), 1024–1035. doi:10.1111/j.1460-9568.2011.07980.x

Collins, W. A. (2003). More than Myth: The Developmental Significance of Romantic Relationships During Adolescence. *Journal of Research on Adolescence*, *13*(1), 1–24. doi:10.1111/1532-7795.1301001

Collins, W. A., Welsh, D. P., & Furman, W. (2009). Adolescent Romantic Relationships. *Annual Review of Psychology*, *60*(1), 631–652. doi:10.1146/annurev.psych.60.110707.163459

Crone, E. A. & Dahl, R. E. (2012). Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nature Reviews Neuroscience*, *13*(9), 636–650. doi:10.1038/nrn3313

Cunningham, W. A. & Brosch, T. (2012). Motivational Salience Amygdala Tuning From Traits, Needs, Values, and Goals. *Current Directions in Psychological Science*, *21*(1), 54–59. doi:10.1177/0963721411430832

Cyders, M. A. (2013). Impulsivity and the Sexes: Measurement and Structural Invariance of the UPPS-P Impulsive Behavior Scale. *Assessment*, *20*(1), 86–97. doi:10.1177/1073191111428762

Cyders, M. A., Littlefield, A. K., Coffey, S., & Karyadi, K. A. (2014). Examination of a short English version of the UPPS-P Impulsive Behavior Scale. *Addictive Behaviors*, *39*(9), 1372–1376. doi:10.1016/j.addbeh.2014.02.013

Davidow, J. Y., Foerde, K., Galván, A., & Shohamy, D. (2016). An Upside to Reward Sensitivity: The Hippocampus Supports Enhanced Reinforcement Learning in Adolescence. *Neuron*, *92*(1), 93–99. doi:10.1016/j.neuron.2016.08.031

Decker, J. H., Lourenco, F. S., Doll, B. B., & Hartley, C. A. (2015). Experiential reward learning outweighs instruction prior to adulthood. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(2), 310–320. doi:10.3758/s13415-014-0332-5

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From Creatures of Habit to Goal-Directed Learners: Tracking the Developmental Emergence of Model-Based Reinforcement Learning. *Psychological Science*, *27*(6), 848–858. doi:10.1177/0956797616639301

Del Giudice, M., Angeleri, R., & Manera, V. (2009). The juvenile transition: A developmental switch point in human life history. *Developmental Review*, *29*(1), 1–31. doi:10.1016/j.dr.2008.09.001

Denton, S. E. & Kruschke, J. K. (2006). Attention and salience in associative blocking. *Learning & behavior*, *34*(3), 285–304. doi:10.3758/bf03192884

van Duijvenvoorde, A. C. K., Zanolie, K., Rombouts, S. A. R. B., Raijmakers, M. E. J., & Crone, E. A. (2008). Evaluating the Negative or Valuing the Positive? Neural Mechanisms Supporting Feedback-Based Learning across Development. *Journal of Neuroscience*, *28*(38), 9495–9503. doi:10.1523/JNEUROSCI.1485-08.2008

Elkind, D. (1967). Egocentrism in Adolescence. *Child Development*, *38*(4), 1025–1034. doi:10.2307/1127100

Ellis, B. J., Del Giudice, M., Dishion, T. J., Figueredo, A. J., Gray, P., Griskevicius, V., … Wilson, D. S. (2012). The evolutionary basis of risky adolescent behavior: Implications for science, policy, and practice. *Developmental Psychology*, *48*(3), 598–623. doi:10.1037/a0026220

Fine, M. & McClelland, S. (2006). Sexuality Education and Desire: Still Missing after All These Years. *Harvard Educational Review*, *76*(3), 297–338. doi:10.17763/haer.76.3.w5042g23122n6703

Fischhoff, B. (2008). Assessing adolescent decision-making competence. *Developmental Review*, *28*(1), 12–28. doi:10.1016/j.dr.2007.08.001

Fischhoff, B., de Bruin, W. B., Parker, A. M., Millstein, S. G., & Halpern-Felsher, B. L. (2010). Adolescents' Perceived Risk of Dying. *The Journal of adolescent health : official publication of the Society for Adolescent Medicine*, *46*(3), 265–269. doi:10.1016/j.jadohealth.2009.06.026. pmid: 20159504

Fischhoff, B., Parker, A. M., de Bruin, W. B., Downs, J., Palmgren, C., Dawes, R., & Manski, C. F. (2000). Teen expectations for significant life events. *The Public Opinion Quarterly*, *64*(2), 189–205.

