DEVELOPMENTAL DIFFERENCES IN SOCIAL LEARNING

Dissertation

zur Erlangung des akademischen Grades
Doctor rerum naturalium
(Dr. rer. nat.)

am Fachbereich Erziehungswissenschaft und Psychologie
der Freien Universität Berlin

vorgelegt von

Dipl.-Psych. Julia M. Rodriguez Buritica

Berlin 2016

GUTACHTER

Erstgutachter: Prof. Dr. Hauke R. Heekeren

Zweitgutachter: Prof. Dr. Ben Eppinger

Datum der Disputation: 24.02.2017

ACKNOWLEDGEMENTS

First and foremost, I wish to thank my supervisors, Ben Eppinger and Hauke Heekeren, for their patient guidance and support throughout my PhD. Thank you for trusting in me and for creating such a stimulating and friendly scientific environment.

I am thankful to Shu-Chen Li for her help setting up the very first important steps into my PhD and the scientific discussions, and my co-authors for helping me accomplish this dissertation.

I am grateful to the members of my dissertation committee for helping me to complete the final step of this dissertation.

Many thanks to the Berlin School of Mind and Brain and the FU Berlin for the funding, teaching and support.

I would like to thank the LIP support team, for their help at all times, and Berndt Wischnewski for his programming support.

I would like to express gratitude to my current and former colleagues at the FU Berlin for the fruitful scientific discussions and the helpful comments. Special thanks to Yulia Oganian for the inspiring scientific discussions and the great times

Finally, I would like to thank my friends and family for their great support and trust.

Table of Contents

ABSTRACT

Learning from other's experiences (i.e., social learning) is fundamental for human development and important for educational practice. In contrast to learning from personal experience (i.e., experience-based learning), social learning has hardly been investigated from a developmental cognitive neuroscience perspective. It is important to utilize this perspective to broaden our knowledge of social learning and to gain a more mechanistic view on fundamental learning principles across development. For this dissertation project, I designed three age-comparative studies (in 8-10 year old children, 13-15 year old adolescents, and young adults) to investigate developmental differences in experience-based and social learning by employing behavioral, electrophysiological and computational analyses. I addressed the questions how peers and non-peers (i.e., adults) influence learning from other's actions and outcomes (observational learning) in children and how observational learning and advice taking from peers vary across development. Our findings provide important novel insights into the development of social learning. Overall, social learning was beneficial (in terms of accuracy and higher earnings) compared to trial-and-error type learning (i.e., experience-based learning) across development. Importantly, social and individual information were weighted differently across development. (1) During middle childhood, we found enhanced event-related potentials (ERPs), as well as, more pronounced imitative choice behavior when observing peers compared to non-peers (i.e., young adults). (2) Children learned more slowly than adults from peer behavior and showed difficulties to use observed feedback for learning; as seen by enhanced medial prefrontal ERPs and no learning-related changes in parietal ERPs. (3) We found specific developmental differences in learning from good peer advice for children, adolescents, and adults. Adolescents were initially highly sensitive to social information of their peers, but quickly used their own experiences for learning and explored more as compared to adults. Although children show higher exploration similarly to adolescents, they could not benefit from it to the same degree. Whereas adolescents (as compared to children and adults) selected choices with higher earning, children showed difficulties using negative feedback for learning. Taken together, adolescents showed benefits and children difficulties combining social and individual information when compared to adults. In terms of educational practice, this highlights that different learning forms might be appropriate for different age groups. Children are highly sensitive to negative experiences and have difficulties learning from them. Moreover, particularly during adolescence, own experiences (positive or negative) and the freedom to explore seem to be beneficial for learning.

ZUSAMMENFASSUNG

Das Lernen aus Erfahrungen Anderer (soziales Lernen) ist fundamental für die menschliche Entwicklung und wichtig für die pädagogische Praxis. Im Gegensatz zum Lernen aus eigenen Erfahrungen (erfahrungsbasiertes Lernen), mangelt es an Untersuchungen des sozialen Lernens über die Entwicklungsspanne aus einer kognitiv-neurowissenschaftlichen Perspektive. Es ist wichtig, diese Lücke zu schließen, um einen mechanistischeren Blick auf fundamentale Prinzipien des sozialen Lernens über die Entwicklungsspanne zu erhalten. Innerhalb der vorliegenden Promotion wurden drei altersvergleichende Studien (in 8-10-, 13-15-Jährigen und jungen Erwachsenen) entwickelt, um Entwicklungsunterschiede im erfahrungsbasierten und sozialen Lernen mit Hilfe von Verhaltens-, elektrophysiologischen und komputationalen Analysen zu untersuchen. Diese Studien konzentrierten sich auf die Forschungsfragen, inwiefern Peers (gleichen Alters und Geschlecht) und Nicht-Peers (d.h. junge Erwachsene) das kindliche Lernen von Handlungen und Handlungskonsequenzen Anderer (d.h. Beobachtungslernen) beeinflussen und inwiefern sich das Lernen aus Beobachtungen und Ratschlägen von Peers über die Entwicklung verändert. Unsere Befunde liefern erste wichtige Einblicke in entwicklungsbezogene Veränderungen im sozialen Lernen: Soziales Lernen war vorteilhaft (im Sinne von richtigen Antworten und Gewinnen) im Vergleich zum erfahrungsbasierten Lernen und das über die Entwicklung hinweg. Allerdings zeigten sich auch spezifische Entwicklungsunterschiede in der Gewichtung eigener und sozialer Informationen. (1) Kinder wiesen erhöhte medial-präfrontale ereigniskorrelierte Potentiale (EKPs) auf und imitierten beobachtete Handlungen von Peers im Vergleich zu Nicht-Peers häufiger. (2) Im Vergleich zu Erwachsenen lernten Kinder langsamer aus beobachteter Information und zeigten Schwierigkeiten im Lernen aus beobachtetem Feedback. Dies spiegelte sich in erhöhten medial-präfrontalen EKPs und fehlenden lernbezogenen Veränderungen in parietalen EKPs wider. (3) Zudem unterschieden sich Kinder, Jugendliche und Erwachsene darin, wie sie einen guten Ratschlag eines Peers zum Lernen nutzten: Jugendliche zeigten eine hohe anfängliche Sensitivität für Peer-Informationen, nutzten jedoch schnell ihre eigenen Erfahrungen zum Lernen und explorierten mehr als Erwachsene. Kinder explorierten ähnlich stark wie Jugendliche, konnten aber davon nicht vergleichbar profitieren. Wohingegen Jugendliche (im Vergleich zu Kindern und Erwachsenen) Optionen mit höheren Erträgen auswählten, hatten Kinder Schwierigkeiten, aus negativem Feedback zu lernen. Zusammenfassend zeigten Jugendliche

Vorteile und Kinder Schwierigkeiten (im Vergleich zu Erwachsenen) darin, soziale und individuelle Informationen adäquat zum Lernen zu nutzen. In Bezug auf die pädagogische Praxis verdeutlichen diese Befunde, dass womöglich unterschiedliche Lernformen für verschiedene Altersgruppen angemessen sind. Kinder im mittleren Kindesalter sind sensitiv für negative Erfahrungen und haben Schwierigkeiten diese zum Lernen zu nutzen. Jugendliche lernen sowohl aus positiven als auch negativen Erfahrungen und zeigen, dass ihre erhöhte Exploration vorteilhaft beim Lernen sein kann.

LIST OF ABBREVIATIONS

EEG	Electroencephalography
fMRI	Functional magnetic resonance imaging
ERP	Event-related potentials
oERP	Observational event-related potentials
FRN	Feedback-related negativity
oFRN	Observational feedback-related negativity
oP300	Observational P300
RL	Reinforcement learning
EL	Experience-based learning
OL	Observational learning
WM	Working memory

1. Introduction

Since our first moments in life we are highly sensitive to social information (Meltzoff, Kuhl, Movellan, & Sejnowski, 2009; Meltzoff & Moore, 1977). As we grow older, learning from social information[1] shapes our cognitive and socio-emotional development (Tomasello, Carpenter, Call, Behne, & Moll, 2005; Nielsen & Tomaselli, 2010; Meltzoff, Waismeyer, & Gopnik, 2012; Frith & Frith, 2003, 2007, 2012; Meltzoff et al., 2009). Beginning with school, social interactions with and learning from peers become increasingly important (Blakemore & Mills, 2014; Steinberg, 2008) and affect learning (Silva, Shulman, Chein, & Steinberg, 2015). During that time particularly negative social feedback of own peers can impact long-term mental health up to 40 years later (Takizawa, Maughan, & Arseneault, 2014). Moreover, maltreatment by peers affects long-term mental health even more strongly than maltreatment by adults (Lereya, Copeland, Costello, & Wolke, 2015). Thus, age similarity to others we are learning from is important for social learning during development. Social learning in children also depends on the consequences of others' actions (whether it resulted in positive or negative outcomes) and on other's recommendations (Lourenco et al., 2015; Morgan, Laland, & Harris, 2015). Although social learning in children has been studied for decades (Bandura, Grusec, & Menlove, 1966; Braaksma, Rijlaarsdam, & van den Bergh, 2002; Coates & Hartup, 1969; Harper & Sanders, 1975; Ladd, 1981; Nagell, Olguin, & Tomasello, 1993; Zimmerman & Rosenthal, 1974), investigations across development and with respect to the underlying neural dynamics are still missing.

Research in adults highlights that social learning is linked to similar underlying neural dynamics as learning from one's own experience (Burke, Tobler, Baddeley, & Schultz, 2010; Cooper, Dunne, Furey, & O'Doherty, 2012). Developmental studies show that learning from one's own experience is related to differences in prefrontal brain areas (e.g. van Duijvenvoorde, Zanolie, Rombouts, Raijmakers, & Crone, 2008), areas that show a protracted maturation until adulthood (Gogtay et al., 2004; Lenroot & Giedd, 2006). Van Duijvenvoorde et al. (2008) suggest that sensitivity to especially negative feedback (signaling the need for behavioral adjustments) develops across adolescence. Whether developmental differences previously described during learning from own experiences apply also to devel-

[1] Social learning is most commonly defined as "learning that is influenced by observation of, or interaction with, another animal (typically a conspecific) or its products" (Heyes, 1994, p. 207).
[2] If the learning context is linked to certain experiences that might scale with age, such as language, (Jaswal &

opmental differences during social learning, is an open question. It is important to enhance our understanding of developmental differences during social learning. This could be highly relevant if we want to know *whether or not* and *why* particularly sensitive periods to negative social feedback exist and how to work with them – from clinical and educational perspectives.

The aim of this doctoral thesis is to advance our understanding of developmental differences in social learning using behavioral, computational and electrophysiological approaches. The thesis is structured along three major research questions, which were addressed in each of the conducted empirical studies: (1) How does observed information about the actions and outcomes of peers (i.e., children) and non-peers (i.e., young adults) affect learning in children (study I)? (2) How does observed information about the actions and outcomes of peers affect learning across development (study II)? (3) How does advice from peers regarding a specific action affect learning across development (study III)?

Before answering these questions, I will briefly outline the existing empirical evidence related to these questions and describe the broader theoretical and empirical background.

2. Theoretical and empirical background

2.1 Social learning from peers, from observation and from advice

How we use social information for learning depends on the characteristics of those we observe, such as similarity to the observed (Fukushima & Hiraki, 2009; Mobbs et al., 2009), the consequences of what we observe (i.e., whether it resulted in positive or negative action-outcomes; Burke et al., 2010) and whether specific actions are recommended by others (Biele, Rieskamp, & Gonzalez, 2009; Biele, Rieskamp, Krugel, & Heekeren, 2011).

As in adults, perceived similarity (or dissimilarity) between the observer and the observed (Bandura, 1977; Owens & Ascione, 1991; Schunk, 1987) influences the integration of social information across development (Hendy & Raudenbush, 2000). Results of developmental studies suggest that children's similarity in age to the observed person (i.e., peers vs. young adult) predicts the degree to which the observed behavior of the other is integrated into one's own actions (Bandura, 1977; Schunk, 1987; Zmyj & Seehagen, 2013). Thus, peers serve as stronger role models for children than adults (Hendy & Raudenbush, 2000; Schunk, 1987; Schunk & Usher, 2012; van Gog & Rummel, 2010; Zmyj & Seehagen, 2013). This should be particularly the case if peers are not judged less competent.[2] Peers become increasingly influential during childhood, particularly during adolescence (Blakemore & Mills, 2014). During adolescence, peers influence not only risk taking, for instance gambling (Smith, Chein, & Steinberg, 2014) or driving (Chein, Albert, O'Brien, Uckert, & Steinberg, 2011; Simons-Morton, Lerner, & Singer, 2005), but also learning from positive and negative feedback (Silva et al., 2015; see also van Hoorn, van Dijk, Meuwese, Rieffe, & Crone, 2016 for review).

Although similarity between the observer and the observed seems important during social learning, action-outcomes (i.e., observational reinforcment learning[3], [OL]; [Burke et al., 2010; Cooper et al., 2012; Hill, Boorman, & Fried, 2016]) and action-recommendations

[2] If the learning context is linked to certain experiences that might scale with age, such as language, (Jaswal & Neely, 2006; Rakoczy, Hamann, Warneken, & Tomasello, 2010), the evaluation of nutritional values of food (Vander Borght & Jaswal, 2009) and expertise (Meshi, Biele, Korn, & Heekeren, 2012), the behavior of older/more experienced models is more likely to be used for one's own behavior.

[3] Observational reinforcement learning has been described as a "subset of response-reinforcer learning (R-S) in which observation of a demonstrator exposes the observer to a relationship between a response and a reinforcer (…)" (Heyes, 1994, p. 225).

(i.e., advice[4] [Biele et al., 2009, 2009; Lourenco et al., 2015; Meshi et al., 2012]) influence social learning (discussed in further detail under 2.2.3.). Using computational modeling and fMRI, these studies show similar learning mechanisms during social learning as during learning from one's own experience (i.e., experience-based reinforcement learning [EL]) and link those to prefrontal areas, which show a protracted maturation until adulthood (Gogtay et al., 2004; Lenroot & Giedd, 2006).

Taken together, peers are important for social learning in children and adolescents. In adults, social learning from other's action-outcomes and other's recommendations are linked to brain areas which are not fully matured in children and adolescents. The underlying neural and computational mechanisms of social learning (i.e., during OL and advice taking) across development are currently unclear. How higher peer sensitivity influences the processing of observed action-outcomes of peers and peer advice across development is an open question. So far, conclusions can only be drawn from existing evidence of EL across development. I will now outline the basic principles and underlying mechanisms of EL in adults and describe them in more detail across development.

2.2 Experience-based reinforcement learning and social learning in adults

I will now discuss the prediction error, an important learning signal during learning action-outcome-contingencies, reinforcement learning (RL), and the underlying mechanisms during EL and social learning in adults with the help of electroencephalography (EEG) studies.

2.2.1 Prediction error as reinforcement learning signal

The appropriateness of behavior can be evaluated over resulting action-outcomes. While positive action-outcomes (e.g. positive feedback or gains) encourage stay behavior, negative action-outcomes (e.g. negative feedback or losses) encourage shift behavior (i.e., shifting towards new behavioral patterns). Learning action-outcome-contingencies can be captured using RL algorithms (Sutton & Barto, 1998). The basic principle is nicely summarized by Niv & Schoenbaum (2008, p. 266): "make the best prediction you can, observe actual events and if your prediction was wrong, update your knowledge-base so that future predictions are more

[4] Advice is mostly considered "as recommendation, from the advisor, favoring a particular option" (Bonaccio & Dalal, 2006, p. 128).

accurate". Thus, discrepancy between what was expected, $Q_a(t)$, and what is actually observed, $r(t)$, is called prediction error (PE):

$$\text{Prediction error} = r(t) - Q_a(t)$$

If action-outcomes are better (worse) than expected, this will generate a positive (negative) PE, which is used to increase (decrease) the expected value, $Q_a(t)$, associated with the chosen option a in the current trial t (see Figure 1 for a schematic illustration).

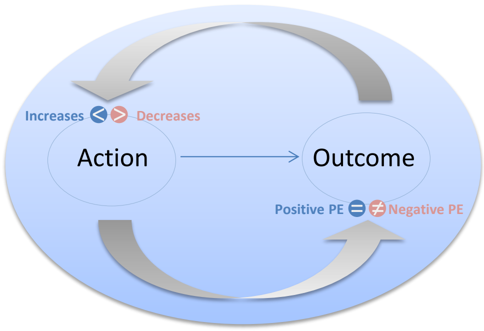

Figure 1. Schematic illustration of the **PE as a learning signal** during learning from action-outcomes.

The PE is related to phasic changes in the activity of midbrain dopamineric neurons (Bayer & Glimcher, 2005; Montague, Dayan, & Sejnowski, 1996; Montague, Hyman, & Cohen, 2004; Sadacca, Jones, & Schoenbaum, 2016). These neurons are reciprocally connected with the striatum (Joel & Weiner, 2000; Menegas et al., 2015; Watabe-Uchida, Zhu, Ogawa, Vamanrao, & Uchida, 2012). fMRI studies reveal that striatal activity is correlated with positive and negative PE (Diederen, Spencer, Vestergaard, Fletcher, & Schultz, 2016; McClure, Berns, & Montague, 2003; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). These striatal PE signals are linked to updated representations of expected values in the medial prefronal cortex (mPFC) (Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Rangel, Camerer, & Montague, 2008; Rushworth & Behrens, 2008). Comparisons of the connectivity between striatum and mPFC across development highlight that age-related differences in PE related activity are not found in the striatum, per se, but in the connectivity strength between striatum and mPFC (van den Bos, Cohen, Kahnt, & Crone, 2012). This finding suggests that developmental differences in EL are linked to differences in the updating of learning signals guiding future expectations and behavior. Furthermore, developmental differences in EL relate to activity changes in prefrontal areas (e.g. Christakou et al., 2013; Crone, Zanolie, Van Leijenhorst, Westenberg, & Rombouts, 2008; Hauser, Iannaccone, Walitza, Brandeis, & Brem, 2015; van den Bos, Güroğlu, Van Den Bulk, Rombouts, & Crone, 2009; van den Bos et al., 2012; van Duijvenvoorde et al., 2008), regions that are not yet fully developed in children (e.g. Gogtay et al., 2004; Lenroot & Giedd, 2006) and are central for cognitive control[5] (Alexander & Brown, 2011; Ridderinkhof, 2004).

The evaluation and updating of action-outcome information during EL have been better temporally dissociated using EEG (Ullsperger, Fischer, Nigbur, & Endrass, 2014). Before I outline developmental differences in EL using EEG, I will briefly introduce adult EEG-correlates of action-outcome processing.

2.2.2 Experience-based reinforcement learning in adults

Action-outcome processing operates in a time range of milliseconds (Keele & Posner, 1968). The investigation of temporal dynamics during action-outcome processing (i.e., of different valences) requires a measure with a high temporal resolution; EEG reflects brain activity (i.e., mainly synaptic potentials; Klee, Offenloch, & Tigges, 1965) with millisecond resolution. Previous studies identified several event-related potentials (ERP) that are sensitive to action-outcome (or short outcome) processing of different valences during EL (see Ferdinand & Kray, 2014 for review). Two ERP components are sensitive to the evaluation and updating of action-outcome or feedback information during learning: the feedback-related negativity (FRN) and the feedback-P300 (short P300 in the following). The FRN is a negative deflection in the EEG (see Figure 2) that is elicited in the medial prefrontal cortex (Hauser et al., 2014; Sambrook & Goslin, 2015) and that is sensitive to negative feedback during learning (Miltner, Braun, & Coles, 1997). The FRN is assumed to reflect an early binary evaluation of feedback along a good-bad dimension (Hajcak, Moser, Holroyd, & Simons, 2006; Nieuwenhuis, 2004; Philiastides, Biele, Vavatzanidis, Kazzer, & Heekeren, 2010; von Borries, Verkes, Bulten, Cools, & de Bruijn, 2013).

[5] Cognitive control can be defined as a "system for selecting contextually relevant information and for organizing and optimizing information processing" (Ridderinkhof, 2004, p. 306).

In addition to the FRN, a later positive deflection (see Figure 2), the P300 (Sutton, Braren, Zubin, & John, 1965), is maximal at parietal electrodes (Polich, 2007; San Martín, 2012). The P300 reflects context updating (Donchin, 1981; Polich, 2007). It scales with the expectedness of events (de Rover et al., 2015; De Taeye et al., 2014; Nieuwenhuis, Aston-Jones, & Cohen, 2005; Nieuwenhuis, De Geus, & Aston-Jones, 2011) and the degree to which information can be used to update reward predictions (Fischer & Ullsperger, 2013; Ullsperger et al., 2014).

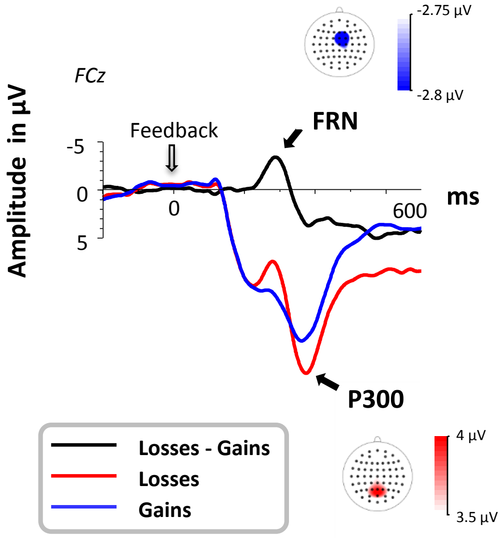

Figure 2. **ERPs to self-experienced feedback**. Grand averages shown for losses (red line) and gains (blue line) for the **FRN** and **P300** displayed at FCz. The topographic map displays the difference (black line) between losses and gains.

Most interestingly, both ERP-components are evoked when observing others' action-outcomes during learning and vary with similarity between the observer and the observed. Recent studies show how others' action-recommendations (i.e., advice) influence social learning. What we know so far about social learning in adults will now be outlined.

2.2.3 Social learning in adults

Previous research showed that adults benefit from social information during OL as compared to EL (Burke et al., 2010) or in form of advice (Biele et al., 2009, 2011). During OL, Burke et al. (2010) further highlighted that the observational PE, similarly to the PE during EL, triggers learning from observed feedback. Thus, during OL, observational and individual PE's are combined. This helps to quickly form expectation about the value associated with choice options and to benefit from additional social information more rapidly as compared to learning without social information (EL). During advice taking, outcomes of options recommended by others are more positively evaluated (i.e., associated with a constant *bonus* after choosing; see 4.3 for further details) as compared to outcomes of non-recommended options (Biele et al., 2011). As a consequence, advice helps choosing an option with higher earnings.