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, 201711978. doi:10.1073/pnas.1711978115

Flournoy, J. C. (2018a). Jflournoy/milav v0.1.0-alpha. doi:10.5281/zenodo.1294090

Flournoy, J. C. (2018b). Jflournoy/rw-model-demo v0.5.0-alpha. doi:10.5281/zenodo.1407189

Foulkes, L., Viding, E., McCrory, E., & Neumann, C. S. (2014). Social Reward Questionnaire (SRQ): development and validation. *Frontiers in Psychology*, *5*. doi:10.3389/fpsyg.2014.00201. pmid: 24653711

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2017). Visualization in Bayesian workflow. arXiv: 1709.01449 [stat]. Retrieved March 7, 2018, from http://arxiv.org/abs/1709.01449

Gelman, A. & Carlin, J. (2014). Beyond Power Calculations Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science*, *9*(6), 641–651. doi:10.1177/1745691614551642

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian Data Analysis, Third Edition* (3rd ed.). Boca Raton, FL: CRC Press.

Gelman, A., Hill, J., & Yajima, M. (2012). Why We (Usually) Don't Have to Worry About Multiple Comparisons. *Journal of Research on Educational Effectiveness*, *5*(2), 189–211. doi:10.1080/19345747.2011.618213

Gershman, S. J. (2015). A Unifying Probabilistic View of Associative Learning. *PLOS Comput Biol*, *11*(11), e1004567. doi:10.1371/journal.pcbi.1004567

Glimcher, P. W., Camerer, C. F., Fehr, E., & Poldrack, R. A. (Eds.). (2009). *Neuroeconomics: decision making and the brain*. Amsterdam: Elsevier.

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, *27*(4), 591–635. doi:10.1207/s15516709cog2704_2

Gregorich, S. E. (2006). Do Self-Report Instruments Allow Meaningful Comparisons Across Diverse Population Groups? Testing Measurement Invariance Using the Confirmatory Factor Analysis Framework. *Medical care*, *44*, S78–S94. doi:10.1097/01.mlr.0000245454.12228.8f

Grossberg, S. (1975). A Neural Model of Attention, Reinforcement and Discrimination Learning. In C. C. Pfeiffer & J. R. Smythies (Eds.), *International Review of Neurobiology* (Vol. 18, pp. 263–327). Academic Press. Retrieved May 14, 2018, from http://www.sciencedirect.com/science/article/pii/S0074774208600379

Guitart-Masip, M., Huys, Q. J., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *NeuroImage*, *62*(1), 154–166. doi:10.1016/j.neuroimage.2012.04.024

Harden, K. P. (2014). A Sex-Positive Framework for Research on Adolescent Sexuality. *Perspectives on Psychological Science*, *9*(5), 455–469. doi:10.1177/1745691614535934

Harden, K. P., Kretsch, N., Mann, F. D., Herzhoff, K., Tackett, J. L., Steinberg, L., & Tucker-Drob, E. M. (2017). Beyond dual systems: A genetically-informed, latent factor model of behavioral and self-report measures related to adolescent risk-taking. *Developmental Cognitive Neuroscience*. Sensitive periods across development, *25*, 221–234. doi:10.1016/j.dcn.2016.12.007

Harden, K. P. & Tucker-Drob, E. M. (2011). Individual differences in the development of sensation seeking and impulsivity during adolescence: Further evidence for a dual systems model. *Developmental Psychology*, *47*(3), 739–746. doi:10.1037/a0023279

Hullinger, R. A., Kruschke, J. K., & Todd, P. M. (2015). An Evolutionary Analysis of Learned Attention. *Cognitive Science*, *39*(6), 1172–1215. doi:10.1111/cogs.12196

James, J., Ellis, B. J., Schlomer, G. L., & Garber, J. (2012). Sex-specific pathways to early puberty, sexual debut, and sexual risk taking: Tests of an integrated evolutionary–developmental model. *Developmental Psychology*, *48*(3), 687–702. doi:10.1037/a0026427

Jarvinen, D. W. & Nicholls, J. G. (1996). Adolescents' social goals, beliefs about the causes of social success, and satisfaction in peer relations. *Developmental Psychology*, *32*(3), 435–441. doi:10.1037/0012-1649.32.3.435

Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Libby, V., Glover, G., … Casey, B. J. (2011). Behavioral and Neural Properties of Social Reinforcement Learning. *Journal of Neuroscience*, *31*(37), 13039–13045. doi:10.1523/JNEUROSCI.2972-11.2011

Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Powers, A., Mehta, N., … Casey, B. J. (2014). Adolescent-specific patterns of behavior and neural activity during social reinforcement learning. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 683–697. doi:10.3758/s13415-014-0257-z

Kang, Y., McNeish, D. M., & Hancock, G. R. (2016). The Role of Measurement Quality on Practical Guidelines for Assessing Measurement and Structural Invariance. *Educational and Psychological Measurement*, *76*(4), 533–561. doi:10.1177/0013164415603764

Kenrick, D. T., Neuberg, S. L., Griskevicius, V., Becker, D. V., & Schaller, M. (2010). Goal-Driven Cognition and Functional Behavior The Fundamental-Motives Framework. *Current Directions in Psychological Science*, *19*(1), 63–67. doi:10.1177/0963721409359281

Kiefer, S. M. & Ryan, A. M. (2008). Striving for social dominance over peers: The implications for academic adjustment during early adolescence. *Journal of Educational Psychology*, *100*(2), 417–428. doi:10.1037/0022-0663.100.2.417

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A Neostriatal Habit Learning System in Humans. *Science*, *273*(5280), 1399–1402.

Kruschke, J. K. (2011). Models of attentional learning. In E. M. Pothos & A. J. Wills (Eds.), *Formal Approaches in Categorization* (pp. 120–152). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511921322.006

LaFontana, K. M. & Cillessen, A. H. N. (2002). Children's perceptions of popular and unpopular peers: A multimethod assessment. *Developmental Psychology*, *38*(5), 635–647. doi:10.1037/0012-1649.38.5.635

LaFontana, K. M. & Cillessen, A. H. N. (2010). Developmental Changes in the Priority of Perceived Status in Childhood and Adolescence. *Social Development*, *19*(1), 130–147. doi:10.1111/j.1467-9507.2008.00522.x

Lakens, D. (2017). Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. doi:10.1177/1948550617697177

Larson, R. W., Richards, M. H., Moneta, G., Holmbeck, G., & Duckett, E. (1996). Changes in adolescents' daily interactions with their families from ages 10 to 18: Disengagement and transformation. *Developmental Psychology*, *32*(4), 744–754. doi:10.1037/0012-1649.32.4.744

Larson, R. & Richards, M. H. (1991). Daily Companionship in Late Childhood and Early Adolescence: Changing Developmental Contexts. *Child Development*, *62*(2), 284–300. doi:10.2307/1131003

Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based Learning: A General Model of Repeated Binary Choice. *Journal of Behavioral Decision Making*, *25*(2), 143–153. doi:10.1002/bdm.722

Li, Y., Xie, H., & Shi, J. (2012). Chinese and American children's perceptions of popularity determinants: Cultural differences and behavioral correlates. *International Journal of Behavioral Development*, *36*(6), 420–429. doi:10.1177/0165025412446393

Lockwood, P. L., Apps, M. A. J., Valton, V., Viding, E., & Roiser, J. P. (2016). Neurocomputational mechanisms of prosocial learning and links to empathy. *Proceedings of the National Academy of Sciences*, *113*(35), 9763–9768. doi:10.1073/pnas.1603198113

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological review*, *82*(4), 276–298. doi:10.1037/h0076778

Mallett, K. A., Lee, C. M., Turrisi, R., & Larimer, M. E. (2009). Assessing college students' perceptions of tolerance to alcohol using social cues: The Social Tolerance Index. *Addictive Behaviors*, *34*(2), 219–222. doi:10.1016/j.addbeh.2008.10.008

Mayeux, L., Sandstrom, M. J., & Cillessen, A. H. N. (2008). Is Being Popular a Risky Proposition? *Journal of Research on Adolescence*, *18*(1), 49–74. doi:10.1111/j.1532-7795.2008.00550.x

McAdams, D. P. (2009). *Power, intimacy, and the life story: Personological inquiries into identity.* (6th). New York: Guilford Press.