Action-outcome processing during social learning has been investigated using EEG (Bellebaum, Kobza, Thiele, & Daum, 2010; Clawson et al., 2014; Fukushima & Hiraki, 2009; Kang, Hirsh, & Chasteen, 2010; Rak, Bellebaum, & Thoma, 2013; Yu & Zhou, 2006). These studies show that adults react similarly to observed feedback of others as compared to self-experienced feedback. That is, the P300 and the FRN are both generated after observed feedback of others and are called observational FRN (oFRN) and observational P300 (oP300) (see Figure 3), respectively.

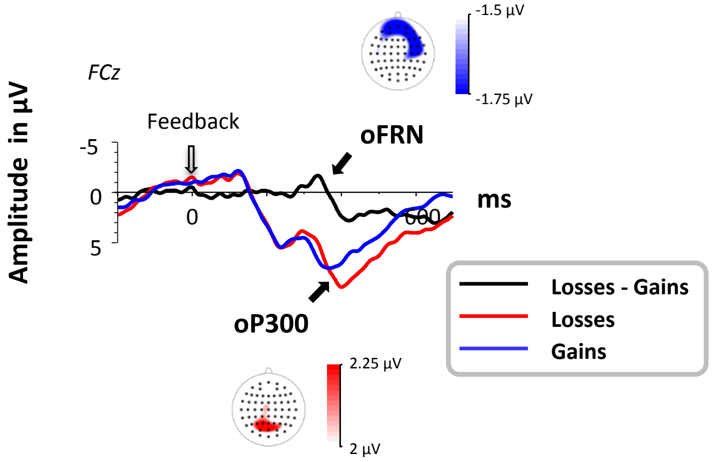

Figure 3. **ERPs to observed feedback**. Grand averages shown for losses (red line) and gains (blue line) for the **oFRN a**nd **oP300** displayed at FCz. The topographic map displays the difference (black line) between losses and gains.

OFRN and oP300 are sensitive to observed feedback of different valences. The oFRN is further linked to similar medial prefrontal source activations as the FRN (Koban, Pourtois, Bediou, & Vuilleumier, 2012). Whether both, observed and self-experienced feedback processing share similar or two distinct parallel mechanisms is still a matter of research (Behrens, Hunt, & Rushworth, 2009; Burke et al., 2010; Cooper et al., 2012; see Ruff & Fehr, 2014 for review). EEG-studies in adults reveal that the evaluation of observed feedback depends also on the characteristics of the observed, who receives the feedback. Here, the oFRN is reported to be larger observing humans vs. computers (Fukushima & Hiraki, 2009) familiar others vs. unfamiliar others (Kang et al., 2010) and similar vs. dissimilar others (Carp, Halenar, Quandt, Sklar, & Compton, 2009). Similarity further influences later behavioral adaptation (Hendy & Raudenbush, 2000).

Whether observed feedback in children would evoke ERPs to observed feedback as shown in adults and whether similarity to the observed person would further modulate this response, is an open question. Similarly, whether children ERPs to observed feedback show similarities to their ERPs to self-experienced feedback has to be investigated. Developmental findings point to specific developmental differences during EL, which will be described in the following.

2.3 Experience-based reinforcement learning across development

Previous developmental research described developmental differences during EL (Crone, Jennings, & van der Molen, 2004; Decker, Lourenco, Doll, & Hartley, 2015; Decker, Otto, Daw, & Hartley, 2016; Eppinger, Mock, & Kray, 2009; Hämmerer, Li, Müller, & Lindenberger, 2010; van den Bos et al., 2012; van Duijvenvoorde et al., 2008). These studies reveal that during EL children adjust their behavior more to current evidence (as compared to adolescents and adults), particularly after negative feedback, reflected in higher learning rates for losses (van den Bos et al., 2012). Children and adolescents rely also more on EL and less on explicit prior instructions as compared to adults (Decker et al., 2015). In the following, I discuss: (1) children's higher sensitivity to negative feedback and (2) how children and adolescents use EL and prior instructions as compared to adults.

2.3.1 Experience-based reinforcement learning: Using *feedback* for learning across development

The ability to effectively use (in particular negative) feedback for learning varies across development (see Ferdinand & Kray, 2014 for review) and is related to behavioral and electrophysiological differences:

Behavior. Children as compared to adults show greater difficulties in extracting relevant outcome information and in adapting their behavior to negative feedback (Crone et al., 2004; Crone, Somsen, Beek, & Van Der Molen, 2004; Crone, Somsen, Zanolie, & Van der Molen, 2006; Eppinger et al., 2009; Hämmerer et al., 2010; van den Bos, 2009; van den Bos et al., 2012; van Duijvenvoorde et al., 2008). Behavioral findings are further supported by children's heart rate slowing following negative feedback (Crone et al., 2004). Interestingly, slowing does not differentiate between response-dependent and uninformative negative feedback, as it is the case in adults. This finding nicely reflects children’s difficulties to assess and use particular negative outcome information to adapt their behavior later on. Developmental studies further suggest that positive and negative feedback are updated asymmetrically prior to adulthood (van den Bos et al., 2012). Van den Bos and colleagues (2012) show an age-related decrease in the impact of negative feedback on expected values using an RL algorithm. Children as compared to adults show higher learning rates for negative feedback, indicating that recent experience has a stronger influence on future predictions than less recent experience. This finding further supports previous literature showing a decreasing influence of irrelevant negative feedback during learning with increasing age (Crone et al., 2004; Eppinger et al., 2009). Developmental EEG-studies support and extend these developmental differences in using feedback for learning. I will discuss their main findings separately for the (1) FRN and (2) P300 in the following.

FRN. Children’s difficulties to disengage from negative feedback during learning are reflected in an enhanced FRN (Eppinger et al., 2009; Hämmerer et al., 2010; Santesso, Dzyundzyak, & Segalowitz, 2011; Zottoli & Grose-Fifer, 2012). Thus, children are less able to use external (particularly negative) feedback during learning (e.g. Crone et al., 2006). fMRI-studies complement EEG studies by adding important aspects of the neurodevelopmental functional changes underlying learning from positive and negative feedback (e.g. Christakou et al., 2013; Crone et al., 2008; Hauser et al., 2015; van den Bos, 2009; van den Bos et al., 2012; van Duijvenvoorde et al., 2008). Van Duijvenvoorde et al. (2008) show developmental differences in areas linked to cognitive control, such as the dorsolateral prefrontal cortex (DLPFC) and superior parietal cortex, regions that are not fully developed in children (e.g. Gogtay et al., 2004; Lenroot & Giedd, 2006). The authors observe that children recruit these areas more when receiving positive compared to negative feedback, whereas adults recruit these regions more after negative feedback. Van Duijvenvoorde et al. (2008) suggest that the greater difficulty in learning from negative feedback in children depends on the additional recruitment of cognitive control after negative feedback. That is, negative feedback signals an erroneous response, but also the need to update. Adolescents, however, show similar patterns in the DLPFC and superior parietal cortex comparable to adults (van Duijvenvoorde et al., 2008). This finding suggests that sensitivity to negative feedback that signals the need for behavioral adjustments develops across adolescence. Jointly, EEG studies show an enhanced FRN and difficulties to disengage from negative feedback during learning, which is assumed to be linked to less developed executive control functions (see Hämmerer & Eppinger, 2012 for review).

P300. Whereas children's FRN is enhanced compared to adults (Ferdinand & Kray, 2014), the reverse pattern is reported for the P300 component: children's P300 is reduced compared to adults (Polich, Ladish, & Burns, 1990; see van Dinteren, Arns, Jongsma, & Kessels, 2014 for review). The reduction in the P300 is linked to developmental differences in working memory (WM) abilities (Polich et al., 1990). Polich and colleagues (1990) show that the P300 response varies as a function of WM performance across development. This is in line with the context-updating framework (Donchin, 1981; Karis, Fabiani, & Donchin, 1984; Polich, 2007) suggesting that the P300 amplitude changes proportionally to the amount of updating in WM. This framework assumes a frequent stimulus-induced updating of a mental model. That is, if a stimulus representation maintained in WM mismatches with a recent stimulus, the model is updated and the P300 changes proportional to the update. More recent studies in adults further extend this view by suggesting that the differences in the P300 reflect the updating of reward expectations during learning (Philiastides et al., 2010). Thus, the P300 seems to reflect changes in reward expectations during EL. The processing of such reward expectations relate to activity in the locus coeruleus (LC), leading to a release of the neurotransmitter norepinephrine (NE) (Berridge & Waterhouse, 2003). Previous findings suggested that the P300 partly reflects the response of the LC-NE system (de Rover et al., 2015; De Taeye et al., 2014; Nieuwenhuis et al., 2005; Nieuwenhuis et al., 2011). That is, after an unexpected stimulus, reflected in a larger P300, phasic NE is released and promotes learning. Thus, in the LC-P3 account and the context-updating framework (Donchin, 1981;

Karis et al., 1984; Polich, 2007) the P300 is assumed to reflect the unexpectedness of stimuli, resulting in updating of an internal model.

Taken together, children's sensitivity to negative feedback is related to enhanced FRN-responses and greater difficulties to disengage from negative feedback. A reduced P300 in children as compared to adults is linked to differences in WM-updating. It is an open question whether children would show these electrophysiological differences also during OL.

2.3.2 Experience-based reinforcement learning: Using *instructions* for learning across development

A recent study by Decker et al. (2015) show that child and adolescent behavior is more influenced by their own experiences during RL and less by prior false instructions as compared to adults. In that particular learning environment, this lead to learning benefits of children and adolescents as compared to adults. The reduced bias toward false instructions helps to learn values of the better (non-instructed) choice alternatives faster. This is in line with studies showing that across development the efficiency in using specific rules for learning increases from early childhood (Munakata, Snyder, & Chatham, 2012) and continues into adolescence (Crone, Donohue, Honomichl, Wendelken, & Bunge, 2006; Huizenga, Crone, & Jansen, 2007). However, the study by Decker et al. (2015) could not rule out the possibility that developmental differences in response to false instructions are related to the higher sensitivity to negative feedback in younger age groups (van den Bos et al., 2012; van Duijvenvoorde et al., 2008). That is, children and adolescents may not have selected the instructed choice option because it was associated with a higher loss-probability as compared to the other options. On the other hand, it could also be suggested that children and adolescents simply explored more alternative options in line with studies showing higher exploration prior to adulthood (Buchsbaum, Bridgers, Weisberg, & Gopnik, 2012; Gopnik et al., 2015; Lucas, Bridgers, Griffiths, & Gopnik, 2014; Thompson-Schill, Ramscar, & Chrysikou, 2009). In part, by engaging in more exploration, children (i.e., preschoolers) learn the use of objects and their causal relationships (Lucas et al., 2014). Other studies show that children and even adolescents attempt unnecessary actions during learning (Nielsen & Tomaselli, 2010). Recent studies, however, suppose that an early period of more exploration might be important for cognitive development (Gopnik et al., 2015) and allow acting more effectively during adulthood (Buchsbaum et al., 2012).

Taken together, children and adolescents rely more on their own experiences and less on prior (false) instructions compared to adults. It is an open question whether this is due to developmental differences in the sensitivity to negative feedback and whether good instructions would result in similar developmental differences. It is unclear, whether prior social advice, particularly of another peer, would influence child and adolescent EL.

3. Research questions and hypotheses

In the previous sections, I outlined that (1) peers influence child and adolescent learning and decision-making. (2) Children show a high sensitivity to external (in particular negative) feedback and difficulties to use (in particular negative) feedback information for learning as compared to adults. (3) Children and adolescents rely more on their experiences during EL and less on prior instructions as compared to adults. It is an open question how these factors connect to each other and how they influence social learning across development.

The aim of this doctoral thesis is to advance our understanding of developmental differences in social learning from observation and advice using behavioral, computational and electrophysiological approaches; a developmental cognitive neuroscience perspective. More specifically, this doctoral thesis investigates how the characteristics, behavior (i.e., actions and outcomes) and advice (i.e., for a specific action) of other individuals affect learning across development. This thesis is structured along three major research questions (see Figure 4 for illustration):

1) How does observed information about the actions and outcomes of peers (i.e., children) and non-peers (i.e., young adults) affect learning in children (study I)?
2) How does observed information about the actions and outcomes of peers affect learning across development (study II)?
3) How does peer's advice regarding a specific action affect learning across development (study III)?

The *first research question* (RQ) concerns whether similarity in age between the observer and the observed individual influences learning in children. Study I compared how observed actions and outcomes of peers and non-peers (i.e., young adults) influences OL in children using EEG. We hypothesised that observing peers as compared to non-peers would result in enhanced oERP responses and greater behavioral adaptation in children.

The *second research question* focuses on developmental differences in social learning from peers' actions and outcomes (OL). To address RQ 2 we compared children and young adults in processing others' and self-experienced action-outcomes using EEG. We examined whether previously reported difficulties in children (as compared to adults) to use action-outcomes during EL also apply to OL. We hypothesised that children should show more difficulties to disengage and use external (observed and self-experienced) action-outcomes for learning. Therefore, we predicted enhanced and less learning sensitive oERP and ERP responses in children as compared to adults.

The *third research question* concerns developmental differences in how advice (others' action-recommendation) from another peer influences learning. If the consequences of others' actions are not observable, the quality of social information has to be evaluated based on one's own experience. Using computational models, we described how children, adolescents and young adults use advice and their own experience for learning. We hypothesised that adolescents (as compared to children and adults) would show the highest initial sensitivity to peer advice, reflected by their choice behavior. We predicted that with more time of learning from own actions, adolescents and children would rely less on the initial advice and more on their own experience. Children (as compared to adolescents and adults) should show more difficulties using negative outcomes for learning, reflected in higher learning rates for losses.

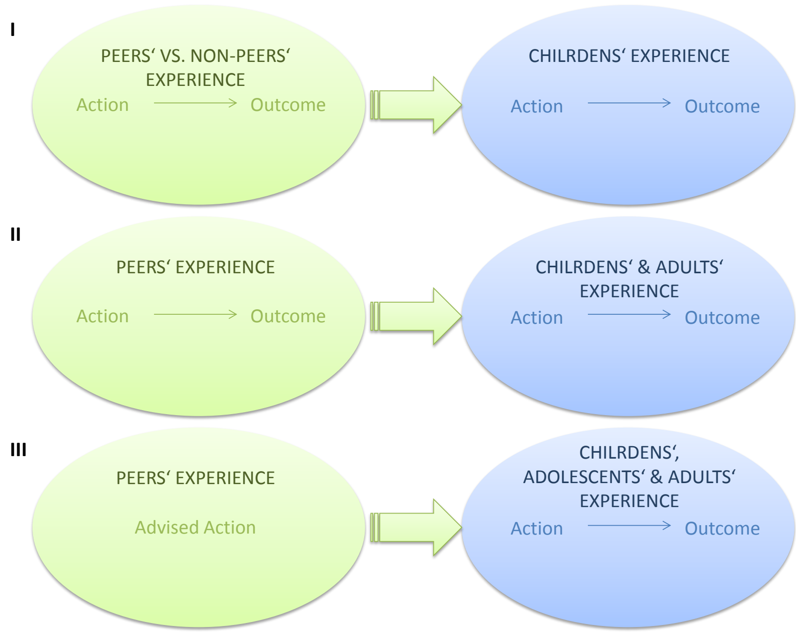

Figure 4. **Illustration of research questions.** How do other's experience influence one's own experience (I) in children when peer's or non-peer's action and outcomes were observable, (II) in children and adults when learning from peer's action and outcomes and (III) in children, adolescents and adults when learning from peer's advice.

4. Methods

I investigated developmental differences social learning from peers – during OL and advice taking - in three empirical studies. The task setup for the three studies will be briefly outlined in the following section. For the *first* and the *second* OL-study, we used the same 2-armed bandit task. For the *third* advice-study, we used a 4-armed bandit task and an extended RL algorithm to describe dynamics during learning from advice and experience. In the following section, I will outline our experimental designs and how we applied RL algorithms.

I aimed to create a more *realistic* social learning setting by testing participants within group sessions and pairing them with other participants for the EEG-session (study I and II; see Figure 5). The advice-study consisted of a single group session (study III).

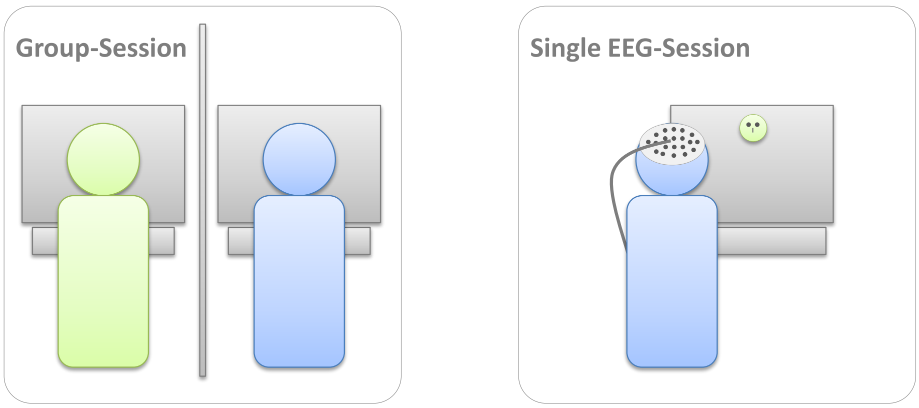

Figure 5. **Experimental setting.** In *study I* and *II*, participants were matched in the single EEG-session with a person from a previous group session and saw a real picture of this person. *Study III* consisted of one group session, were participants were told that they would receive advice of another peer out of a previous group session.

In the following I will describe the experimental designs of the three studies in further details:

4.1 Observational reinforcement learning task

In *study I* and *II* we used a probabilistic reward-based observational learning paradigm (Burke et al., 2010; Rodriguez Buritica, Eppinger, Schuck, Heekeren, & Li, 2016). Participants were asked to choose one out of two abstract stimuli (colored snowflakes [Windell, 2008]). One stimulus was associated with a high probability (80% gains, 20% losses) and one associated with a low probability (20% gains, 80% losses) of gaining points (see Figure 6A). Before they could choose, they observed another peer (who participated in the same previous group session) dealing with the same two abstract stimuli. Participants were told that the other player had already performed the task and that they could observe the recorded choices, which were - unbeknownst to participants - computer generated using a RL model (see mean learning curve of the “other player” in Figure 6C; publication I and II for further details). Participants were debriefed about the cover story after the experiment.

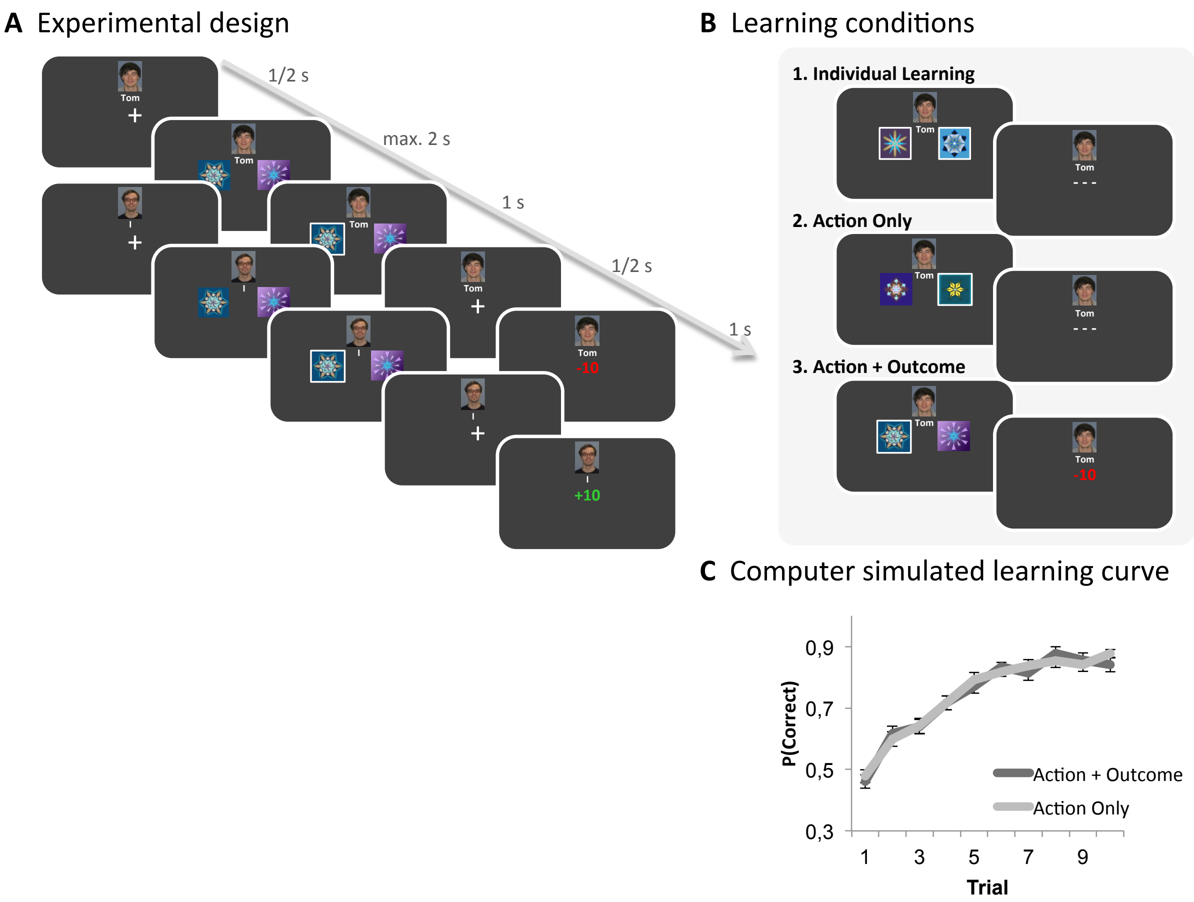

Figure 6. **Experimental Design. (A) Trial procedure. (B) Learning conditions.** 1: Individual Learning (IL); 2: Action Only (A), 3: Action + Outcome (AO). **(C) Computer simulated averaged learning curve** for the two observational conditions.

The amount of observable information of the other player was gradually manipulated across three learning conditions (see Figure 6B): (1) *individual learning* (neither the actions nor the outcomes of the other player were observable [IL]), (2) learning from observing only the other player’s actions (*action only* [A]), (3) observing both the other player’s actions and outcomes (*action + outcome* [AO]) (see Figure 6B). Each of the three conditions were associated with one stimulus pair, and presented within one block, for 10 trials (resulting in a block of 30 trials per block), for a total of 12 blocks. While the participants performed the task (controlled by PsychToolBox-3, Brainard, 1997), EEG was recorded continuously (Brain Amp DC, BrainVision Recorder software) from 64 Ag/AgCl electrodes (10-10 System, American Electroencephalographic Society, 1994) in an elastic cap (Braincap, BrainVision) (see publication I and II for further details regarding the EEG data analysis).