McCabe, C. J., Louie, K. A., & King, K. M. (2015). Premeditation moderates the relation between sensation seeking and risky substance use among young adults. *Psychology of Addictive Behaviors*, *29*(3), 753–765. doi:10.1037/adb0000075

McCormick, E. M. & Telzer, E. H. (2017). Adaptive Adolescent Flexibility: Neurodevelopment of Decision-making and Learning in a Risky Context. *Journal of Cognitive Neuroscience*, *29*(3), 413–423. doi:10.1162/jocn_a_01061

Meuwese, R., Cillessen, A. H. N., & Güroğlu, B. (2017). Friends in high places: A dyadic perspective on peer status as predictor of friendship quality and the mediating role of empathy and prosocial behavior. *Social Development*, *26*(3), 503–519. doi:10.1111/sode.12213

Mõttus, R., Briley, D. A., Zheng, A., Mann, F. D., Engelhardt, L. E., Tackett, J. L., … Tucker-Drob, E. M. (2018). Kids becoming less alike: A behavioral genetic analysis of developmental increases in personality variance from childhood to adolescence. *Journal of Personality and Social Psychology*. doi:10.1037/pspp0000194

Neel, R., Kenrick, D. T., White, A. E., & Neuberg, S. L. (2015). Individual Differences in Fundamental Social Motives. *Journal of Personality and Social Psychology*, *110*(6), 887–907. doi:10.1037/pspp0000068

Nelson, E. E., Jarcho, J. M., & Guyer, A. E. (2016). Social re-orientation and brain development: An expanded and updated view. *Developmental Cognitive Neuroscience*, *17*, 118–127. doi:10.1016/j.dcn.2015.12.008

Nelson, E. E., Leibenluft, E., McClure, E. B., & Pine, D. S. (2005). The social re-orientation of adolescence: a neuroscience perspective on the process and its relation to psychopathology. *Psychological Medicine*, *35*(2), 163–174. doi:10.1017/s0033291704003915

Ojanen, T., Grönroos, M., & Salmivalli, C. (2005). An Interpersonal Circumplex Model of Children's Social Goals: Links With Peer-Reported Behavior and Sociometric Status. *Developmental Psychology*, *41*(5), 699–710. doi:10.1037/0012-1649.41.5.699

Owens, L. & Duncan, N. (2009). They might not like you but everyone knows you: Popularity among teenage girls. *The Journal of Student Wellbeing*, *3*(1), 14–39. doi:10.21913/jsw.v3i1.408

Palminteri, S., Kilford, E. J., Coricelli, G., & Blakemore, S.-J. (2016). The Computational Development of Reinforcement Learning during Adolescence. *PLOS Computational Biology*, *12*(6), e1004953. doi:10.1371/journal.pcbi.1004953

Patrick, M. E. & Schulenberg, J. E. (2011). How trajectories of reasons for alcohol use relate to trajectories of binge drinking: National panel data spanning late adolescence to early adulthood. *Developmental Psychology*, *47*(2), 311–317. doi:10.1037/a0021939

Peters, S. & Crone, E. A. (2017). Increased striatal activity in adolescence benefits learning. *Nature Communications*, *8*(1), 1983. doi:10.1038/s41467-017-02174-z

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*, *17*(2), 117–133. doi:10.1007/BF01537962

Pfeifer, J. H. & Berkman, E. T. (2018). The Development of Self and Identity in Adolescence: Neural Evidence and Implications for a Value-Based Choice Perspective on Motivated Behavior. *Child Development Perspectives*. doi:10.1111/cdep.12279

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. Retrieved from https://www.R-project.org/

Rescorla, R. A. & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (6th, pp. 64–99). New York: Appleton-Century-Crofts.

Revelle, W. (2017). *Psych: Procedures for Psychological, Psychometric, and Personality Research*. Evanston, Illinois. Retrieved from https://CRAN.R-project.org/package=psych

Richards, M. H., Crowe, P. A., Larson, R., & Swarr, A. (1998). Developmental Patterns and Gender Differences in the Experience of Peer Companionship during Adolescence. *Child Development*, *69*(1), 154–163. doi:10.2307/1132077. JSTOR: 1132077

Roberti, J. W. (2004). A review of behavioral and biological correlates of sensation seeking. *Journal of Research in Personality*, *38*(3), 256–279. doi:10.1016/S0092-6566(03)00067-9

Rodkin, P. C., Ryan, A. M., Jamison, R., & Wilson, T. (2013). Social goals, social behavior, and social status in middle childhood. *Developmental Psychology*, *49*(6), 1139–1150. doi:10.1037/a0029389

Romer, D. (2010). Adolescent risk taking, impulsivity, and brain development: Implications for prevention. *Developmental Psychobiology*, *52*(3), 263–276. doi:10.1002/dev.20442

Romer, D., Reyna, V. F., & Satterthwaite, T. D. (2017). Beyond stereotypes of adolescent risk taking: Placing the adolescent brain in developmental context. *Developmental Cognitive Neuroscience*, *27*, 19–34. doi:10.1016/j.dcn.2017.07.007

Rose, A. J., Swenson, L. P., & Waller, E. M. (2004). Overt and Relational Aggression and Perceived Popularity: Developmental Differences in Concurrent and Prospective Relations. *Developmental Psychology*, *40*(3), 378–387. doi:10.1037/0012-1649.40.3.378

Rosenblau, G., Korn, C. W., & Pelphrey, K. A. (2017). A computational account of optimizing social predictions reveals that adolescents are conservative learners in social contexts. *The Journal of Neuroscience*, 1044–17. doi:10.1523/JNEUROSCI.1044-17.2017

Rosseel, Y. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. Retrieved from http://www.jstatsoft.org/v48/i02/

Ruch, W. & Zuckerman, M. (2001). Sensation Seeking and adolescence. In J. Raithel (Ed.), *Risikoverhaltensweisen Jugendlicher. Erklärungen, Formen udn Prävention* (pp. 97–110). Opladen: Leske + Budrich.

Santelli, J. S. & Beilenson, P. (1992). Risk factors for adolescent sexual behavior, fertility, and sexually transmitted diseases. *Journal of School Health*, *62*(7), 271–279. doi:10.1111/j.1746-1561.1992.tb01243.x

Savin-Williams, R. C. (1979). Dominance hierarchies in groups of early adolescents. *Child development*, *50*(4), 923–935. doi:10.2307/1129316

Schalet, A. T. (2011). *Not Under My Roof: Parents, Teens, and the Culture of Sex*. Chicago: University Of Chicago Press.

Selfhout, M., Burk, W., Branje, S., Denissen, J., Van Aken, M., & Meeus, W. (2010). Emerging Late Adolescent Friendship Networks and Big Five Personality Traits: A Social Network Approach. *Journal of Personality*, *78*(2), 509–538. doi:10.1111/j.1467-6494.2010.00625.x

Sercombe, H. (2014). Risk, adaptation and the functional teenage brain. *Brain and Cognition*. Special Issue on Reward and Regulatory Processes in Adolescence, *89*, 61–69. doi:10.1016/j.bandc.2014.01.001

Shanks, D. R. (2010). Learning: From Association to Cognition. *Annual Review of Psychology*, *61*(1), 273–301. doi:10.1146/annurev.psych.093008.100519

Shi, Y. & Moody, J. (2017). Most Likely to Succeed: Long-run Returns to Adolescent Popularity. *Social Currents*, *4*(1), 13–33. doi:10.1177/2329496516651642

Shulman, E. P., Harden, K. P., Chein, J. M., & Steinberg, L. (2015). Sex Differences in the Developmental Trajectories of Impulse Control and Sensation-Seeking from Early Adolescence to Early Adulthood. *Journal of Youth and Adolescence*, *44*(1), 1–17. doi:10.1007/s10964-014-0116-9

Shulman, E. P., Smith, A. R., Silva, K., Icenogle, G., Duell, N., Chein, J., & Steinberg, L. (2016). The dual systems model: Review, reappraisal, and reaffirmation. *Developmental Cognitive Neuroscience*, *17*, 103–117. doi:10.1016/j.dcn.2015.12.010

Smith, D. K., Leve, L. D., & Chamberlain, P. (2006). Adolescent Girls' Offending and Health-Risking Sexual Behavior: The Predictive Role of Trauma. *Child Maltreatment*, *11*(4), 346–353. doi:10.1177/1077559506291950

Srivastava, S. & Anderson, C. (2011). Accurate when it counts: Perceiving power and status in social groups. In J. L. Smith, W. Ickes, J. Hall, S. D. Hodges, & W. Gardner (Eds.), *Managing interpersonal sensitivity: Knowing when—and when not—to understand others* (pp. 41–58). Hauppage, NY: Nova Science Publishers.