In *study I* we further varied, in addition to the amount of observable information, the age of the observed model player (another sex-matched, age-matched and randomly chosen peer (i.e., child) or a non-peer (i.e., adult), who participated in the same previous group session). This manipulation examines effects of similarity in age on the use of observed information for learning in children.

4.2 Advice taking task

As can be seen in Figure 7, in *study III* we used a probabilistic reward-based learning task (modified after Biele et al., 2011), where participants are supposed to gain as many points as possible by choosing more beneficial decks over the course of 210 trials.

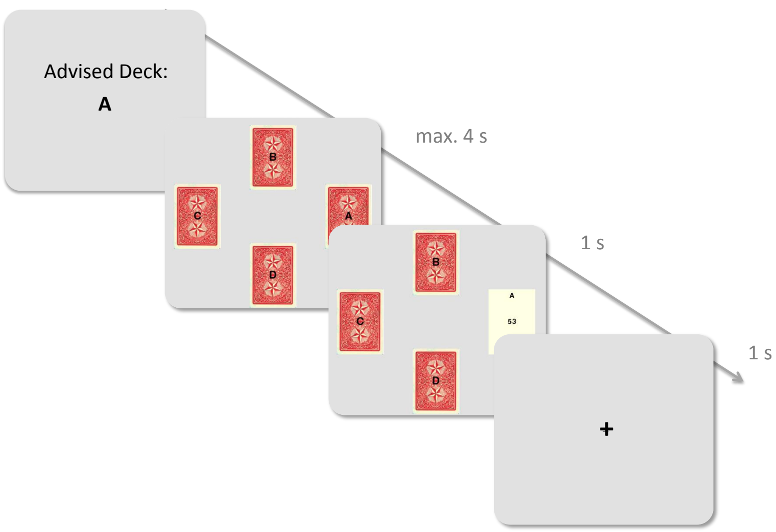

Figure 7. **Experimental design.** Participants received advice prior to when they were asked to play a 4-armed bandit task. Every trial started with the presentation of 4 card decks, where one should be selected within max. 4 seconds. Afterwards the associated feedback was presented. Before a new trial started a fixation cross was displayed for 1 second.

Decks that were more beneficial were associated with higher expected values, although each deck had a 50% probability of losses. Unbeknownst to the participants, two of four decks were associated with higher expected positive values ("good decks") than the other two ("bad decks"). At the beginning of the experiment, participants received a good advice for one of the "good decks" from another peer (see publication III for further details). Thus, the preference for the advised deck over the other good deck, would be a clear indicatior for the advice-effect.

4.3 Advice taking in reinforcement learning models

Learning action-outcome-contingencies can be computationally captured using RL models (Sutton & Barto, 1998). During RL learning the discrepancy between what was expected, $\mathrm{Q_a(t)}$, and what is actually observed, $\mathrm{r(t)}$, is called prediction error:

$$\text{Prediction error} = \mathrm{r(t)}\text{-}\mathrm{Q_a(t)}$$

If feedback is better (worse) than expected, the model will generate a positive (negative) prediction error, which is used to increase (decrease) the predicted value, $Q_a(t)$, associated with the chosen option a in the current trial t.

The impact of the prediction errors on forming new expectations is scaled by the learning rate α as follows:

$$Q_a\,(t+1) = Q_a(t) + \alpha[r(t) - Q_a(t)]$$

A high learning rate (~1) indicates that a new experience (i.e. prediction error) has a stronger impact on future predictions whereas a low learning rate (~0) means that a prediction error only weakly influences the expected value.

This basic RL-algorithm has been further extended to describe social influences during learning: by learning from other's choices and outcomes (Burke et al., 2010) or during advice-taking (Biele et al., 2009, 2011). In study III we investigated the influence of advice (i.e., for a certain option) on learning by using different extensions of RL models: We first compared how well an *outcome-bonus model*, a *prior model*, described participants' choices; and then a combined *prior & outcome-bonus model* (Biele et al., 2011).

The *outcome-bonus model* differs from the standard RL model by assuming that there is a constant bonus associated with choosing the advised option. The constant bonus was added to the objective reward (highlighted in bold) according to:

$$Q_a(t+1) = Q_a(t) + \alpha[r(t) + \boldsymbol{g(i)\mu\beta_b} - Q_a(t)$$

Here $g(i)$ serves as an indicatior for the advice, which takes the value 1 if the option was advised and the value 0 if it was not advised. The outcome-bonus parameter β_b captures the degree to which social influence leads to an outcome bonus, and μ is the expected payoff from choosing randomly among all options.

The simple *prior model* assumes an an initial strong positive prior for the advised option. The initial reward expectation in the prior model is defined as:

$$\boldsymbol{g(i)\mu\beta_p N}$$

Here β_p reflects the social influence on the prior expectations and N, the number of trials in the experiement serving as a scaling factor.

Finally, the combined *prior & outcome-bonus model* assumes an increased prior β_p for the advised option, and the advice is also associated with a constant bonus β_b.

We furthermore tested versions of the *social influence models* that assumed that the outcome-bonus associated with the advised option would decline over time (highlighted in bold). To model this decreasing influence of advice over time we modulated the bonus parameter:

$$Q_a\,(t+1) = Q_a(t) + \alpha[r(t) + g(i)\mu\beta_b \left(\boldsymbol{\frac{1}{t}}\right)^{\boldsymbol{\pi}} - Q_a(t)]$$

where π is the free parameter that captures how quickly the effect of influence decays, with smaller values indicating less decline ($0 < \pi < \infty$).

We extend all *social influence models* additionally with two independent learning rates: a learning rate for positive feedback (α_{pos}) and a learning rate for negative feedback (α_{neg}). Based on previous developmental studies (Palminteri, Kilford, Coricelli, & Blakemore, 2016; van den Bos et al., 2012) we expected that gains and losses are asymmetrically updated.

The initial expected values were set to 0, except for the the *prior model*. In addition, all *social influence models* assume that participants who received advice will always choose the advised option in their first trial accoring to (accoring to Biele et al., 2011). This was implemented by setting the probability of choosing the advised option to 1 for the first trial.

For all social influence models, the probability of the model choosing option a from a pair - according to the ratio of the Q values linked to each stimulus (a and c) - was computed using a softmax function (O'Doherty, 2004):

$$P_{(a)} = \frac{e^{Q_{a\,(t)/\beta}}}{\sum_{c \in A} e^{Q_{c\,(t)/\beta}}}$$

The probability of selecting option a is influenced by the expected value of option a in trial t divided by the sum of the expected values of all possible options A. The β parameter reflects the sensitivity of the subject to the differences in expected values. The lower β, the more exploratory choices appear.

5. Empirical studies

5.1 STUDY I: Observational reinforcement learning from peers vs. non-peers in children

Rodriguez Buritica, J. M., Eppinger, B., Schuck, N. W., Heekeren, H. R., & Li, S.-C. (2016). Electrophysiological correlates of observational learning in children. *Developmental Science*, *19*(5), 699–709.

Aims. Whether observed behavior of others is integrated in one's own behavior is modulated - next to other factors - by the perceived similarity (or dissimilarity) between the observer and the observed person (Bandura, 1977; Owens & Ascione, 1991; Schunk, 1987; Schunk & Usher, 2012). Whether 8-10 year old children show an ERP-response while observing other's feedback and whether similarity to the observed person (either another same-aged peer or an adult) further modulates this ERP-response, was unknown.

Hypotheses. Based on previous findings in adults (Burke et al., 2010) we predicted that children should benefit from additional social information during learning. We expected that social information is further evaluated by similarity in age (i.e., high similarity for peers): Children should prefer social information of similar others (i.e., peer) over dissimilar others (i.e., adults) (Hendy & Raudenbush, 2000; Schunk, 1987; van Gog & Rummel, 2010). Therefore, we expected larger oFRN observing similar others (i.e., peers) compared to dissimilar others (i.e., adult).

Methods. To address this question we used a probabilistic reward-based OL paradigm (adopted from Burke et al., 2010) in combination with EEG in 8-10 year old children (*N* = 31, 15 female, mean age = 8.94, *SD* = 0.85). We used a factorial 3 (learning condition) x 2 (model player) within-subject design to measure behavior (i.e. accuracy and imitative choice behavior) and ERPs for self-experienced (i.e., FRN and P300) and observed feedback (i.e., oFRN and oP300).

Results. The results show that children's accuracy increased as the amount of observable information across learning conditions increased (see Figure 8A). Interestingly, when analyzing how often other's choices were chosen independently of the feedback (called "imitative choice behavior"), children imitated the behavior of the same-aged child model player more often than of an adult model (see Figure 8B) and imitation was higher for the A as compared to the AO condition. Due to restrictions in available trials, we could not test whether children differently imitated with respect to the correctness of observed choices.

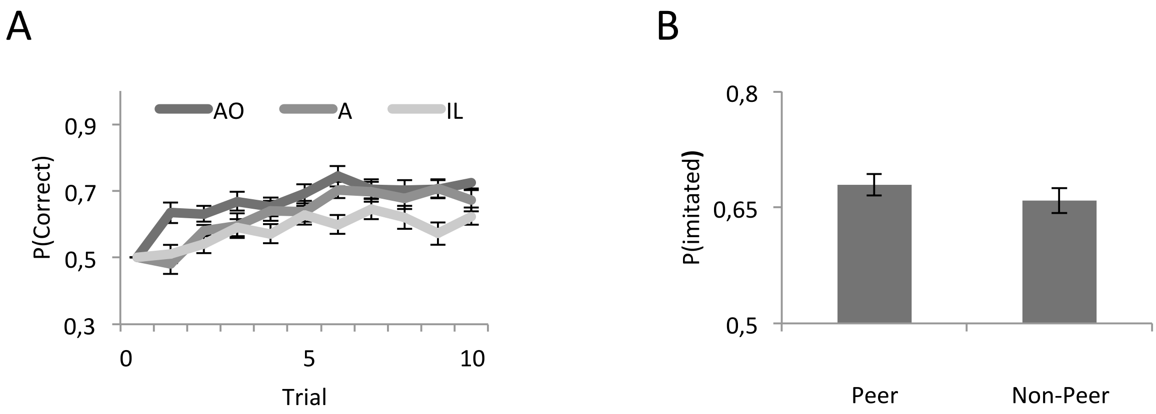

Figure 8. **Behavior. (A) Learning condition effects**. Accuracy in proportion correct for learning condition and trial (IL – Individual Learning; A –Action Only, AO – Action + Outcome). **(B) Imitative choice behavior.** Proportion of imitative choice behavior observing another peer (i.e., another child of the same age and same sex) and non-peer (i.e., young adult).

The FRN in children distinguished between losses and gains (as previously shown: e.g. Crone, 2014; Eppinger et al., 2009; see Ferdinand & Kray, 2014 for review; Hämmerer & Eppinger, 2012; Santesso et al., 2011) and was comparable across all learning conditions. In contrast, and in line with the behavioral learning effects, the feedback-locked P300 increased with the amount of observable information over the learning conditions. Thus, the feedback-locked P300 amplitude decreased from AO to IL learning condition. Most interestingly, we found that 8-10 year old children showed ERP responses (i.e., oFRN) to observed feedback, which were similar to their ERP responses to self-experienced-feedback. In contrast to previous adult studies, children's oFRN was not diminished compared to their FRN. This was especially true for observing peers, as oFRNs in response to observed feedback given to peers showed a trend of being larger compared to those when observing an adult (see Figure 9A vs. 9B).

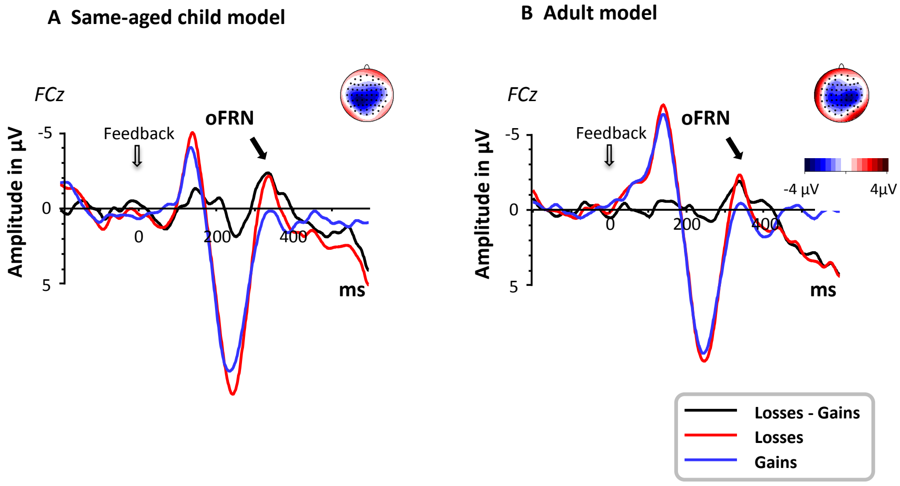

Figure 9. **OFRN to the to be observed model.** OFRN for gains and losses after observing a same-aged child or adult model. Recorded at FCz as peak-to-peak measures. Grand averages are shown for losses (red line) and gains (blues line). The topographic map displays the difference between the oFRN for losses and gains (black line) within a time window of 50 ms around the peak separately for the **(A) same-aged child model player** and **(B) adult model player** condition.

Taken together, study I answered RQ I as follows: children's oFRN was enhanced when observing feedback received by peers (i.e., another sex- and age-matched child) compared to non-peers (i.e., young adults). Children also imitated behavior of their peers as compared to non-peers more frequently.

5.2 STUDY II: Observational reinforcement learning from peers across development

Rodriguez Buritica, J. M., Heekeren, H. R., Li, S.-C. & Eppinger, B. (under revision). Developmental differences in the neural dynamics of observational learning.

Aims. In the second study, we described developmental changes during OL using EEG in 8-10 year old children and young adults. More specifically, we investigated whether developmental differences in FRN and P300 during EL (e.g. Eppinger et al., 2009; Hämmerer & Eppinger, 2012) also apply to OL. That is, (1) whether children show an enhanced oFRN (as they show an enhanced FRN) in comparisons to adults and (2) whether their oP300 is reduced (like their P300 [Polich et al., 1990; van Dinteren et al., 2014]) as compared to adults. We used 3 (learning condition) x 2 (learning changes by contrasting the beginning and the end across all learning blocks) within-subject factors and age (child vs. young adult) as a between-subject factor to measure behavior (i.e., accuracy) and ERPs for self-experienced (i.e., FRN and P300) and observed feedback (i.e., oFRN and oP300).

Hypotheses. We predicted that both groups should benefit (i.e., in terms of accuracy) from the additional social information during learning. Children, however, should show more difficulties in using (particularly negative) feedback for learning as compared to adults (e.g. Eppinger et al., 2009). This should be reflected in an enhanced FRN response as compared to adults (Ferdinand & Kray, 2014). In contrast to adults (Bellebaum et al., 2010), children's FRN and oFRN should not differ in amplitude (Rodriguez Buritica et al., 2016). We expected that children should be less able (as compared to adults) to efficiently update feedback information during learning (Ferdinand & Kray, 2014), reflected over a diminished P300 response in children compared to adults (van Dinteren et al., 2014). Developmental differences in updating of feedback information should further be related to developmental differences in WM capacity (Polich et al., 1990).

Methods. The effective sample consisted of 23 adults (20-30 years old; 11 female, mean age = 23.52, SD = 2.81) and 22 children (8-10 years old; 10 female, mean age = 9.05, SD = 0.79). General cognitive abilities of the sample were assessed using several psychometric tests, such as a modified version of the spatial n-back task to investigate WM capacity (described in detail by Li et al., 2008). Children had proportionally lower scores than adults on the WM test, which is in line with previous developmental studies (Fry & Hale, 1996) .

Results. Both age groups differed in behavior and ERP-response: Children showed more gradual observational learning compared to adults (see Figure 10).

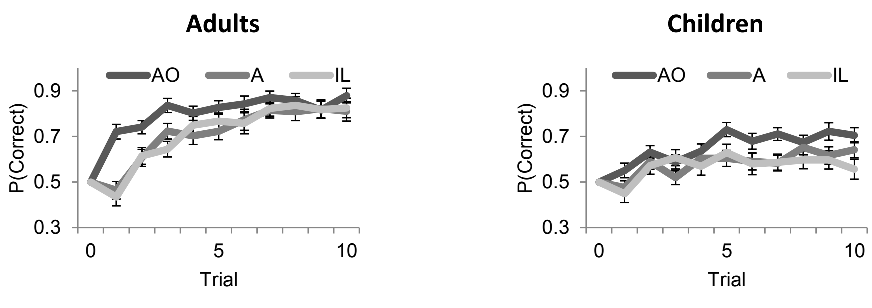

Figure 10. **Behavior. Learning & Condition Effects.** Accuracy in proportion correct for age group and learning condition displayed per trial (IL – Individual Learning; A –Action Only, AO – Action + Outcome).

Furthermore, they showed greater sensitivity to observed (particularly negative) feedback, reflected over their oFRN (see Figure 11A). Adults' more pronounced OL effects (see Figure

10) were related to reduced sensitivity to observed feedback (i.e., no oFRN valence effect; see Figure 11A). Childrens' oERPs (oFRN and oP300) were generally not diminished compared to their ERPs (FRN and P300; see Figure 11B), like it was for adults (see Figure 11).

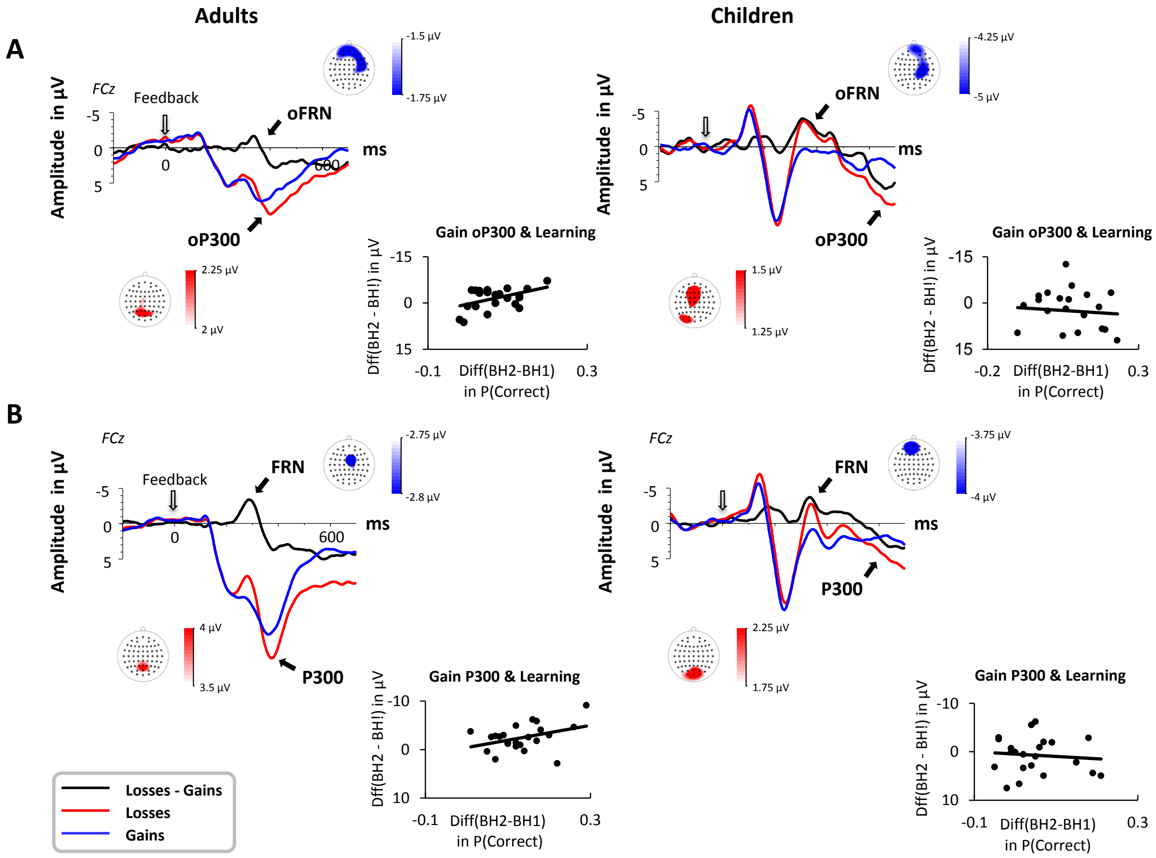

Figure 11. **ERPs to observed and experienced feedback**. Grand averages shown for losses (red line) and gains (blue line) for the **(A) oFRN/oP300** and **(B) FRN/P300** displayed at FCz separately for both age groups. The topographic map displays the difference (black line) between losses and gains of the oFRN/FRN (within 50 ms) and oP300/P300 (within 200 ms). *Correlation effects.* Scatter plots illustrate the correlation between difference scores of proportion of correct choice (second – first block half) on the x-axis and the difference score of the mean oP300/P300 amplitude to gains (second – first block half) on the y-axis separately for the two age groups.

The P300 to observed and self-experienced gains decreased with learning in adults, but not in children (see Figure 11). Interestingly in adults, the relation between learning and changes of the P300 was moderated by WM (see Figure 12).

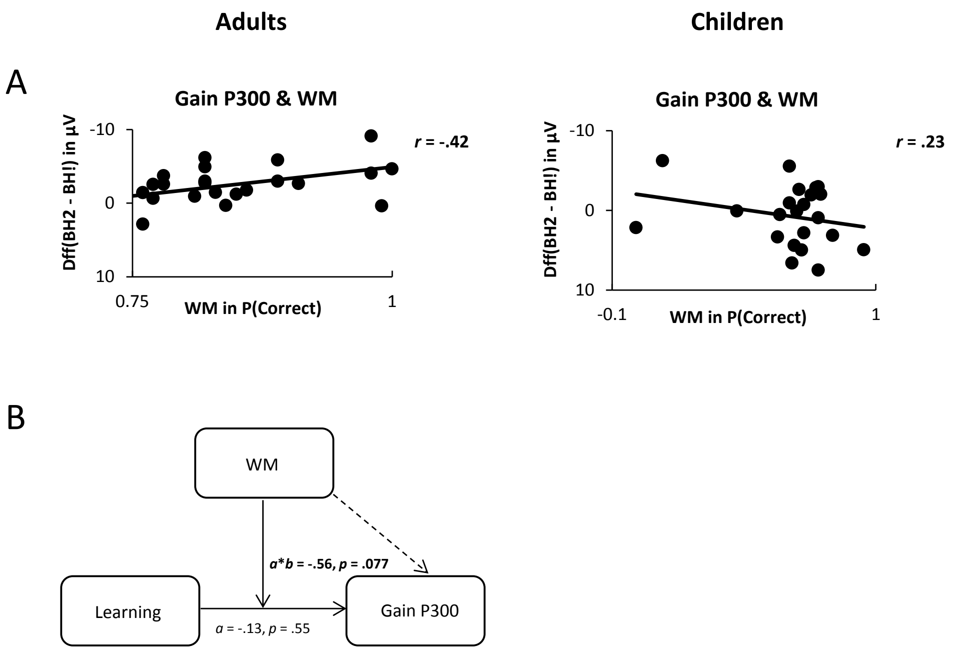

Figure 12. **Effects of learning on the P300**. **(A) Correlation effects**. Scatter plots illustrate the correlation between difference score of proportion correct in the WM task on the x-axis and the difference score of the mean P300 amplitude to gains (second – first block half) on the y-axis separately for the two age groups. **(B) Moderation model of P300 learning effects.** Working memory (WM) as a moderator for the relationship between learning and learning related changes in the P300 to gains: *a***b* path predicted learning related changes in the P300 to gains (neither the *a* path (learning as predictor) nor the *b* path (WM as predictor) reached significance).

Taken together, study II answered RQ II as follows: children's oERP responses were similarly to their ERP responses and showed specific developmental differences. Children, as compared to adults, seemed less able to disengage from observed negative feedback and to use observed feedback information to update their predictions, probably due to developmental differences in WM capacities.

5.3 STUDY III: Advice taking from peers across development

Rodriguez Buritica, J. M., Heekeren, H. R., & van den Bos, W. (under revision). Adolescents are sensitive to peer influence, but only for so long.

Aims. In this age-comparative study, we compared behavioral difference in the integration of advice, EL and the exploration of alternatives (in 8-10 year old children, 13-15 year old adolescents and 18-22 year old adults). Furthermore, we aimed to computationally describe developmental differences in advice taking by an extension of the *outcome-bonus* *model* (previously described in adults by Biele et al., 2011; see Chapter 4.3 for further details).

Hypotheses. All age groups, particularly adolescents, should be sensitive to the initial peer advice. As previously described (Biele et al., 2009, 2011) the advised option should be associated with a continuous *bonus* (i.e., the expected value of an option is always a bit higher, when being advised by another person) while choosing it. Children and adolescents, however, should be less biased than adults towards the initial advice and instead explore more alternative options (Decker et al., 2015). As a result, the two younger groups are more likely to discover the other (equally) “good deck”. Due to the fact that all decks were equally associated with gains and losses, children should encounter the other positive deck less often than adolescents due to their greater difficulties to use negative feedback for learning (van Duijvenvoorde et al., 2008). This should be further captured in higher learning rates for losses compared to gains in children (van den Bos et al., 2012).

Methods. In this study, we tested 25 adults (18-22 years old; 13 female, mean age = 20.32, *SD* = 1.15), 24 adolescents (13-15 years old; 12 female, mean age = 13.71, *SD* = 0.75) and 24 children (8-10 years old; 10 female, mean age = 9.08, *SD* = 0.83). Here, we applied and extended RL models (see Study I and 4.3. for further details) to capture developmental differences in the influence of advice, experience and exploration on learning. We measured P(chosen) per deck across 210 trials as a within-subject factor and age (children, adolescents, adults) as a between subject-factor.

Results. All age groups followed the good advice at the beginning of the task. With more time of sampling-behavior the age groups differed their choice behavior: Adolescents showed the highest initial sensitivity to peer advice (see Figure 13), whereas adults most consistently followed the advice throughout the task (Decker et al., 2015). This age-difference during initial advice taking was captured with the prior + bonus dual RL model (which was the winning model across our model comparison, see Figure 14A). This model revealed higher prior expectations based on the advice in adolescents compared to the other two age groups (see Figure 14B). Adolescents and children showed higher exploration rates (see Figure 14B), but due to children’s difficulties using negative feedback for learning (i.e., indicated by higher learning rates for losses; see Figure 14B), only adolescents explored the good alternative within the task (see Figure 13). Adolescents also selected the bad decks less often and showed highest expected earnings compared to the other two age groups.

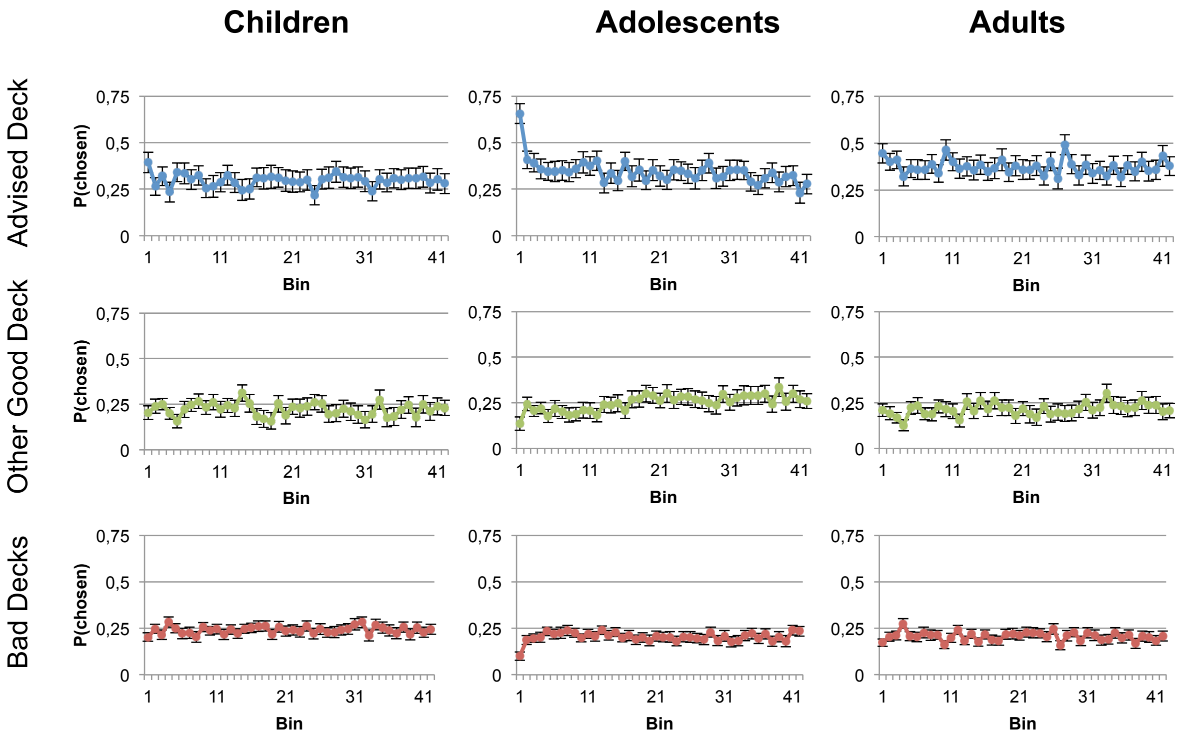

Figure 13. **Choice Behavior.** Proportion chosen separately for bin (of 5 trials each) and age group seperately for the advised, other good deck and bad decks.

No differences between age groups were found neither for their learning rate for gains, nor for their bonus parameter (see Figure 14B).

Taken together, the results of this study answered RQ III: peer advice influenced adolescents' behavior initially more strongly as compared to children and adults. However, only adults showed a more consistent influence of advice over time (as compared to children and adolescents). Children and adolescents showed more exploratory behavior. Whereas adolescents benefited from that, children did not and showed higher learning rates for losses. Adolescents were better able to use their own experience to choose options with higher expected values and to discover the other good choice alternative within the task. Our social learning model combined both (apparently) contradicting findings - the higher sensitivity to peer information and higher exploration.

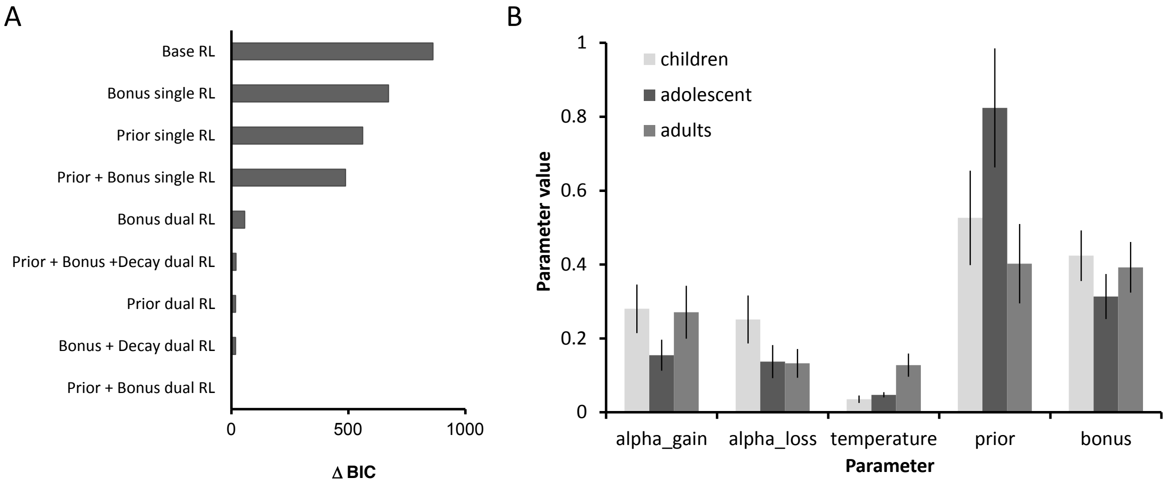

Figure 14. **(A) Relative BIC's for model comparision.** The relative difference in BIC values for each models compared to the model with the lowest over BIC value. The Bayes factor for comparing the best (lowest BIC) and second best model is 6494, which indicates the best fitting model is very strongly favored over all other models tested. Note that this model also is the winning model if we perform these comparisons on the level of age groups separately. **(B) Parameter estimates for the prior + bonus dual RL model.** Estimates separatly for the three age groups and the two learning rates (alpha_gain, alpha_loss), temperature, prior and bonus.

6. Discussion

This thesis provides a preliminary description of developmental differences in social learning, especially OL and advice taking. In three age-comparative studies I employed behavioral, EEG and computational analyses in 8-10 year old children, 13-15 year old adolescents and young adults. I mainly addressed three research questions: How (1) other's similarity in age to the learner (high for peers), (2) action and outcomes (during OL) and (3) action-recommendations (in form of advice) influence social learning across development. The answers to these questions will be discussed in the following:

6.1 Observational reinforcement learning: Using peers' and non-peers' information for learning in children

Whether social information is used for learning is influenced additionally by the characteristics of its sources (i.e., the observed model or the adviser), such as by similarity in age to the observer (Hendy & Raudenbush, 2000; Schunk, 1987). Developmental studies suggest that similarity in age is used as an indicator for the appropriateness of an observed behavior (see Schunk, 1987 for review). Study I supported this view: 8-10 year old children imitated choices of their peers as compared to choices of young adults more frequently. Study I extended this view by showing that higher peer-sensitivity also influenced the processing of observed feedback.

Children's oFRN to observed peer-feedback was enhanced as compared to their oFRN to observed *non*-peer-feedback (i.e. by young adults). Comparing oERPs (i.e., oFRN, oP300) to ERPs (i.e., FRN, P300), children's oERPs to their own peers were similar in magnitude to their ERPs differently from findings in adults (Bellebaum et al., 2010; Fukushima & Hiraki, 2009). The higher peer sensitivity to the observed feedback, reflected in childrens' oERPs, aligns with our finding that children imitated choices of own peers more frequently than choices of young adults. Both findings support the assumption that peers can serve as stronger role models for children than non-peers (Hendy & Raudenbush, 2000; Schunk, 1987; Schunk & Usher, 2012; van Gog & Rummel, 2010; Zmyj & Seehagen, 2013). Our ERP-results extended previous findings in adults (Carp et al., 2009; Mobbs et al., 2009) by showing that the oFRN is also (and probably more strongly) modulated by social factors in children as compared to adults.

6.2 Observational reinforcement learning: Using peers' information for learning across development

Developmental differences in using of observed feedback information for learning show similarities to developmental differences in using self-experienced feedback information for learning. Developmental studies on EL show that children, as compared to adults, have greater difficulties in extracting relevant feedback information and using (particularly negative) feedback for learning (Crone, 2014; Eppinger et al., 2009; Hämmerer & Eppinger, 2012; van den Bos, 2009; van den Bos et al., 2012; van Duijvenvoorde et al., 2008). Study II extended this view by showing that children (as compared to adults) also benefited more slowly from social information. Importantly, study II added electrophysiological evidence to this behavioral finding: Children showed (1) enhanced sensitivity to external (particular negative) observed and self-experienced feedback and (2) had difficulties to use this information during learning.

oFRN/FRN. During OL, children showed an enhanced oFRN-response (particularly to negative feedback) as compared to adults. This finding is consistent with the general notion that children are more susceptible to external (and especially to negative) feedback information during learning than adults (Eppinger et al., 2009; Ferdinand & Kray, 2014; Hämmerer et al., 2010). Moreover, it extended this notion by showing that this greater sensitivity to negative feedback in children can be generalized to observed negative feedback (Rodriguez Buritica et al., 2016). This might suggest that children (i.e., 8-10 years old) do not yet appropriately weight and integrate external information of different (and in particular negative) valences compared to adults (Eppinger et al., 2009; Hämmerer et al., 2010; van den Bos et al., 2012). Our findings showed that children do not yet appropriately assess the information value of feedback (that is, how much it should guide learning) and thus are less efficient in integrating (particularly negative) outcomes into their behavioral strategies than adults. Therefore, they might generally up-regulate their response to external (particularly negative) feedback, independently of who receives the feedback. Thus, developmental differences during OL link to a greater sensitivity to negative feedback, comparable to findings during EL.

oP300/P300. Our ERP findings suggest that children as compared to adults were less able to use observed feedback for learning. Children learned more gradually from OL and their oP300 to gains did not decreased with learning as found in adults. In adults, this learning-related decrease in the P300 to gains was positively correlated with their success in OL. Similarly, adult's P300 to gains decreased as a function of learning; WM capacity modulated this relationship. These findings are consistent with prior research showing a correlation of the P300 (but not the FRN) with reward prediction errors generated during RL (Fischer & Ullsperger, 2013; Ullsperger et al., 2014) and that reward-based learning depends on WM abilities (Collins & Frank, 2012). Collins and Frank (2012) show that WM abilities explain differences during RL. That is, learning was slower the greater the WM load. The fact that children P300 was not modulated by learning nor WM, suggested that children might not be as efficient, compared to adults, in using feedback information to update their predictions during learning (McGuire, Nassar, Gold, & Kable, 2014; van den Bos et al., 2012). This is possibly due to developmental differences in WM, which increase until young adulthood (Fry & Hale, 1996). Thus, developmental differences in learning may not be the consequence of differences in prediction error signaling, per se (Cohen et al., 2010; Hauser et al., 2015), but rather of determining how much to learn from a given prediction error (McGuire et al., 2014; van den Bos et al., 2012) based on WM abilities. Thus, extracting how much to learn from outcomes (worse or better than expected) might require sufficient WM abilities. Future studies should directly investigate to what extend are developmental differences during RL link to WM and whether children's difficulties to use negative feedback for learning are partly due to their WM abilities.

The results of *study II* support previous accounts of the P300 as a marker of context updating (Donchin, 1981). It is also in line with more recent accounts, suggesting that the P300 reflects the expectedness of events (de Rover et al., 2015; De Taeye et al., 2014; Nieuwenhuis et al., 2005, 2011) and the degree to which different information (of one's own and others' experiences) can be used to update reward predictions (Fischer & Ullsperger, 2013; Ullsperger et al., 2014). Combining these accounts, the P300 may be a marker of the computational mechanisms that determine the degree to which behavior is updated based on a given prediction error, which might depend on WM capacity (Nassar et al., 2016).

6.3 Advice taking: Using peers' information for learning across development

A recent study by Decker et al. (2015) show that children and adolescents rely more on their own experiences and less on prior (false) instructions than adults during EL. It was unclear whether this is due to developmental differences in the sensitivity to negative feedback and whether good instructions would result in similar developmental differences. It was an open question, whether a prior social advice would and particularly of another peer would have another influence on children's and adolescent's EL. Study III solved these questions and investigated how good peer advice and own experience was used for learning in children, adolescents and adults.

Adolescents initially showed greater sensitivity to peer advice and selected the advised deck more often than children and adults. Peer advice affected adolescents' initial expectations (i.e., their priors) most strongly as compared to the other age groups. These results support the view that adolescence may be a developmental period with a particularly high sensitivity to social influence (Blakemore & Mills, 2014; Jones et al., 2014; van Hoorn et al., 2016). Although social information (i.e., in form of single initial advice) was highly influential in adolescents at the beginning of learning, social information was less influential the more own experience they collected - in line with a recent study by Decker et al. (2015). The authors showed that adolescents are highly sensitive to social influence, but also more likely to stop recommended behavior, if they are not positively reinforced. In line with the study by Decker et al. (2015) adolescents and children explored more (i.e., showed higher exploration rates). That is, adolescents increasingly selected the other good deck with learning and chose the bad decks less often (as compared to adults and children). Consequently, and in contrast to adults and children, adolescents selected choices with higher earnings. Although children also showed increased explorative behavior this did not result in choosing the other good deck more often. This might be due to of children's greater sensitivity to negative feedback (as compared to adolescents and adults [van Duijvenvoorde et al., 2008]); children's higher learning rates for losses, as compared to the other two age groups, supports this view. Previous studies suggest that children's performance decreases as the probability of negative feedback increases (Eppinger et al., 2009). Thus, children's difficulties to use negative feedback for learning should be particularly salient in the current task, where each card deck was associated with 50% losses and 50% gains (although they differ in their magnitude; see publication III for further details). Thus, probably due to their higher sensitivity to negative feedback (i.e., higher learning rates) children were not able to benefit from their experience to the same degree as adolescents in this learning environment.

In contrast to children and adolescents, adults followed recommended actions more consistently over time, which was adaptive in our experiment. Decker et al. (2015) suggest that adults will also do so if it would be less adaptive (i.e., if the recommendation was not good). This adult-bias towards prior instructions is linked to the frontal/hippocampal system that trains the reinforcement system; amplifying instruction-consistent feedback and dismissing instruction-inconsistent feedback (Doll, Jacobs, Sanfey, & Frank, 2009). Developmental literature link the protracted maturation of prefrontal cognitive control functions to reduced abilities using rules to guide behavior (Crone et al., 2006; Wendelken, Munakata, Baym, Souza, & Bunge, 2012) and higher exploration rates during development (Decker et al., 2015; Thompson-Schill et al., 2009). Thus, lower levels of cognitive control might have led to a reduced advice-bias and more exploration in children and adolescents as compared to adults. Exploratory behavior can have positive aspects in dynamic and unknown environments, and is suggested to be an important adaptation in human development (Thompson-Schill et al., 2009). Higher exploration in children and adolescents, however, was not linked to similar learning benefits in both age groups. That is, adolescents but not children selected to other good (non-advised) option increasingly and chose more choices with higher earning (as compared to children and adults). To what extend children's behavior might reflect random behavior cannot be drawn from their exploration rate. Recent studies in adults use different computationally approaches to investigate exploration in uncertain learning environments (Speekenbrink & Konstantinidis, 2015) and differentiate between direct (i.e., in which sampling is encouraged by information seeking) and random exploration (i.e., in which sampling is encouraged by chance) (Wilson, Geana, White, Ludvig, & Cohen, 2014). Future studies need to investigate exploration across development.

6.4 Conclusion

In the final section I will highlight main findings and implications of our empirical studies and will discuss them with respect to their limitations, and possible direction of future research.

6.4.1 Social learning across development: Implications of the empirical studies

The empirical studies constituting this thesis show that social learning, across development, is influenced by peers, their behavior and their advice (see Figure 15). Children and particularly adolescents were highly sensitive to social information of peers. That is, it affected children's sensitivity to observed feedback more strongly compared to adults, children imitated choices of peers more frequently than non-peers and adolescents were initially highly sensitive to peer-advice. How other's behavior and recommendations were used for learning was related to specific developmental differences: children benefited from the additional social information as compared to learning without additional social information (EL), but they benefited more slowly from others behavior (action and outcomes) than adults. Our electrophysiological results gave first insights in *why* that might be the case. Children showed a greater sensitivity to negative observed feedback, reflected in enhanced FRN to observed feedback (as compared to adults). Children were less able than adults to disengage and use observed feedback for learning, by showing no learning related changes in their P300 to observed and self-experienced feedback. In contrast, in adults the P300 to observed and experienced-feedback changed with learning and this relationship was moderated by WM. WM has been recently shown to explain learning differences during EL (Collins & Frank, 2012). Our study showed that WM also explained learning differences during OL und suggested that WM might limit children's abilities to use social information for learning. Future studies should address this question more directly.

If behavior was not observable and social information was only given in form of good peer advice, social information had to be evaluated based own experiences. Our findings point to specific developmental differences in using advice and experience for learning. Initially, adolescents were strongly influenced by their peer's advice as compared to children and adults. Adults, however, stayed more consistently with the good advice. Children and adolescents were more exploratory and relied less on the initial advice with more time of learning from own actions. Adolescents benefited from that. They chose the good (non-advised) alternative more and the bad alternative less often (as compared to children and adults). Therefore, they selected choices with higher earnings more than adults and children. Children did not benefit similarly to adolescents from higher exploration. Children showed comparably to our OL-studies a higher sensitivity to negative feedback (i.e., higher learning rate). That might have caused difficulties to distinguish good and bad alternatives comparably to adolescents and adults.