Stan Development Team. (2018). RStan: the R interface to Stan. Retrieved from http://mc-stan.org/

Steinberg, L., Albert, D., Cauffman, E., Banich, M., Graham, S., & Woolard, J. (2008). Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: Evidence for a dual systems model. *Developmental Psychology*, *44*(6), 1764–1778. doi:10.1037/a0012955

Steinberg, L. & Morris, A. S. (2001). Adolescent Development. *Annual Review of Psychology*, *52*(1), 83–110. doi:10.1146/annurev.psych.52.1.83. pmid: 11148300

Stevenson, F., Zimmerman, M. A., & Caldwell, C. H. (2007). Growth Trajectories of Sexual Risk Behavior in Adolescence and Young Adulthood. *American Journal of Public Health*, *97*(6), 1096–101. doi:10.2105/AJPH.2005.074609

Suleiman, A. B. & Deardorff, J. (2015). Multiple Dimensions of Peer Influence in Adolescent Romantic and Sexual Relationships: A Descriptive, Qualitative Perspective. *Archives of Sexual Behavior*, *44*(3), 765–775. doi:10.1007/s10508-014-0394-z

Suleiman, A. B., Galván, A., Harden, K. P., & Dahl, R. E. (2017). Becoming a sexual being: The 'elephant in the room' of adolescent brain development. *Developmental Cognitive Neuroscience*. Sensitive periods across development, *25*, 209–220. doi:10.1016/j.dcn.2016.09.004

Suleiman, A. B. & Harden, K. P. (2016). The importance of sexual and romantic development in understanding the developmental neuroscience of adolescence. *Developmental Cognitive Neuroscience*, *17*, 145–147. doi:10.1016/j.dcn.2015.12.007

Susman, E. J. & Dorn, L. D. (2009). Puberty. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of Adolescent Psychology* (pp. 116–151). Hoboken, NJ: John Wiley & Sons Inc.

Todorov, A. (2008). Evaluating Faces on Trustworthiness. *Annals of the New York Academy of Sciences*, *1124*(1), 208–224. doi:10.1196/annals.1440.012

Udry, J. R. (1988). Biological Predispositions and Social Control in Adolescent Sexual Behavior. *American Sociological Review*, *53*(5), 709–722. doi:10.2307/2095817

Underwood, M. K., Kupersmidt, J. B., & Coie, J. D. (1996). Childhood Peer Sociometric Status and Aggression as Predictors of Adolescent Childbearing. *Journal of Research on Adolescence*, *6*(2), 201–223.

Vehtari, A., Gabry, J., Yao, Y., & Gelman, A. (2018). Loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. Retrieved from https://CRAN.R-project.org/package=loo

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. doi:10.1007/s11222-016-9696-4

Westfall, J., Nichols, T. E., & Yarkoni, T. (2017). Fixing the stimulus-as-fixed-effect fallacy in task fMRI [version 2; referees: 1 approved, 2 approved with reservations]. *Wellcome Open Research*, *1*(23). doi:10.12688/wellcomeopenres.10298.2

Westfall, J. & Yarkoni, T. (2016). Statistically Controlling for Confounding Constructs Is Harder than You Think. *PLOS ONE*, *11*(3), e0152719. doi:10.1371/journal.pone.0152719

Whiteside, S. P., Lynam, D. R., Miller, J. D., & Reynolds, S. K. (2005). Validation of the UPPS impulsive behaviour scale: a four-factor model of impulsivity. *European Journal of Personality*, *19*(7), 559–574. doi:10.1002/per.556

Wickham, H. (2017). *Tidyverse: Easily Install and Load the 'Tidyverse'*. Retrieved from https://CRAN.R-project.org/package=tidyverse

Wiggins, J. S. (1991). Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. In *Thinking clearly about psychology: Essays in honor of Paul E. Meehl, Vol. 1: Matters of public interest; Vol. 2: Personality and psychopathology.* (pp. 89–113). Minneapolis, MN, US: University of Minnesota Press.

Willoughby, T., Good, M., Adachi, P. J., Hamza, C., & Tavernier, R. (2013). Examining the link between adolescent brain development and risk taking from a social–developmental perspective. *Brain and Cognition*, *83*(3), 315–323. doi:10.1016/j.bandc.2013.09.008

Yuen, K. K. (1974). The two-sample trimmed t for unequal population variances. *Biometrika*, *61*(1), 165–170. doi:10.2307/2334299

Zuckerman, M., Eysenck, S. B., & Eysenck, H. J. (1978). Sensation seeking in England and America: Cross-cultural, age, and sex comparisons. *Journal of consulting and clinical psychology*, *46*(1), 139–149. doi:https://doi.org/10.1037%2F0022-006x.46.1.139