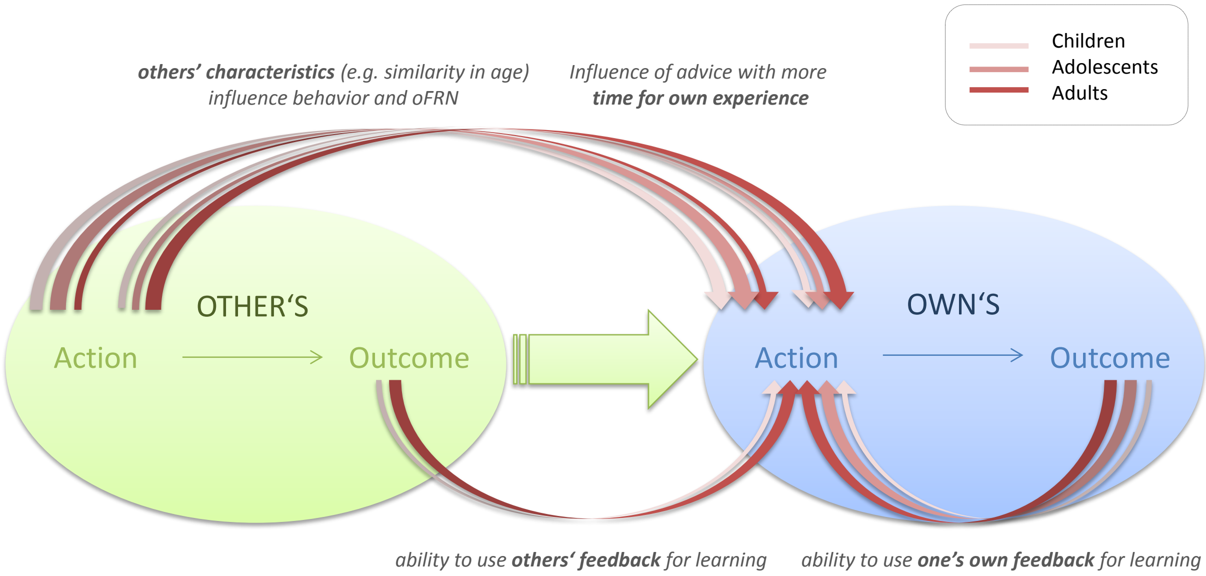

Figure 15. **Framework for developmental differences in social learning.** How others' experience influenced one's own experience is modulated by: time for one's own experience, others' feedback, and others' characteristics. Thickness of arrows represents strength of effect.

Taken together, although children and particularly adolescents are highly sensitive to peer information, they might rely more strongly on what they experience during EL. Children are more sensitive to (negative) experiences and have more difficulties to use them for learning, whereas adolescents and adults are able to. Adults showed that they are more efficient to use other's behavior to update their expectations during learning. When transferring the current findings to an applied context, one tempting interpretation of our results could be that educational interventions programs that are supposed to enhance learning in children should focus on their ability to use experienced feedback for learning and should be *particularly* cautious with using negative social feedback, which might lead to reduced learning. During adolescence positive and negative feedback were integrated more balanced. Although adolescents where initially highly sensitive to other's recommendations, they quickly relied more on their own experiences. Future studies should investigate whether this finding has implications for adolescent instruction-based learning in school.

Although these findings have enhanced the knowledge about developmental differences during social learning, they are also limited in their conclusions, for instance, with respect to educational practice. In the last paragraph, I will describe limitations and future directions studying social learning across development.

6.4.2 Social learning across development: Limitations of the empirical studies

Even though our studies broaden the understanding of developmental differences in social learning, they are also limited regarding (1) the understanding of the underlying neural dynamics in the neural substrates and (2) their broader implications for educational practice.

First, our EEG-studies did not allow us to draw inferences about the neural substrates underlying EL and OL. Recent literature suggests that fronto-striatal areas play a role in the generation of the FRN during learning in adults (Becker, Nitsch, Miltner, & Straube, 2014). That is, the FRN might serve as an index of learning dynamics in this network, in line with recent findings showing that developmental differences in action-outcome processing and RL are due to changes in the fronto-striatal-networks (Hämmerer & Eppinger, 2012; van den Bos, 2009; van den Bos et al., 2012; van Duijvenvoorde et al., 2008). As compared to the FRN, the oFRN is similarly sensitive to gains and losses (Kobza, Thoma, Daum, & Bellebaum, 2011) and shares similar medial prefrontal source activation (Koban et al., 2012). Whether social and individual feedback-processing are linked to similar or two distinct parallel mechanisms (according to Ruff & Fehr, 2014) is still a matter of research. For instance, social and individual feedback-processing are similarly linked to PE signals, but their PE signals are not linked to the same neural substrates (Apps, Lesage, & Ramnani, 2015; Burke et al., 2010; Cooper et al., 2012; Dunne, D'Souza, & O'Doherty, 2016; Hill et al., 2016). Importantly, it has to be considered that the social context further modulates whether action-outcomes of others are perceived as rewarding or aversive (Burke et al., 2010; Koban et al., 2012). These studies show that the oFRN for losses is enhanced for cooperating partners, whereas the oFRN for gains is amplified for competing partners (Itagaki & Katayama, 2008; Marco-Pallarés, Krämer, Strehl, Schröder, & Münte, 2010). In line with this view, Burke and colleagues report that reward prediction error of other's action-outcomes is inversely coded in the ventral striatum. This is supported by previous findings highlighting the role of the ventral striatum in competitive social situations (Fliessbach et al., 2007). Thus, although both processes might be embedded in same or similar neural dynamics, they might be further modulated differently, particularly by the social context. Developmental studies have just started to investigate the influence of the social context on the neural dynamics during social learning.

This should be extended using different methods, such as fMRI, DTI and computational modeling to get a more mechanistic view on the underlying processes involved in developmental changes during social learning. Additionally, using multiple repeated assessments of developmental changes in social learning will give a better approximation of intra-individual changes. This could be for instance important for a better approximation of probable sensitive periods (e.g. for negative social feedback) in different developmental phases and their relation to intra-individual changes in the underlying mechanisms.

Second, in school, children and adolescents learn directly from and in interaction with their teachers or peers. In our laboratory setting, however, children and adolescents learn more indirectly (mainly in the OL-task) from other individuals (young adults and peers) using computer-generated information. In school, social information is mostly given via instruction and advice, but rarely by a contingently observation of other's behavior, where teachers or peers make errors. Thus, a generalization of our findings for educational practice is not given yet. If we want to understand and design programs to enhance learning from, for instance, errors and negative feedback in children, we have to study social learning in classroom-settings with more real-world designs and to manipulate factors that might drive learning in school.

For instance, whereas educational studies suggest that errors are helpful for learning (see Metcalfe, 2016 for review), developmental neuroscience studies suggest that errors and negative feedback are not helpful for learning in children (see Ferdinand & Kray, 2014 for review). Thus, one research question for developmental neuroscience could be, *how* children can benefit from their errors in certain more school-like contexts and whether advice from other's in a constructive way (e.g. due to corrective instead of negative feedback) can help to enhance learning in children. Importantly, developmental neuroscience and educational science can profit from each other: Educational science can provide applicable approaches for learning in school, whereas developmental neuroscience can provide a more mechanistic view of the underlying processes. This exchange will help elucidating questions that arise during educational practice, which developmental neuroscience might answer.

Taken together, social learning has been studied for decades, but it has rarely been studied across development and with respect to its underlying mechanisms. In this thesis, we outlined that others' experience influence one's own experience differently depending on: on the characteristics of the observed other, on the ability to use others' feedback for learning and the ability to use others' advice and one's own experience for learning. How these bits of information are weighted against each other seemed to be scaled by age. This thesis provides preliminary steps in outlining possible mechanisms underlying social learning across development. Future studies are needed to gain a more mechanistic view on developmental differences in social learning. This will help to design interventions, when negative social feedback might lead to mental health issues or if we want to enhance learning in children with learning difficulties.

REFERENCES

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, *14*(10), 1338–1344.

American Electroencephalographic Society. (1994). Guideline thirteen: guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society*, *11*(1), 111–113.

Apps, M. A. J., Lesage, E., & Ramnani, N. (2015). Vicarious Reinforcement Learning Signals When Instructing Others. *Journal of Neuroscience*, *35*(7), 2904–2913.

Bandura, A. (1977). *Social learning theory* (Vol. viii). Oxford, England: Prentice-Hall.

Bandura, A., Grusec, J. E., & Menlove, F. L. (1966). Observational Learning as a Function of Symbolization and Incentive Set. *Child Development*, *37*(3), 499.

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), 129-141.

Becker, M. P. I., Nitsch, A. M., Miltner, W. H. R., & Straube, T. (2014). A Single-Trial Estimation of the Feedback-Related Negativity and Its Relation to BOLD Responses in a Time-Estimation Task. *Journal of Neuroscience*, *34*(8), 3005–3012.

Behrens, T. E. J., Hunt, L. T., & Rushworth, M. F. S. (2009). The Computation of Social Behavior. *Science*, *324*(5931), 1160–1164.

Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2010). It was not MY fault: event-related brain potentials in active and observational learning from feedback. *Cerebral Cortex (New York, N.Y.: 1991)*, *20*(12), 2874–2883.

Berridge, C. W., & Waterhouse, B. D. (2003). The locus coeruleus–noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews*, *42*(1), 33-84.

Biele, G., Rieskamp, J., & Gonzalez, R. (2009). Computational Models for the Combination of Advice and Individual Learning. *Cognitive Science*, *33*(2), 206–242.

Biele, G., Rieskamp, J., Krugel, L. K., & Heekeren, H. R. (2011). The Neural Basis of Following Advice. *PLOS Biol*, *9*(6), e1001089.

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

Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes*, *101*(2), 127–151.

Braaksma, M. A. H., Rijlaarsdam, G., & van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology*, *94*(2), 405–415.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Buchsbaum, D., Bridgers, S., Weisberg, D. S., & Gopnik, A. (2012). The power of possibility: Causal learning, counterfactual reasoning, and pretend play. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2202-2212.

Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010). Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences*, *107*(32), 14431–14436.

Carp, J., Halenar, M. J., Quandt, L. C., Sklar, A., & Compton, R. J. (2009). Perceived similarity and neural mirroring: Evidence from vicarious error processing. *Social Neuroscience*, *4*(1), 85–96.

Chein, J., Albert, D., O'Brien, L., Uckert, K., & Steinberg, L. (2011). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry: Peer influence on risk taking. *Developmental Science*, *14*(2), F1–F10.

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.

Clawson, A., Clayson, P. E., Worsham, W., Johnston, O., South, M., & Larson, M. J. (2014). How about watching others? Observation of error-related feedback by others in autism spectrum disorders. *International Journal of Psychophysiology*, *92*(1), 26–34.

Coates, B., & Hartup, W. W. (1969). Age and verbalization in observational learning. *Developmental Psychology*, *1*(5), 556–562.

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.

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. *European Journal of Neuroscience*, *35*(7), 1024–1035.

Cooper, J. C., Dunne, S., Furey, T., & O'Doherty, J. P. (2012). Human Dorsal Striatum Encodes Prediction Errors during Observational Learning of Instrumental Actions. *Journal of Cognitive Neuroscience*, *24*(1), 106–118.

Crone, E. A., Jennings, J. R., & Van der Molen, M. W. (2004). Developmental change in feedback processing as reflected by phasic heart rate changes. *Developmental psychology*, *40*(6), 1228.

Crone, E. A., Somsen, R. J., Beek, B. V., & Van Der Molen, M. W. (2004). Heart rate and skin conductance analysis of antecendents and consequences of decision making. *Psychophysiology*, *41*(4), 531-540.

Crone, E. A., Donohue, S. E., Honomichl, R., Wendelken, C., & Bunge, S. A. (2006). Brain Regions Mediating Flexible Rule Use during Development. *Journal of Neuroscience*, *26*(43), 11239–11247.

Crone, E. A., Zanolie, K., Van Leijenhorst, L., Westenberg, P. M., & Rombouts, S. A. (2008). Neural mechanisms supporting flexible performance adjustment during development. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(2), 165-177.

Crone, E. A. (2014). The role of the medial frontal cortex in the development of cognitive and social-affective performance monitoring. *Psychophysiology*, *51*(10), 943–950.

de Rover, M., Brown, S. B. R. E., Band, G. P., Giltay, E. J., van Noorden, M. S., van der Wee, N. J. A., & Nieuwenhuis, S. (2015). Beta receptor-mediated modulation of the oddball P3 but not error-related ERP components in humans. *Psychopharmacology*, *232*(17), 3161–3172.

De Taeye, L., Vonck, K., van Bochove, M., Boon, P., Van Roost, D., Mollet, L., … Raedt, R. (2014). The P3 Event-Related Potential is a Biomarker for the Efficacy of Vagus Nerve Stimulation in Patients with Epilepsy. *Neurotherapeutics*, *11*(3), 612–622.

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.

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.

Diederen, K. M., Spencer, T., Vestergaard, M. D., Fletcher, P. C., & Schultz, W. (2016). Adaptive prediction error coding in the human midbrain and striatum facilitates behavioral adaptation and learning efficiency. *Neuron*, *90*(5), 1127-1138.

Doll, B. B., Jacobs, W. J., Sanfey, A. G., & Frank, M. J. (2009). Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. *Brain Research*, *1299*, 74–94.

Donchin, E. (1981). Surprise!… Surprise? *Psychophysiology*, *18*(5), 493–513.

Dunne, S., D'Souza, A., & O'Doherty, J. P. (2016). The involvement of model-based but not model-free learning signals during observational reward learning in the absence of choice. *Journal of Neurophysiology*, *115*(6), 3195–3203.

Eppinger, B., Mock, B., & Kray, J. (2009). Developmental differences in learning and error processing: Evidence from ERPs. *Psychophysiology*, *46*(5), 1043–1053.

Ferdinand, N. K., & Kray, J. (2014). Developmental changes in performance monitoring: How electrophysiological data can enhance our understanding of error and feedback processing in childhood and adolescence. *Behavioural Brain Research*, *263*, 122–132.

Fischer, A. G., & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, *79*(6), 1243–1255.

Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., & Falk, A. (2007). Social Comparison Affects Reward-Related Brain Activity in the Human Ventral Striatum. *Science*, *318*(5854), 1305–1308.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1431), 459-473.

Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology*, *17*(16), R724-R732.

Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual review of psychology*, *63*, 287-313.

Fry, A. F., & Hale, S. (1996). Processing Speed, Working Memory, and Fluid Intelligence: Evidence for a Developmental Cascade. *Psychological Science*, *7*(4), 237–241.

Fukushima, H., & Hiraki, K. (2009). Whose loss is it? Human electrophysiological correlates of non-self reward processing. *Social Neuroscience*, *4*(3), 261–275.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., … Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences*, *101*(21), 8174–8179.

Gopnik, A., Griffiths, T. L., & Lucas, C. G. (2015). When younger learners can be better (or at least more open-minded) than older ones. *Current Directions in Psychological Science*, *24*(2), 87-92.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, *71*(2), 148–154.

Hämmerer, D., & Eppinger, B. (2012). Dopaminergic and prefrontal contributions to reward-based learning and outcome monitoring during child development and aging. *Developmental Psychology*, *48*(3), 862–874.

Hämmerer, D., Li, S.-C., Müller, V., & Lindenberger, U. (2010). Life Span Differences in Electrophysiological Correlates of Monitoring Gains and Losses during Probabilistic Reinforcement Learning. *Journal of Cognitive Neuroscience*, *23*(3), 579–592.

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *Journal of neuroscience*, *28*(22), 5623-5630.

Harper, L. V., & Sanders, K. M. (1975). The effect of adults' eating on young children's acceptance of unfamiliar foods. *Journal of Experimental Child Psychology*, *20*(2), 206–214.

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: new insights into the localization, meaning and network organization. *NeuroImage*, *84*, 159–168.

Hauser, T. U., Iannaccone, R., Walitza, S., Brandeis, D., & Brem, S. (2015). Cognitive flexibility in adolescence: neural and behavioral mechanisms of reward prediction error processing in adaptive decision making during development. *NeuroImage*, *104*, 347–354.

Hendy, H. M., & Raudenbush, B. (2000). Effectiveness of teacher modeling to encourage food acceptance in preschool children. *Appetite*, *34*(1), 61–76.

Heyes, C. M. (1994). Social Learning in Animals: Categories and Mechanisms. *Biological Reviews*, *69*(2), 207–231.

Hill, M. R., Boorman, E. D., & Fried, I. (2016). Observational learning computations in neurons of the human anterior cingulate cortex. *Nature Communications*, *7*, 12722.

Huizenga, H. M., Crone, E. A., & Jansen, B. J. (2007). Decision-making in healthy children, adolescents and adults explained by the use of increasingly complex proportional reasoning rules. *Developmental science*, *10*(6), 814-825.

Itagaki, S., & Katayama, J. (2008). Self-relevant criteria determine the evaluation of outcomes induced by others: *NeuroReport*, *19*(3), 383–387.

Jaswal, V. K., & Neely, L. A. (2006). Adults Don't Always Know Best: Preschoolers Use Past Reliability Over Age When Learning New Words. *Psychological Science*, *17*(9), 757–758.

Joel, D., & Weiner, I. (2000). The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience*, *96*(3), 451-474.

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.

Kang, S. K., Hirsh, J. B., & Chasteen, A. L. (2010). Your mistakes are mine: Self-other overlap predicts neural response to observed errors. *Journal of Experimental Social Psychology*, *46*(1), 229–232.

Karis, D., Fabiani, M., & Donchin, E. (1984). "P300" and memory: Individual differences in the von Restorff effect. *Cognitive Psychology*, *16*(2), 177-216.

Keele, S. W., & Posner, M. I. (1968). PROCESSING OF VISUAL FEEDBACK IN RAPID MOVEMENTS. *Journal of Experimental Psychology*, *77*(1), 155–158.

Klee, M. R., Offenloch, K., & Tigges, J. (1965). Cross-Correlation Analysis of Electroencephalographic Potentials and Slow Membrane Transients. *Science*, *147*(3657), 519–521.

Koban, L., Pourtois, G., Bediou, B., & Vuilleumier, P. (2012). Effects of social context and predictive relevance on action outcome monitoring. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(3), 460–478.

Kobza, S., Thoma, P., Daum, I., & Bellebaum, C. (2011). The feedback-related negativity is modulated by feedback probability in observational learning. *Behavioural Brain Research*, *225*(2), 396–404.

Ladd, G. (1981). Effectiveness of a social learning method for enhancing children's social interaction and peer acceptance. *Child Development*, *52*(1), 171–178.

Lenroot, R. K., & Giedd, J. N. (2006). Brain development in children and adolescents: Insights from anatomical magnetic resonance imaging. *Neuroscience & Biobehavioral Reviews*, *30*(6), 718–729.

Lereya, S. T., Copeland, W. E., Costello, E. J., & Wolke, D. (2015). Adult mental health consequences of peer bullying and maltreatment in childhood: two cohorts in two countries. *The Lancet Psychiatry*, *2*(6), 524–531.

Li, S.-C., Schmiedek, F., Huxhold, O., Röcke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: Practice gain, transfer, and maintenance. *Psychology and Aging*, *23*(4), 731–742.

Lourenco, F. S., Decker, J. H., Pedersen, G. A., Dellarco, D. V., Casey, B. J., & Hartley, C. A. (2015). Consider the source: Adolescents and adults similarly follow older adult advice more than peer advice. *PloS One*, *10*(6), e0128047.

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition*, *131*(2), 284-299.

Marco-Pallarés, J., Krämer, U. M., Strehl, S., Schröder, A., & Münte, T. F. (2010). When decisions of others matter to me: an electrophysiological analysis. *BMC Neuroscience*, *11*(1), 86.

McGuire, J. T., Nassar, M. R., Gold, J. I., & Kable, J. W. (2014). Functionally dissociable influences on learning rate in a dynamic environment. *Neuron*, *84*(4), 870–881.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, *38*(2), 339-346.

Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a New Science of Learning. *Science*, *325*(5938), 284–288.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science (New York, N.Y.)*, *198*(4312), 74–78.

Meltzoff, A. N., Waismeyer, A., & Gopnik, A. (2012). Learning about causes from people: observational causal learning in 24-month-old infants. *Developmental psychology*, *48*(5), 1215.

Menegas, W., Bergan, J. F., Ogawa, S. K., Isogai, Y., Venkataraju, K. U., Osten, P., ... & Watabe-Uchida, M. (2015). Dopamine neurons projecting to the posterior striatum form an anatomically distinct subclass. *Elife*, *4*.

Meshi, D., Biele, G., Korn, C. W., & Heekeren, H. R. (2012). How Expert Advice Influences Decision Making. *PLoS ONE*, *7*(11), e49748.

Metcalfe, J. (2017). Learning from Errors. *Annual Review of Psychology*, *68*(1).

Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-Related Brain Potentials Following Incorrect Feedback in a Time-Estimation Task: Evidence for a “Generic” Neural System for Error Detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798. https://doi.org/10.1162/jocn.1997.9.6.788

Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., … Dalgleish, T. (2009). A Key Role for Similarity in Vicarious Reward. *Science*, *324*(5929), 900–900.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of neuroscience*, *16*(5), 1936-1947.

Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, *431*(7010), 760.

Morgan, T. J. H., Laland, K. N., & Harris, P. L. (2015). The development of adaptive conformity in young children: effects of uncertainty and consensus. *Developmental Science*, *18*(4), 511–524.

Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current directions in psychological science*, *21*(2), 71-77.

Nagell, K., Olguin, R. S., & Tomasello, M. (1993). Processes of social learning in the tool use of chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Journal of Comparative Psychology*, *107*(2), 174–186.

Nassar, M. R., Bruckner, R., Gold, J. I., Li, S.-C., Heekeren, H. R., & Eppinger, B. (2016). Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nature Communications*, *7*, 11609.

Nielsen, M., & Tomaselli, K. (2010). Overimitation in Kalahari Bushman children and the origins of human cultural cognition. *Psychological science*, *21*(5), 729-736.

Nieuwenhuis, S. (2004). Sensitivity of Electrophysiological Activity from Medial Frontal Cortex to Utilitarian and Performance Feedback. *Cerebral Cortex*, *14*(7), 741–747.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus--norepinephrine system. *Psychological Bulletin*, *131*(4), 510–532.

Nieuwenhuis, S., De Geus, E. J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response: P3 and orienting response. *Psychophysiology*, *48*(2), 162–175.

Niv, Y., & Schoenbaum, G. (2008). Dialogues on prediction errors. *Trends in Cognitive Sciences*, *12*(7), 265–272.

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Current Opinion in Neurobiology*, *14*(6), 769–776.

Owens, C. R., & Ascione, F. R. (1991). Effects of the Model's Age, Perceived Similarity, and Familiarity on Children's Donating. *The Journal of Genetic Psychology*, *152*(3), 341–357.

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.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042.

Philiastides, M. G., Biele, G., Vavatzanidis, N., Kazzer, P., & Heekeren, H. R. (2010). Temporal dynamics of prediction error processing during reward-based decision making. *NeuroImage*, *53*(1), 221–232.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148.

Polich, J., Ladish, C., & Burns, T. (1990). Normal variation of P300 in children: Age, memory span, and head size. *International Journal of Psychophysiology*, *9*(3), 237–248.

Rak, N., Bellebaum, C., & Thoma, P. (2013). Empathy and feedback processing in active and observational learning. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(4), 869–884.

Rakoczy, H., Hamann, K., Warneken, F., & Tomasello, M. (2010). Bigger knows better: Young children selectively learn rule games from adults rather than from peers. *British Journal of Developmental Psychology*, *28*(4), 785–798.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature reviews neuroscience*, *9*(7), 545.

Ridderinkhof, K. R. (2004). The Role of the Medial Frontal Cortex in Cognitive Control. *Science*, *306*(5695), 443–447.

Rodriguez Buritica, J. M., Eppinger, B., Schuck, N. W., Heekeren, H. R., & Li, S.-C. (2016). Electrophysiological correlates of observational learning in children. *Developmental Science*, *19*(5), 699–709.

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, *15*(8), 549–562.

Rushworth, M. F., & Behrens, T. E. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature neuroscience*, *11*(4), 389.

Sadacca, B. F., Jones, J. L., & Schoenbaum, G. (2016). Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. *Elife*, *5*.

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235.

San Martín, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, *6*.

Santesso, D. L., Dzyundzyak, A., & Segalowitz, S. J. (2011). Age, sex and individual differences in punishment sensitivity: Factors influencing the feedback-related negativity. *Psychophysiology*, *48*(11), 1481–1489.

Schunk, D. H. (1987). Peer Models and Children's Behavioral Change. *Review of Educational Research*, *57*(2), 149–174.

Schunk, D. H., & Usher, E. L. (2012). Social Cognitive Theory and Motivation. Retrieved from http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780195399820.001.0001/oxfordhb-9780195399820-e-2

Silva, K., Shulman, E. P., Chein, J., & Steinberg, L. (2015). Peers Increase Late Adolescents' Exploratory Behavior and Sensitivity to Positive and Negative Feedback. *Journal of Research on Adolescence*, *26*(4), 696-705.

Simons-Morton, B., Lerner, N., & Singer, J. (2005). The observed effects of teenage passengers on the risky driving behavior of teenage drivers. *Accident Analysis & Prevention*, *37*(6), 973–982.

Smith, A. R., Chein, J., & Steinberg, L. (2014). Peers increase adolescent risk taking even when the probabilities of negative outcomes are known. *Developmental Psychology*, *50*(5), 1564–1568.

Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and Exploration in a Restless Bandit Problem. *Topics in Cognitive Science*, *7*(2), 351–367.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, *28*(1), 78–106.

Sutton, R. S., & Barto, A. G. (1998). *Introduction to Reinforcement Learning* (1st ed.). Cambridge, MA, USA: MIT Press.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-Potential Correlates of Stimulus Uncertainty. *Science*, *150*(3700), 1187–1188.

Takizawa, R., Maughan, B., & Arseneault, L. (2014). Adult Health Outcomes of Childhood Bullying Victimization: Evidence From a Five-Decade Longitudinal British Birth Cohort. *American Journal of Psychiatry*, *171*(7), 777–784.

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition Without Control When a Little Frontal Lobe Goes a Long Way. *Current Directions in Psychological Science*, *18*(5), 259–263.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(05).

Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends in Cognitive Sciences*, *18*(5), 259–267.

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.

van den Bos, W., Cohen, M. X., Kahnt, T., & Crone, E. A. (2012). Striatum–Medial Prefrontal Cortex Connectivity Predicts Developmental Changes in Reinforcement Learning. *Cerebral Cortex*, *22*(6), 1247–1255.

van Dinteren, R., Arns, M., Jongsma, M. L. A., & Kessels, R. P. C. (2014). P300 Development across the Lifespan: A Systematic Review and Meta-Analysis. *PLoS ONE*, *9*(2), e87347.

van Duijvenvoorde, A. C. K. van, 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. *The Journal of Neuroscience*, *28*(38), 9495–9503.

van Gog, T., & Rummel, N. (2010). Example-Based Learning: Integrating Cognitive and Social-Cognitive Research Perspectives. *Educational Psychology Review*, *22*(2), 155–174.

van Hoorn, J., van Dijk, E., Meuwese, R., Rieffe, C., & Crone, E. A. (2016). Peer Influence on Prosocial Behavior in Adolescence. *Journal of Research on Adolescence*, *26*(1), 90–100.

VanderBorght, M., & Jaswal, V. K. (2009). Who knows best? Preschoolers sometimes prefer child informants over adult informants. *Infant and Child Development*, *18*(1), 61–71.

von Borries, A. K. L., Verkes, R. J., Bulten, B. H., Cools, R., & de Bruijn, E. R. A. (2013). Feedback-related negativity codes outcome valence, but not outcome expectancy, during reversal learning. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(4), 737–746.

Watabe-Uchida, M., Zhu, L., Ogawa, S. K., Vamanrao, A., & Uchida, N. (2012). Whole-brain mapping of direct inputs to midbrain dopamine neurons. *Neuron*, *74*(5), 858-873.

Wendelken, C., Munakata, Y., Baym, C., Souza, M., & Bunge, S. A. (2012). Flexible rule use: Common neural substrates in children and adults. *Developmental Cognitive Neuroscience*, *2*(3), 329–339.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, *143*(6), 2074–2081.

Windell, O. (2008). Vector Snowflake Application. Retrieved September 7, 2012, from http://www.evilmadscientist.com/2008/vector-snowflake-application/.

Yu, R., & Zhou, X. (2006). Brain responses to outcomes of one's own and other's performance in a gambling task. *Neuroreport*, *17*(16), 1747–1751.

Zimmerman, B. J., & Rosenthal, T. L. (1974). Observational learning of rule-governed behavior by children. *Psychological Bulletin*, *81*(1), 29–42.

Zmyj, N., & Seehagen, S. (2013). The Role of a Model's Age for Young Children's Imitation: A Research Review: The Role of a Model's Age for Young Children's Imitation. *Infant and Child Development*, *22*(6), 622–641.

Zottoli, T. M., & Grose-Fifer, J. (2012). The feedback-related negativity (FRN) in adolescents. *Psychophysiology*, *49*(3), 413-420.

SUPPLEMENTS

I EIDESSTATTLICHE ERKLÄRUNG

Hiermit erkläre ich an Eides statt,

- dass ich die vorliegende Arbeit selbstständig und ohne unerlaubte Hilfe verfasst habe,
- dass ich mich nicht bereits anderwärts um einen Doktorgrad beworben habe und keinen Doktorgrad in dem Promotionsfach Psychologie besitze und
- dass ich die zugrunde liegende Promotionsordnung vom 08.08.2016 kenne.

Berlin, 28. November 2016

Julia M. Rodriguez Buritica

II LIST OF PUBLICATIONS

The Thesis is based on the following articles:

Rodriguez Buritica, J. M., Eppinger, B., Schuck, N. W., Heekeren, H. R., & Li, S.-C. (2016). Electrophysiological correlates of observational learning in children. *Developmental Science*, *19*(5), 699–709. https://doi.org/10.1111/desc.12317

Rodriguez Buritica, J. M., Heekeren, H. R., Li, S.-C. & Eppinger, B. (under revision). Developmental differences in the neural dynamics of observational learning.

Rodriguez Buritica, J. M., Heekeren, H. R., & van den Bos, W. (under revision). Adolescents are sensitive to peer influence, but only for so long.

III ORIGINAL PUBLICATIONS

Title: Developmental differences in the neural dynamics of observational learning

Running title: Age differences in vicarious learning

Julia M. Rodriguez Buritica[1,2,3], Hauke R. Heekeren[1,3], Shu-Chen Li[2,3,4], and Ben Eppinger[4]

[1]Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany

[2]Center of Lifespan Psychology, Max-Planck-Institute for Human Development, Berlin, Germany

[3]Berlin School of Mind and Brain, Berlin, Germany

[4]Chair of Lifespan Developmental Neuroscience, Department of Psychology, TU Dresden, Dresden, Germany

Address for correspondence:

Julia M. Rodriguez Buritica, buritica@zedat.fu-berlin.de, phone: +49 (0) 30 838 – 71781

Department of Education Sciences and Psychology, Freie Universität, Habelschwerdter Allee 45, 14195 Berlin, Germany

Abstract

Learning from vicarious experience is central for educational practice, but not well understood with respect to its ontogenetic development and underlying neural dynamics. In this age-comparative study we compared behavioral and electrophysiological markers of learning from vicarious and one's own experience in children (age 8-10) and young adults. Behaviorally both groups benefitted from integrating vicarious experience into their own choices however, adults learned much faster from social information than children. The electrophysiological results show learning-related changes in the P300 to experienced and observed rewards in adults, but not in children, indicating that adults were more efficient in integrating observed and experienced information during learning. In comparison to adults, children showed an enhanced FRN for observed and experienced feedback, indicating that they focus more on valence information than adults. Taken together, children compared to adults seem to be less able to rapidly assess the informational value of observed and experienced feedback during learning and consequently up-regulate their response to both, observed and experienced (particularly negative) feedback. When transferring the current findings to an applied context, educational practice should strengthen children's ability to use feedback information for learning and be particularly cautious with negative social feedback during supervised learning.

Keywords: development, experience-based learning, observational learning, FRN/P300 & observational FRN/P300

Highlights

- Adults benefit more from observed information during learning than children.
- ERP: Children have problems in disengaging from observed and experienced feedback.
- ERP: Adults are better in integrating observed and experienced information.
- Educational practice should enhance children's ability to use feedback for learning.
- Educational practice should be cautious with the negative social feedback

1. Introduction

Learning through observation is a prerequisite for the acquisition of new behavior (Rendell et al., 2011) and central for educational practice (Groenendijk, 2013). Observational learning is particularly of interest from a developmental perspective because it may serve as an important mechanism for human cognitive and social-emotional development (Marshall et al., 2011; Meltzoff et al., 2012; Nielsen and Tomaselli, 2010). Compared to learning from own experience (Bos et al., 2012; Duijvenvoorde et al., 2008; Eppinger et al., 2009; Hämmerer et al., 2010), however, the underlying neural dynamics and their ontogenetic development are far less well understood. In the current study we therefore investigated developmental differences in experience-based and observational learning using an electrophysiological approach.

Developmental studies on experience-based learning suggest that children in comparison to adults seem to have greater difficulties in extracting and using relevant feedback information for learning (Bos et al., 2012; Crone et al., 2004; Duijvenvoorde et al., 2008; Eppinger et al., 2009; Hämmerer et al., 2010). Consistent with these findings, electrophysiological studies in children point to a reduced ability to disengage from negative feedback during learning as reflected in an enhanced medial prefrontal ERP component, the feedback-related negativity (FRN) (Eppinger et al., 2009; Hämmerer et al., 2010). Interestingly, the reverse pattern is observed for the later parietal P300 component, which has been associated with the updating of working memory (WM) representations (Morgan et al., 2008). The parietal P300 component tends to be reduced in children compared to adults, which has been interpreted in terms of developmental differences in working memory capacities (Polich et al., 1990)

In adults, the two ERP components described in experience-based learning (FRN and P300) can also be observed during vicarious learning (referred to as observational FRN (oFRN) and observational P300 (oP300), respectively; (Bellebaum et al., 2010; Rodriguez Buritica et al., 2016; Yu and Zhou, 2006). Developmental differences in these ERP components during observational learning have, however, just started to be investigated (Rodriguez Buritica et al., 2016).

Here we examined whether children and younger adults differ in their abilities to learn from experienced and observed feedback information and whether these developmental differences in learning are associated with separable underlying neurophysiological mechanisms. To address these questions we used a probabilistic reward- based observational learning paradigm (Burke et al., 2010; Rodriguez Buritica et al., 2016) in combination with EEG in 8-10 year-old children and 20-30 year-old adults. The paradigm consists of two observational and an individual learning condition (see Figure 1B). In the two observational conditions we varied the amount of observable information: in the "action only" (A) condition, only the actions of the other player were observable; whereas, in the "action + outcome" (AO) condition, both the actions *and* the outcomes of the other player were presented. In the individual learning condition neither the actions nor the outcomes of the other player were observable (see Figure 1B).

Based on prior work (Burke et al., 2010; Rodriguez Buritica et al., 2016), we predicted that learning performance should scale with the amount of observable information and that children should be impaired in learning compared to adults. Furthermore, we expected that learning impairments in children should be associated with a greater sensitivity to observed as well as experienced negative feedback, as reflected in enhanced medial prefrontal (FRN) activity (Ferdinand and Kray, 2014). Based on our own prior work, we predict that in children medial prefrontal brain responses to external feedback information should not be modulated by its relationship to one's own action (Rodriguez Buritica et al., 2016). In contrast, in adults medial prefrontal brain responses to external feedback information should vary with its relationship to one's own action (Bellebaum et al., 2010; Yu and Zhou, 2006). Thus, the FRN and oFRN should not differ in amplitude in children, whereas in adults the oFRN should be reduced compared to the FRN.

Based on previous results, suggesting that reward-based learning depends on WM abilities (Collins and Frank, 2012) and given several findings indicating that developmental differences in WM updating are reflected in the parietal P300 components (Polich et al., 1990; van Dinteren et al., 2014), we expected reduced P300 components during the processing observed and experienced feedback during learning in children compared to adults.

2. Methods

2.1. Participants

The effective sample of the study consisted of 23 adults between 20-30 years of age (11 female, mean age = 23.52, SD = 2.81) and 22 children between 8-10 years of age (10 female, mean age = 9.05, SD = .79). Data of one child were excluded from further analyses, due to technical problems during data collection. Data from another child had to be excluded because the individual ERP components of interest differed more than three standard deviations from the group mean. All participants were right-handed (Oldfield, 1971), had normal or corrected-to-normal vision and no neurological or psychological disorders. Prior to the experiment we obtained informed consent from the participants and their parents (in case of children). The study was approved by the Ethics Committee of the Max-Planck-Institute for Human Development, Berlin. Subjects participated in two experimental sessions: a behavioral group session (together with same-aged participants) in which we assessed psychometric covariate measures and an individual EEG session in which we assessed observational learning performance. The participants received a compensation of 14 Euro for the first and 24 Euro for the second session. General cognitive abilities of the sample were assessed using several psychometric tests: *I.* Identical pictures test (Ekstrom et al., 1976) as a marker for cognitive speed; *II.* Raven's Progressive matrices (Raven and Court, 1998) for adults and the Colored Progressive Matrices (CPM; Raven and Raven, 2002) for children as a measure of fluid intelligence; *III.* a modified version of the Spot-the-Word test (Lehrl, 2005; Lindenberger et al., 1993) as a marker for verbal knowledge; *IV.* a modified version of the spatial n-back task described in detail by Li et al. (2008) to investigate working memory (WM) capacity. Children had lower scores than adults on the Identical Pictures (children: M = 20.32 (SD = 2.95), adults: M = 32.83 (SD = 3.63); $F(1, 43) = 160.24$, $p < .001$, $\eta_p^2 = .79$), Spot-the-Word test (children: M = 2.09 (SD = 1.49), adults: M = 17.57 (SD = 6.77); $F(1, 43) =$ 100.55, $p < .001$, $\eta_p^2 = .7$) and in proportion correct on the WM test (children: M = .64 (SD = .21), adults: M = .84 (SD = .08); $F(1, 43) = 18.47$, $p < .001$, $\eta_p^2 = .3$). These age differences are consistent with previous findings from larger population-based lifespan samples (Li et al., 2004). One-way analysis of variance (ANOVA) yielded no age differences in the normalized IQ scores, $F(1, 43) = .72$, $p = .4$, $\eta_p^2 = .02$ (children: M = 63.23 (SD = 30.53), adults: M = 70.61 (SD = 27.71)).

2.2. Experimental Design and Procedure

2.2.1. Design. The task involved three learning conditions: Individual Learning (*IL*), Action Only (*A*), and Action and Outcome (*AO*). As shown in Figure 1A, participants were asked to choose one of two abstract stimuli (generated with Vector Snowflake Application; (Windell, 2008) that may result in a positive or negative feedback. Within each stimulus pair, one stimulus was associated with a high probability (80% rewards, 20% losses) and one associated with a low probability (20% rewards, 80% losses) of gaining points. In the two observational learning conditions (*A* and *AO*), prior to making their own choices, the computer program presented the participants a picture of the face of a randomly chosen sex- and age-matched "model player" (i.e., another participant who took part in the same first group session of cognitive covariate assessments together with the to be tested participant). The participants were told that the other player had already performed the task and that they would observe recorded choices of this other player. In fact, however, the to be observed behavior was computer-generated (see Figure 1C; see Appendix A for further details). The participants were debriefed about the cover story after the experiment.

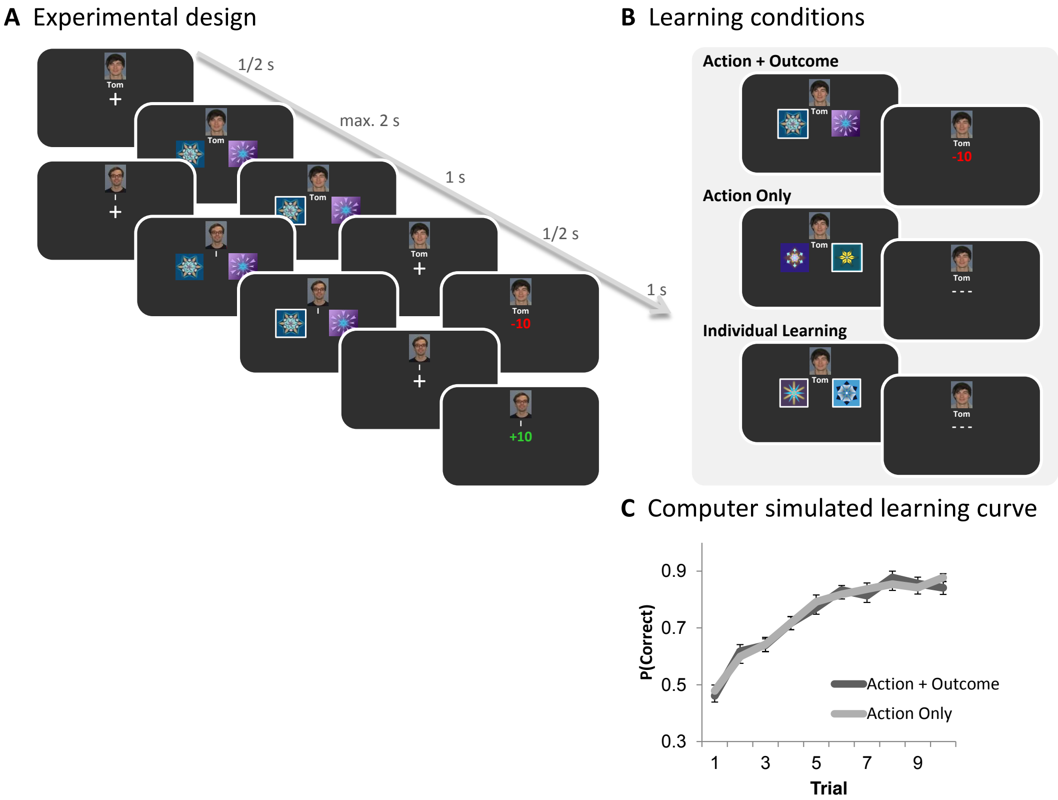

Figure 1. **Design.** *(A) Trial procedure. (B) Learning conditions.* 1: Action + Outcome (AO), 2: Action Only (A), 3: Individual Learning (IL). *(C)* Computer simulated averaged learning curve for the two observational conditions.

2.2.2. Trial procedure: As shown in Figure 1A, the participants first saw the picture and the name of the model player for 500 ms. They were told that if they pressed the response key within 2 seconds they could see the choices of the model player (to ensure that they paid attention during the task). Then the model player's choice was indicated using a colored frame (1 sec), which was followed by a fixation cross for 500 ms and the outcome (reward / loss of 10 Points) for 1 sec. The position of the stimuli was randomized across and within trials and blocks. The amount of information about the model behavior differed between the three different learning conditions: In the *AO* condition, full information about the choices and outcomes of the model players was provided (see Figure 1B). In the *A* condition, information about the choices of the model players was shown but not information about the associated choice outcomes (see Figure 1B). In the *IL* condition, no information about the model player's choices and the outcomes was provided (see Figure 1B). At the end of each of the learning conditions was an action stage indicated by the picture of the participant (displayed for 500 ms) for the participant to take his or her own actions (within 2 sec) about the same pair of stimuli as in the prior learning condition. The timing of the action stage was identical to the prior learning condition. Each block included 10 trials per learning condition. Each condition was assigned to one stimulus pair (3 different pairs per block). The orders of the learning conditions were pseudo-randomized over each of the 12 blocks.

2.3. Electrophysiological recording

While the participants performed the task (controlled by using the psychtoolbox for Matlab; psychtoolbox, Brainard, 1997) EEG was recorded continuously (Brain Amp DC, BrainVision Recorder software) from 64 Ag/AgCl electrodes (American Electroencephalographic Society, 1994) in an elastic cap (Braincap, BrainVision). The sampling rate was 1000 Hz, with a bandpass filter (0.01 to 100 Hz) applied. EEG recordings were referenced online to the right mastoid. Vertical and horizontal eye movements were recorded from electrodes placed next to each eye and below the eye. Impedances were kept below 5 kΩ.

2.4. Data analysis

2.4.1. Behavioral Data. Responses faster than 100 ms (children: 4.22 %, adults: .09%) and exceeding the response deadlines (2000 ms) in the action stage (children: 5.31 %, adults: 1.38 %) were excluded from further analyses.

2.4.2. EEG Data. The recorded data were re-referenced offline to averaged mastoids and further analyzed using BrainVision Analyzer software (Brain Products, Germany). The data were bandpass-filtered in a range of .01 to 20 Hz (according to suggestions of Luck, 2012) and segmented into epochs (200 to 700 ms) after feedback onset. Ocular artifacts were removed using a linear regression approach (Gratton et al., 1983). Additional artifacts were rejected based on a maximum admissible voltage step (50 µV), and by a maximum admissible difference between 2 values on a segment (200 µV). The data were baseline corrected (200 ms pre-stimulus). For eight participants, the data from one to seven malfunctioning electrodes (AF7, AF8, FP1, FP2, FPz, FP2, F8, FT8, P1, P2, PO7, O1, O2, Oz) were replaced via spherical spline interpolation (Perrin et al., 1989). ERPs were averaged for each condition and each participant first, and then across participants. The FRN was determined in all three learning conditions. The oFRN was measured in the *AO* condition.

As shown in Figure 3A & 6A FRN and oFRN peak latencies differed between age groups. To evaluate these effects we determined the peak latencies of FRN and oFRN for each individual and valence (reward, loss). We then used mixed-effects analysis of variances (ANOVA) with the between-subject factor age group (adults, children), and the within-subject factors valence (reward, loss) to compare FRN and oFRN latencies statistically. The analyses revealed main effects of age group (p's < .001, η_p^2's >.3) for both components, but no main effects of valence (p's > .1). To account for these group differences in peak latencies we measured and analyzed the FRN and oFRN as mean amplitudes in 50 ms time windows centered on the peaks of the components (Reinhart and Woodman, 2014) at electrode FCz (adults: FRN: 248 to 298 ms & oFRN: 275 to 325 ms; children: FRN: 317 to 367 ms & oFRN: 327 to 377 ms).

Similar to the FRN/oFRN, visual inspection of the EEG waveforms of the P300 and oP300 (see Figure 3B & 6B) suggested differences in peak latencies between groups and additionally between valences. To verify this we performed a latency analysis on the individual peaks latencies of P300 and oP300 using a mixed-effects analysis of variances (ANOVA) with the between-subject factor age group (adults, children), and the within-subject factors valence (reward, loss). The analyses showed main effects of age (p's < .001, η_p^2's >.3), and of valence (p's < .01, η_p^2's >.2) for both components. To account for the significantly different peaks for rewards and losses in the P300 and oP300 we adjusted the time windows accordingly. The components were measured and analyzed as mean amplitudes within an age -group and valence-specific 100 ms time window centered on the peaks of the components at electrode Pz (adults: P300: rewards: 309 to 409 ms & losses: 325 to 425 ms; oP300: rewards 313 to 413 & losses 366 to 466 ms; children: P300: rewards: 443 to 543 ms & losses 472 to 572 ms; oP300: rewards: 484 to 584 ms & losses 472 to 572 ms) (see Figure 3B und 6B). Difference waves were calculated by subtracting the ERPs following gains from those following losses.

2.4.3. Statistical analyses. Statistical analyses were performed using SPSS (SPSS Inc., Chicago, IL). Accuracy (proportion correct) was averaged into two block halves (i.e., trials 1-5 versus trials 6-10 for each learning condition, collapsed across 12 blocks) and was analyzed using a mixed-effects analysis of variances (ANOVA) with the between-subject factor age group (adults, children), and the within-subject factors learning condition [Action + Outcome

(AO), Action Only (A), Individual Learning (IL)] and block half [first (BH1), second (BH2)]. The accuracy data involved an average trial number of $M = 54.16$, $SD = 3.53$ trials per block half in children and $M = 59.13$, $SD = 1.88$ in adults.

ERPs to experienced feedback (FRN, P3) were analyzed using a mixed-effects ANOVA with the between-subject factor age group (adults, children), and the within-subject factors learning condition (AO, A, IL), valence (reward, loss), and block half (first, second). ERPs with respect to observed feedback (AO condition) were investigated using a mixed-effects ANOVA with the between-subjects factor age group (adults, children), and the within-subjects factors valence (rewards, loss), block half (first, second). The ANOVA comparison between experienced and observed outcomes involved the additional factor agency (FRN, oFRN). For this comparison the FRN was averaged across the three learning conditions. Given that we were interested in valence-dependent learning effects in the ERPs rather than overall amplitude changes with learning, we focused the analyses of ERP components on interactions including the factor valence. To understand the resulting interactions separate ANOVAs and paired samples *t*-tests were conducted. Effect sizes (partial eta squared, η_p^2) are reported, and Pearson's *r* was computed for correlation analysis. The Greenhouse-Geisser correction for non-sphericity and Bonferroni-corrections were applied when necessary (Geisser and Greenhouse, 1958) and the corrected *p*-values are reported.

3. Results

3.1. Behavioral results

3.1.1. Learning Effects. Both age groups benefited (in terms of accuracy) from integrating the observed information into their own choices (learning condition: $F(2, 86) = 27.09$, $p < .001$, $\eta_p^2 = .39$). As shown in Figure 2, accuracy was greater in the AO than the other learning conditions (two-tailed *t* test, *p's* <. 001). However, accuracy did not differ significantly between the A and IL conditions ($p > .1$).

Adults performed significantly better than children (age group: $F(1, 43) = 28.94$, $p < .001$, $\eta_p^2 = .40$) and, across conditions, showed greater learning effects (age group x block half: $F(1, 43) = 15.42$, $p < .001$, $\eta_p^2 = .26$). Most interestingly, we found a significant three-way interaction between age group, learning condition and block half ($F(2, 86) = 8.18$, $p < .001$, $\eta_p^2 = 0.16$). Separate analyses for each age group showed that learning effects differed across conditions in adults, ($p = .001$, $\eta_p^2 = 0.27$), but not in children, ($p = .14$). As displayed in Figure 2 younger adults showed faster learning in the AO compared to the other two learning conditions. In contrast, children learned more gradually in the AO condition.

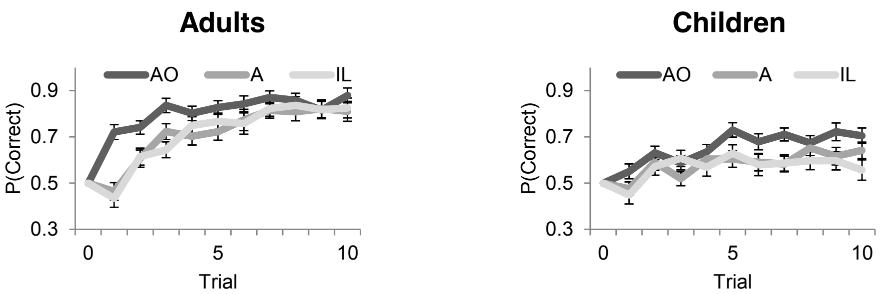

Figure 2. **Behavior.** *Learning & Condition Effects.* Accuracy in proportion correct separately for age group and learning condition displayed per trial.

3.2. ERPs to experienced feedback

3.2.1. FRN. As shown in Figure 3A, across age groups, the FRN was significantly larger (more negative going) for losses than for rewards (valence: $F(1, 43) = 22.45$, $p < .001$, $\eta_p^2 = .34$), and across conditions, FRN amplitudes were larger for children than adults (age group: $F(1, 43) = 43.86$, $p < .001$, $\eta_p^2 = .51$).

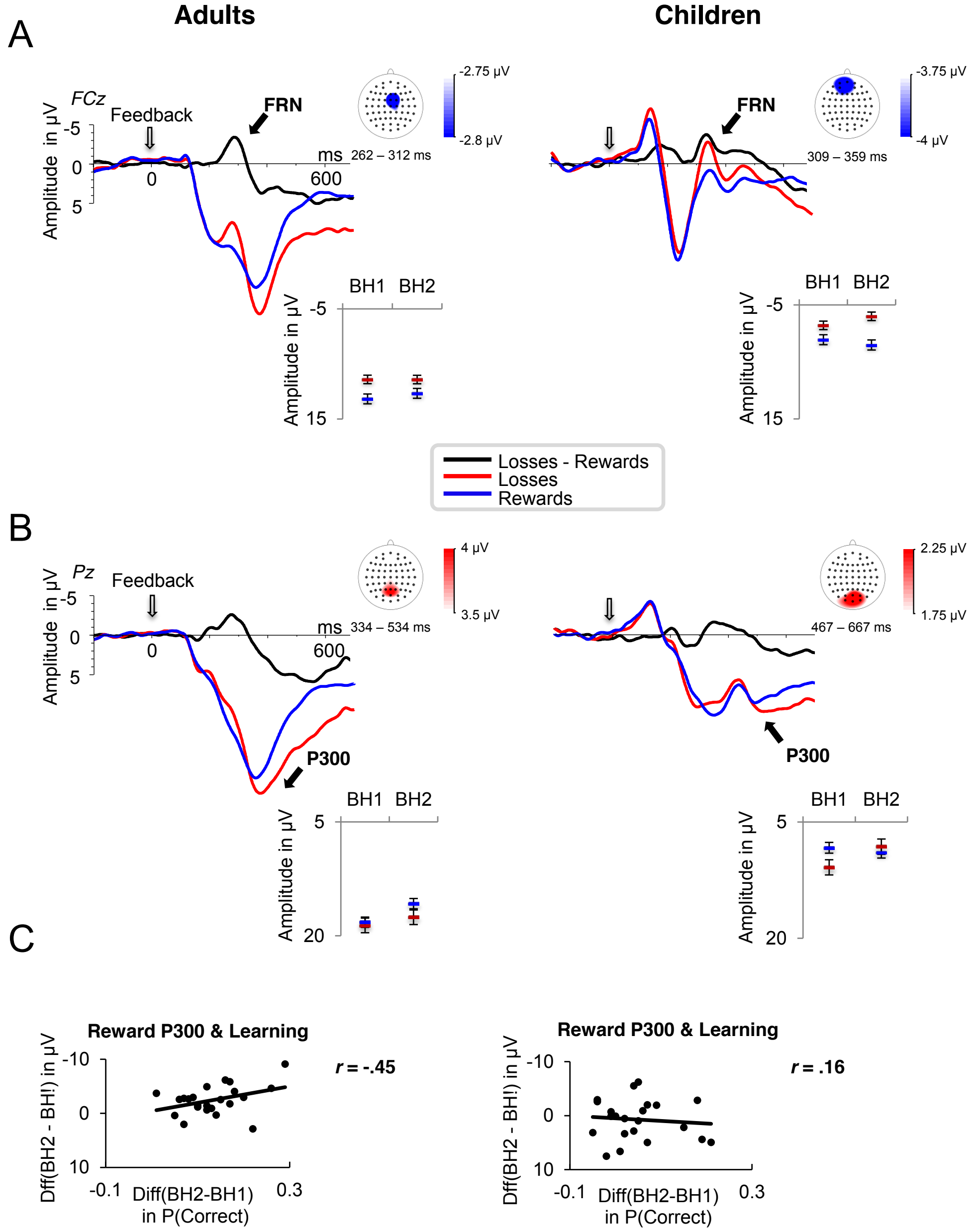

Figure 3. **ERPs experienced feedback**. Grand averages shown for losses (red line) and rewards (blue line) for the *(A) FRN* displayed at FCz and the *(B) P300* displayed at Pz separately for both age groups, as well as the learning related changes for losses and rewards separately for both block halves (BH1 and BH2). The topographic map displays the difference (black line) between losses and rewards for the FRN (within 50 ms) and P300 (within 200 ms). *(C) Correlation effects.* Scatter plots illustrating the correlation between difference score of proportion of correct choice (second – first block half) on the x-axis and the difference score of the mean P300 amplitude to rewards (second – first block half) on the y-axis separately for the two age groups.

3.2.2. Feedback-related P300. Across age groups, the analysis revealed main effects of learning condition $F(2, 86) = 5.08$, $p = .008$, $\eta_p^2 = .11$ and valence: $F(1, 43) = 9.25$, $p = .004$, $\eta_p^2 = .18$ as well as a significant learning condition by valence interaction in the P300, $F(2, 86) = 7.55$, $p < .001$, $\eta_p^2 = .15$. Separate analyses for reward and loss trials showed a condition effect for both trial types ($p = .002$ $\varepsilon = .93$, $\eta_p^2 = .13$, for reward trials and $p = .006$, $\eta_p^2 = .11$, for loss trials). T-tests between conditions that were performed separately for the two valences show that the P300 to rewards decreases with increasing amount of observable information (two-tailed *t* test, AO – A + IL: *p* =. 004; A – IL: *p* =. 067), whereas as the P300 response to losses increases, two-tailed *t* test, AO – A + IL: *p* =. 045; A – IL: *p* =. 017. Thus, the P300 difference between losses and rewards increased with the amount of information that could be used for learning (see Figure 4A).

As shown in Figure 3B, the P300 was larger for adults compared to children (age group: $F(1, 43) = 23.54$, $p < 0.001$, $\eta_p^2 = .35$) and for losses compared to rewards (valence: $F(1, 43) = 9.25$, $p = .004$, $\eta_p^2 = .18$). Moreover, as for performance, we obtained a significant interaction between the factors age group, valence, and block half ($F(1, 43) = 7.82$, $p = .008$, $\eta_p^2 = .15$). Separate analyses for the factor valence revealed a significant interaction between block half and age group, ($p = .002$, $\eta_p^2 = .2$) for rewards, but not for losses ($p = .37$). As shown in Figure 3B, comparisons between block halves revealed a decrease in the P300 to rewards as a function of learning in adults (two-tailed *t* test, $p < .001$), but not in children ($p = .37$).

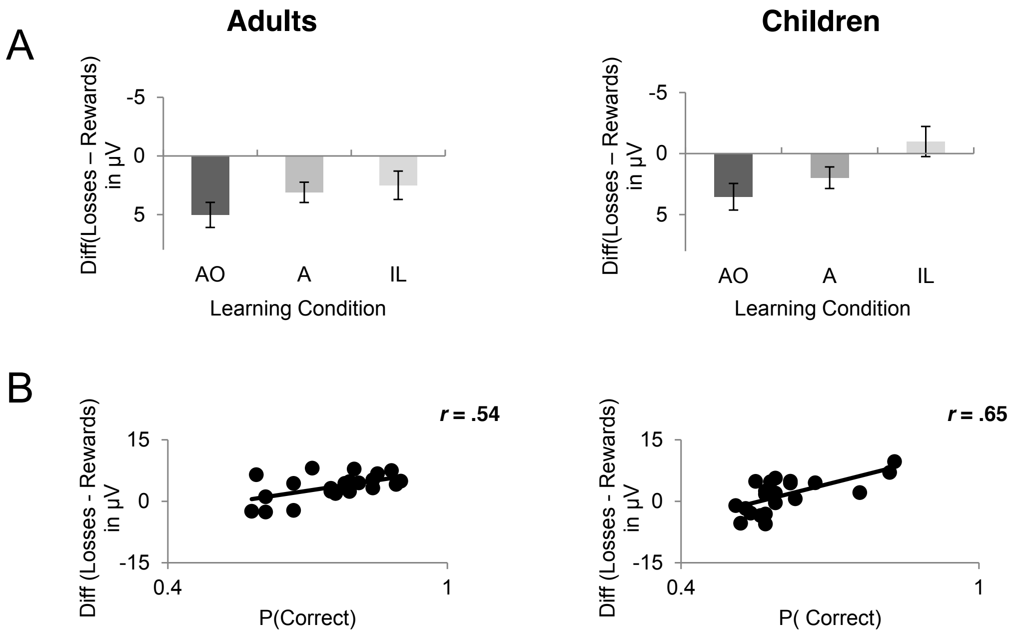

Figure 4. **P300 valence effects**. *(A)* Amplitude difference for the P300 between losses and rewards separately for age group and learning condition. *(B) Correlation effects.* Scatter plots illustrating the correlation of proportion of correct choice on the x-axis and the mean P300 amplitude difference between losses and rewards on the y-axis separately for the two age groups.

3.2.3.Correlation analyses. As shown in Figure 4B, in adults the P300 difference between losses and rewards correlated positively with accuracy in both age groups (adults: $r(23) = .43$, $p = .04$, children: $r(22) = .6$, $p = .003$). Thus, the more participants differentiate between outcomes in the P300 the better their performance. No such effects were observed for the FRN (p's > .3). This result suggests that the P300 but not the FRN reflects the degree to which individuals update outcome predictions (see Figure 4B).

Moreover, in adults learning effects in the P300 to rewards (amplitude difference between block halves) were positively correlated with behavioral learning effects (accuracy difference between block halves) ($r(23) = -.45$, $p = .03$) (see Figure 3C) and WM ($r(23) =$ -.42, $p = .045$) (Figure 5A). In children the correlations were not significant, neither for behavioral learning effects ($r(22) = .16$, $p = .47$) nor for WM ($r(22) = .23$, $p = .31$). Fisher's Z test showed that the correlation coefficients between both age groups were significantly different from each other (two-tailed test for behavioral learning: $z = -1.91$, $p = .056$ & WM: $z = -2.22$, $p = .03$).

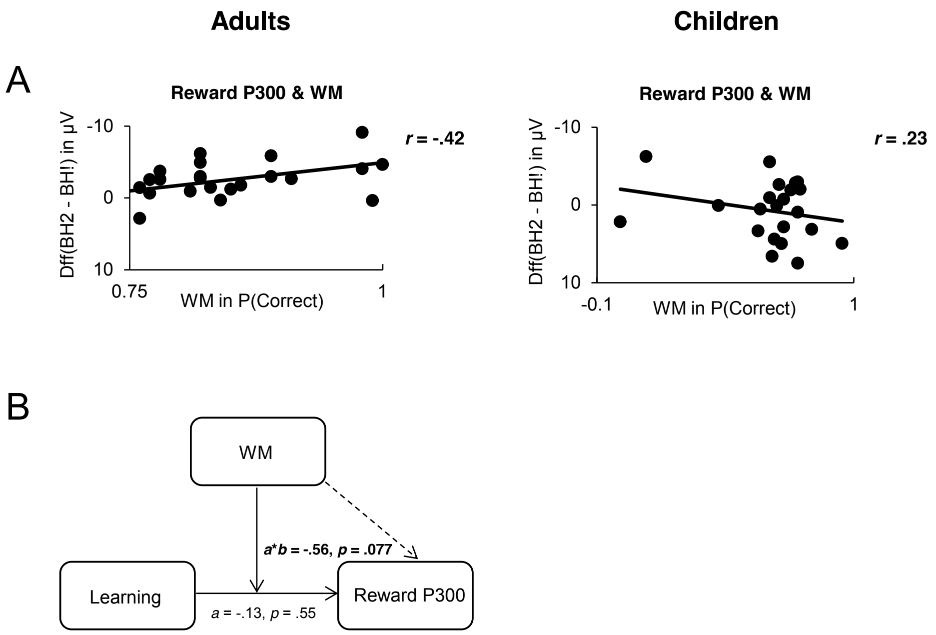

Figure 5. **P300 learning effects**. *(A) Correlation effects.* Scatter plots illustrating the correlation between difference score of proportion correct in the WM task on the x-axis and the difference score of the mean P300 amplitude to rewards (second – first block half) on the y-axis separately for the two age groups. (*B) Moderation model of P300 learning effects*. Working memory (WM) as a moderator for the relationship between learning and learning related changes in the P300 to rewards: *a***b* path predicted learning related changes in the P300 to rewards (neither the *a* path (learning as predictor) nor the *b* path (WM as predictor) reached significance).

3.2.4. Moderator analysis. To investigate whether WM might moderate the relationship between learning and the P300 (Collins and Frank, 2012) we performed a linear regression separately for each age group. In this analysis learning effects in the P300 to rewards were the dependent variable, and behavioral learning and WM the independent variables. The results of this analysis showed that in adults neither behavioral learning nor WM separately predicted changes in the P300 to rewards (p's > .3). However, for the interaction between learning and WM we found a marginally significant moderation effect (learning * WM: $\beta = -.56$, $p = .077$; linear regression model: $R^2 = .41$, $F(3, 22) = 4.38$, $p = .02$). In children no significant effects were obtained (p's > .2). Thus in adults WM seems to moderate the relationship between learning and learning-related changes in the P300 to rewards, whereas no such effect was obtained in children (see Figure 5B).

3.3. ERPs to observed feedback

3.3.1. oFRN. As shown in Figure 6A we found significant main effects of age group and valence (p's $\leq .05$, η_p^2's $> .08$) as well as a significant age group by valence interaction, $F(1, 43) = 5.37$, $p = .03$, $\eta_p^2 = .11$. The oFRN was larger for losses than rewards in children ($t(21) = -3.22$, $p = .004$,) but not in adults*,* ($p = .57$) (see Figure 6A).

3.3.2. Feedback-related oP300. Similar to the P300 we found a significant interaction between the factors age group, valence, and block half (see Figure 6B) in the oP300 ($F(1, 43) = 4.05$, $p = .05$, $\eta_p^2 = .09$). Separate analyses for the factor valence showed a significant interaction between block half and age group, ($p = .015$, $\eta_p^2 = .13$) for rewards, but not for losses ($p = .64$). As shown in Figure 6B, comparisons between block halves revealed a decrease in the oP300 to rewards as a function of learning in adults (two-tailed t test, $p = .05$), but not in children ($p = .09$). No further main effects or interactions reached significance (p's > .2).

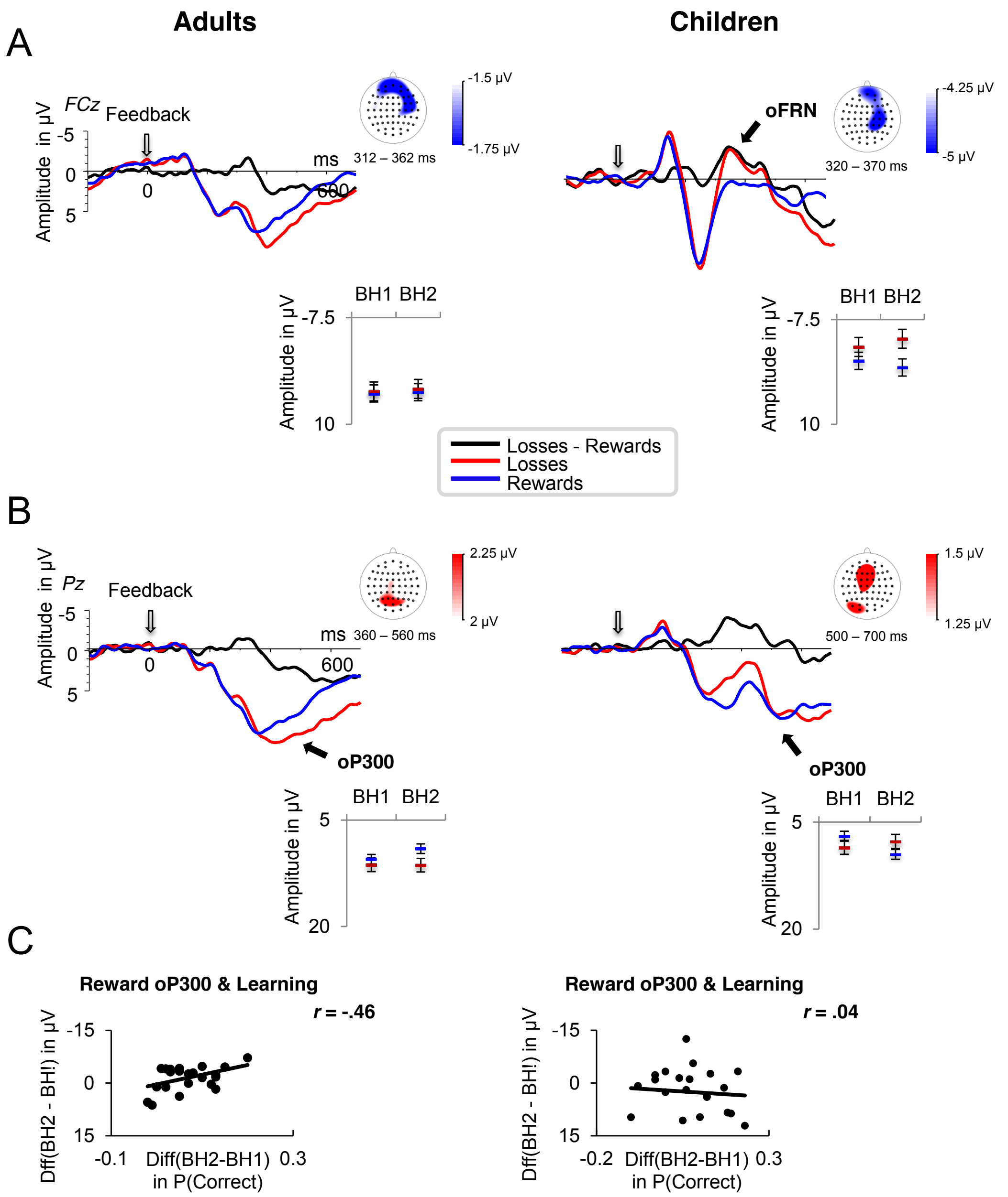

Figure 6. **ERPs observed feedback**. Grand averages shown for losses (red line) and rewards (blue line) for the *(A) oFRN* displayed at FCz and the (*B) oP300* displayed at Pz separately for both age groups, as well as the learning related changes for losses and rewards separately for both block halves (BH1 and BH2). The topographic map displays the difference (black line) between losses and rewards for the oFRN (within 50 ms) and oP300 (within 200 ms). *(C) Correlation effects.* Scatter plots illustrating the correlation between difference score of proportion of correct choice (second – first block half) on the x-axis and the difference score of the mean oP300 amplitude to rewards (second – first block half) on the y-axis separately for the two age groups.

3.3.3. Correlation analyses. As shown in Figure 6C the learning effects in the oP300 to rewards (amplitude difference between block halves) were positively correlated with behavioral learning effects (accuracy difference between block halves) in the AO condition ($r(23) = -.46$, $p = .03$) in adults, but not in children ($r(22) = .04$, $p = .85$). For the oFRN no significant correlations were observed (p's > .3). Similar to our findings in the P300 to experienced feedback this result suggests that the oP300 (but not the oFRN) reflects the degree to which individuals can use information from vicarious experience during learning to update feedback predictions (see Figure 6). Fisher's Z test showed a marginally significant difference between the correlation coefficients of the two age groups (two-tailed test for behavioral learning: $z = -1.67$, $p = .09$).

3.4. Comparison between ERPs to experienced versus observed feedback

In a final analysis step we compared ERPs to experienced feedback (FRN, P300) and ERPs to observed feedback (oFRN, oP300). As previously reported we found significant main effects of valence (*p's* $\leq$.001, η_p^2's > .3), as well as main effects of age group and agency (*p's* $\leq$.001, η_p^2's > .3). For the P300/oP300, similar to the previously reported results, we found valence effects for the P300 ($F(1, 44) = 52.65$, $p < .001$, $\eta_p^2 = .55$), but not for the oP300 ($F(1, 44) = .98$, $p = .33$, $\eta_p^2 = .02$), indicated by an interaction between valence and agency ($F(1, 43) = 18.76$, $p < .001$, $\eta_p^2 = .3$). Moreover, age group and agency interacted significantly for the FRN/oFRN, $F(1, 43) = 4.89$, $p = .03$, $\eta_p^2 = .1$, and the P300/oP300, $F(1, 43) = 24.46$, $p < .001$, $\eta_p^2 = .36$ (see Figure 3 & 6). Separate analyses for the two age groups showed larger ERPs to experienced (FRN/P300) than observed (oFRN/oP300) feedback in adults (*p's* < .001), but not in children (*p's* >.2). These results indicate that adults were more sensitive to experienced than observed feedback, whereas children attribute similar weight to both types of feedback.

4. Discussion

In this age-comparative study we used an observational learning paradigm (Burke et al., 2010; Rodriguez Buritica et al., 2016) to compare behavioral and electrophysiological markers of experience-based and observational learning in 8-10 year old children and younger adults. Although both age groups benefitted from information about the other's actions and outcomes during learning, adults learned much faster from vicarious information (see Figure 2). The ERP results indicate that these age differences in observational learning are associated with a developmental shift in the processing of experienced and observed feedback during learning: children are less able than adults to disengage from processing the valence of experienced feedback during learning. This effect is further exaggerated when learning from observed information. In contrast, young adults are better able to disengage from valence processing during learning and are more efficient in integrating experienced and observed feedback information.

4.1. Benefits of observing other's feedback for learning

In line with previous developmental studies (Bos et al., 2012; Crone et al., 2004; Duijvenvoorde et al., 2008; Eppinger et al., 2009; Hämmerer et al., 2010), adults performed better than children and showed more learning than children across all conditions (see also Figure 2). As hypothesized, both age groups benefited from observing the actions *and* outcomes of the other player (in the AO condition) (Burke et al., 2010; Rodriguez Buritica et al., 2016). However, this effect was exaggerated in younger adults, who showed rapid learning in the AO condition, whereas children learned more gradually from the observed information. In line with recent findings on developmental differences in behavioral learning strategies (Decker et al., 2016), children might be less able to use the additional information of the other player (particularly in the AO condition) to form and test hypotheses about their own outcomes. Previous studies refer to this type of learning as goal-directed or model-based learning (Eppinger et al., 2013; Otto et al., 2013) and the findings by Decker and colleagues suggest that the ability to use such a learning strategy emerges around adolescence (see also Li and Eppinger, 2016). One cognitive ability that has been linked to model-based learning is WM capacity (Eppinger et al., 2013; Otto et al., 2013). In relating these findings to the current results it could be that developmental differences in WM functions (Fry and Hale, 1996; Kwon et al., 2002) are one of the primary sources of the developmental differences in observational learning. Future studies should address this hypothesis more directly and should also try to clarify whether age differences in the ability to integrate social information during learning relate to processes involved in model-based learning.

4.2. ERPs to experienced feedback

4.2.1. FRN. In line with previous developmental findings we observed a larger FRN for losses compared to rewards and for children compared to adults (see Figure 3A), suggesting that children react more strongly to the valence of external feedback than adults (Eppinger et al., 2009; Hämmerer et al., 2010). However, we found no evidence for valence by condition or age by condition interactions in the FRN. This corroborates several previous results indicating that the feedback negativity reflects a rapid and relatively coarse evaluation of events along a valence (good vs. bad) dimension (Eppinger et al., 2008; Hajcak et al., 2006; Hämmerer et al., 2010).

4.2.2. P300. In contrast to the FRN, the P300 valence effects varied as a function of learning condition (Rodriguez Buritica et al., 2016). As shown in Figure 4A across age groups the P300 difference between losses and rewards was larger in the AO compared to the other learning conditions. Thus, the P300 reward effect increases with increasing amount of observable information. We think that this might reflect the fact that with better performance the probability of positive feedback increases whereas the likelihood of negative feedback decreases. That is, rewards may be experienced as more expected whereas negative outcomes may be experienced as more surprising (Mars et al., 2008; Neville et al., 1986). Consistent with this interpretation the P300 differentiation between losses and rewards correlated with behavioral performance (see Figure 4B). That is, the more subjects expect themselves to be correct, the greater the P300 for losses and the smaller the P300 for rewards.

Our P300 results also point to specific developmental differences: In line with prior research (Polich et al., 1990; van Dinteren et al., 2014), adults showed a larger P300 than children (see Figure 3B). Moreover, the P300 to rewards decreased as a function of learning in adults but not in children and this learning-related decrease in the P300 correlated with behavioral learning effects as well as working memory capacity (WM). Together, these findings suggest that the P300 in adults may reflect the updating of reward predictions (Philiastides et al., 2010; Wu and Zhou, 2009) and that this effect is sensitive to individual differences in WM (Gevins and Smith, 2000; Nittono et al., 1999). In line with this interpretation, results of a regression analysis suggest that in younger adults WM capacity may moderate the relationship between behavioral learning effects and learning-related changes in the P300 (see Figure 5B). In contrast no such effect is observed in children. These findings are consistent with prior research showing that the P300 rather than the FRN correlates with reward prediction errors generated during reinforcement learning (Fischer and Ullsperger, 2013; Ullsperger et al., 2014) and that reward-based learning depends on WM abilities (Collins and Frank, 2012). Thus, our electrophysiological findings suggest that WM abilities might be a limiting factor for reward-based (probabilistic) learning in children compared to adults.

Taken together, the current results suggest that the P300 reflects the degree to which individuals can use information from own and observed experience to update outcome predictions (Fischer and Ullsperger, 2013; Ullsperger et al., 2014). However, there are also alternative accounts suggesting that the P300 may reflect context updating (Donchin, 1981). One way of combining these views would be to assume that the P300 may reflect the computational mechanisms that determine the degree to which behavior should be updated based on a given prediction error (i.e. areas involved in determining the rate of learning) and that this process may in part depend on WM capacity (Nassar et al., 2016). According to such an interpretation the absence of this effect in children may reflect difficulties in determining (or adjusting) the optimal rate of learning in a given environment (McGuire et al., 2014; Nassar et al., 2012) as well as developmental differences in WM functions (Fry and Hale, 1996; Kwon et al., 2002). On a more general level such an interpretation suggests that developmental differences in learning may not be the consequence of differences in prediction error signaling per se (Cohen et al., 2010; Hauser et al., 2015) but rather in determining how much to learn from a given prediction error (van den Bos et al., 2012; McGuire et al., 2014).

4.3. ERPs to observed feedback

4.3.1. oFRN. Children showed a larger oFRN than adults when observing the outcomes of others' actions, confirming previous results that children are generally more susceptible to external (experienced or observed) negative feedback information during learning than adults (van den Bos et al., 2012; Crone et al., 2004; Duijvenvoorde et al., 2008; Eppinger et al., 2009; Ferdinand and Kray, 2014; Hämmerer et al., 2010; Rodriguez Buritica et al., 2016). In adults, we found no significant differentiation between losses and rewards in the oFRN, indicating that the valence information of observed feedback is less salient for them (see Figure 6A). One interpretation of these findings is that children are less able to rapidly assess the information value of feedback (that is, how much the feedback should guide learning). As a consequence they up-regulate their response to negative feedback, independently of who receives the feedback. In contrast, the younger adults may use the observed feedback in a more goal-directed manner and focus less on the valence information.

4.3.2. oP300. Similar to the findings in the P300 to experienced outcomes, in adults but not in children the P300 to observed rewards decreased with learning (see Figure 6B). Moreover and similar to the findings in the P300 to experienced outcomes, in adults but not in children the oP300 valence effect was positively related to learning effects in the AO condition (see Figure 6C). Thus, consistent with the interpretation for the P300 to experienced outcomes our results suggest that adults are able to use the observed information to update reward predictions. In contrast, children seem to be stuck in processing the valence of the observed outcomes but are unable to determine how much they should learn from this information.

4.3.3. Comparison of ERPs & oERPs. Children's ERP responses to observed outcomes were similar in magnitude to their ERPs to experienced feedback (see Figure 3 & 6). In adults the ERPs to observed outcomes were overall significantly smaller than the ERPs to experienced feedback (Bellebaum et al., 2010). This finding is consistent with the general notion that children compared to adults are more susceptible to external (and especially to negative) feedback information during learning than adults (Eppinger et al., 2009; Ferdinand and Kray, 2014; Hämmerer et al., 2010). Moreover, it extends this notion by suggesting that this greater sensitivity to negative feedback in children is also found when they observe feedback that other individuals receive (Rodriguez Buritica et al., 2016).

Taken together, our findings suggest that children in the age of 8 to 10 years do not yet appropriately assess the information value of feedback (that is, how much it should guide learning) and thus are less efficient in integrating (particularly negative) outcomes into their behavioral strategies than adults (van den Bos et al., 2012; Eppinger et al., 2009; Hämmerer et al., 2010; Rodriguez Buritica et al., 2016). As a consequence, they might generally up-regulate their response to external negative feedback, independently of who receives the feedback. When transferring the current findings to an applied context one tempting interpretation of our results is that educational intervention programs designed to enhance learning in children should focus on their ability to use experienced feedback for learning and should be *particularly cautious* with using negative social feedback. Thus, at least in the age range between 8 and 10 years, supervised learning may be less effective than its prevalence in education may suggest.

5. Conclusion

Our results indicate a developmental shift in the processing of experienced and observed feedback during learning: Although both age groups benefitted from observational information during learning, adults learned much faster compared to children. Our electrophysiological results suggest that adults are more efficient in using the additional information in the fully observational (AO) condition, which is in line with their accelerated learning effects in this condition. Children also benefit from the observational information, but rely more on external (particularly negative) feedback (oFRN) during learning than adults. Thus, their behavior seems to be driven more by the valence of the observed feedback rather than the information content of the feedback. These developmental differences may reflect a more general developmental trend in the ability to use observed and experienced feedback for goal-directed learning. According to our findings, educational practice should enhance children's ability to use experienced feedback for learning and be particularly cautious with negative social feedback during supervised learning.

Acknowledgments

We would like to thank Anna Würtz, Lissa Mapouyat and Katharina Wermuth for help during data acquisition, as well as our participants for their contribution to the study. This research was in part funded by grants of the German Federal Ministry of Education and Research |BMBF; FKZ 01GQ0913, FKZ 01GQ1313] awarded to Shu-Chen Li and Hauke Heekeren and by the Freie Universität Berlin. Julia Rodriguez Buritica's doctoral fellowship was supported by the Berlin School of Mind and Brain. We are grateful to the Center of Lifespan Psychology at the Max Planck Institute for Human Development for making EEG data collections possible.

References

American Electroencephalographic Society, 1994. Guideline thirteen: guidelines for standard electrode position nomenclature. J. Clin. Neurophysiol. Off. Publ. Am. Electroencephalogr. Soc. 11, 111–113.

Bellebaum, C., Kobza, S., Thiele, S., Daum, I., 2010. It was not MY fault: event-related brain potentials in active and observational learning from feedback. Cereb. Cortex N. Y. N 1991 20, 2874–2883.

Bos, W. van den, Cohen, M.X., Kahnt, T., Crone, E.A., 2012. Striatum–Medial Prefrontal Cortex Connectivity Predicts Developmental Changes in Reinforcement Learning. Cereb. Cortex 22, 1247–1255.

Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vis. 10, 433–436.

Burke, C.J., Tobler, P.N., Baddeley, M., Schultz, W., 2010. Neural mechanisms of observational learning. Proc. Natl. Acad. Sci. 107, 14431–14436.

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. Nat. Neurosci. 13, 669–671.

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. Eur. J. Neurosci. 35, 1024–1035.

Crone, E.A., Richard, J., Van der Molen, M.W., 2004. Developmental Change in Feedback Processing as Reflected by Phasic Heart Rate Changes. Dev. Psychol. 40, 1228–1238.

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. Psychol. Sci. 956797616639301.

Donchin, E., 1981. Surprise!… Surprise? Psychophysiology 18, 493–513.

Duijvenvoorde, A.C.K. van, 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. J. Neurosci. 28, 9495–9503.

Ekstrom, R. B., French, J. W, Harman, H. H., Dermen, D., 1976. Manual for kit of factor referenced cognitive tests. Princeton, NJ: Educational Testing Service.

Eppinger, B., Kray, J., Mock, B., Mecklinger, A., 2008. Better or worse than expected? Aging, learning, and the ERN. Neuropsychologia 46, 521–539.

Eppinger, B., Mock, B., Kray, J., 2009. Developmental differences in learning and error processing: Evidence from ERPs. Psychophysiology 46, 1043–1053.

Eppinger, B., Walter, M., Heekeren, H.R., Li, S.-C., 2013. Of goals and habits: age-related and individual differences in goal-directed decision-making. Front. Neurosci. 7.

Ferdinand, N.K., Kray, J., 2014. Developmental changes in performance monitoring: How electrophysiological data can enhance our understanding of error and feedback processing in childhood and adolescence. Behav. Brain Res. 263, 122–132.

Fischer, A.G., Ullsperger, M., 2013. Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. Neuron 79, 1243–1255.

Fry, A.F., Hale, S., 1996. Processing Speed, Working Memory, and Fluid Intelligence: Evidence for a Developmental Cascade. Psychol. Sci. 7, 237–241.

Geisser, S., Greenhouse, S.W., 1958. An Extension of Box's Results on the Use of the F Distribution in Multivariate Analysis. Ann. Math. Stat. 29, 885–891.

Gevins, A., Smith, M.E., 2000. Neurophysiological Measures of Working Memory and Individual Differences in Cognitive Ability and Cognitive Style. Cereb. Cortex 10, 829–839.

Gratton, G., Coles, M.G., Donchin, E., 1983. A new method for off-line removal of ocular artifact. Electroencephalogr. Clin. Neurophysiol. 55, 468–484.

Groenendijk, T., Janssen, T., Rijlaarsdam, G., Bergh, H., 2013. Learning to be creative. The effects of observational learning on students' design products and processes. Learn. Instr. 28, 35–47.

Hajcak, G., Moser, J.S., Holroyd, C.B., Simons, R.F., 2006. The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. Biol. Psychol. 71, 148–154.

Hämmerer, D., Li, S.-C., Müller, V., Lindenberger, U., 2010. Life Span Differences in Electrophysiological Correlates of Monitoring Gains and Losses during Probabilistic Reinforcement Learning. J. Cogn. Neurosci. 23, 579–592.

Hauser, T.U., Iannaccone, R., Walitza, S., Brandeis, D., Brem, S., 2015. Cognitive flexibility in adolescence: neural and behavioral mechanisms of reward prediction error processing in adaptive decision making during development. NeuroImage 104, 347–354.

Kwon, H., Reiss, A.L., Menon, V., 2002. Neural basis of protracted developmental changes in visuo-spatial working memory. Proc. Natl. Acad. Sci. 99, 13336–13341.

Lehrl, S., 2005. Manual zum MWT-B: [Mehrfachwahl-Wortschatz-Intelligenztest], 5., unveränd. Aufl. ed. Spitta-Verl, Balingen.

Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., Baltes, P.B., 2004. Transformations in the Couplings Among Intellectual Abilities and Constituent Cognitive Processes Across the Life Span. Psychol. Sci. 15, 155–163.

Li, S.-C., Schmiedek, F., Huxhold, O., Röcke, C., Smith, J., Lindenberger, U., 2008. Working memory plasticity in old age: Practice gain, transfer, and maintenance. Psychol. Aging 23, 731–742.

Li, S.-C., Eppinger, B., 2016. Lifespan development of adaptive neurocognitive representations: reciprocal interactions between cognition and motivation. In Todd Braver (Ed.), Motivation and cognitive control (pp.283-312). New York: Psychological Press Taylor & Francis.

Lindenberger, U., Mayr, U., Kliegl, R., 1993. Speed and intelligence in old age. Psychol. Aging 8, 207–220.

Luck, S. J., 2012. The Oxford Handbook of Event-Related Potential Components. Oxford University Press, USA.

Mars, R.B., Debener, S., Gladwin, T.E., Harrison, L.M., Haggard, P., Rothwell, J.C., Bestmann, S., 2008. Trial-by-Trial Fluctuations in the Event-Related Electroencephalogram Reflect Dynamic Changes in the Degree of Surprise. J. Neurosci. 28, 12539–12545.

Marshall, P.J., Young, T., Meltzoff, A.N., 2011. Neural correlates of action observation and execution in 14-month-old infants: an event-related EEG desynchronization study. Dev. Sci. 14, 474–480.

McGuire, J.T., Nassar, M.R., Gold, J.I., Kable, J.W., 2014. Functionally dissociable influences on learning rate in a dynamic environment. Neuron 84, 870–881.

Meltzoff, A.N., Waismeyer, A., Gopnik, A., 2012. Learning about causes from people: Observational causal learning in 24-month-old infants. Dev. Psychol. 48, 1215–1228.

Morgan, H.M., Klein, C., Boehm, S.G., Shapiro, K.L., Linden, D.E.J., 2008. Working memory load for faces modulates P300, N170, and N250r. J. Cogn. Neurosci. 20, 989–1002.

Nassar, M.R., Bruckner, R., Gold, J.I., Li, S.-C., Heekeren, H.R., Eppinger, B., 2016. Age differences in learning emerge from an insufficient representation of uncertainty in older adults. Nat. Commun. 7, 11609.

Nassar, M.R., Rumsey, K.M., Wilson, R.C., Parikh, K., Heasly, B., Gold, J.I., 2012. Rational regulation of learning dynamics by pupil-linked arousal systems. Nat. Neurosci. 15, 1040–1046.

Neville, H.J., Kutas, M., Chesney, G., Schmidt, A.L., 1986. Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. J. Mem. Lang. 25, 75–92.

Nielsen, M., Tomaselli, K., 2010. Overimitation in Kalahari Bushman Children and the Origins of Human Cultural Cognition. Psychol. Sci.

Nittono, H., Nageishi, Y., Nakajima, Y., Ullsperger, P., 1999. Event-related potential correlates of individual differences in working memory capacity. Psychophysiology 36, 745–754.

Oldfield, R.C., 1971. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9, 97–113.

Otto, A.R., Gershman, S.J., Markman, A.B., Daw, N.D., 2013. The Curse of Planning Dissecting Multiple Reinforcement-Learning Systems by Taxing the Central Executive. Psychol. Sci. 956797612463080.

Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current density mapping. Electroencephalogr. Clin. Neurophysiol. 72, 184–187.

Philiastides, M.G., Biele, G., Vavatzanidis, N., Kazzer, P., Heekeren, H.R., 2010. Temporal dynamics of prediction error processing during reward-based decision making. NeuroImage 53, 221–232.

Polich, J., Ladish, C., Burns, T., 1990. Normal variation of P300 in children: Age, memory span, and head size. Int. J. Psychophysiol. 9, 237–248.

Raven, J.C., Court, J.H., 1998. Raven's progressive matrices and vocabulary scales. Oxford Psychologists Press.

Raven, J.C., Raven, J., 2002. Court JH Coloured Progressive Matrices (CPM) (Dt. Bearbeitung von Bulheller S, Häcker HO, Pearson Assessment, Frankfurt 2002).

Reinhart, R.M.G., Woodman, G.F., 2014. Causal control of medial-frontal cortex governs electrophysiological and behavioral indices of performance monitoring and learning. J. Neurosci. Off. J. Soc. Neurosci. 34, 4214–4227.

Rendell, L., Fogarty, L., Hoppitt, W.J.E., Morgan, T.J.H., Webster, M.M., Laland, K.N., 2011. Cognitive culture: theoretical and empirical insights into social learning strategies. Trends Cogn. Sci. 15, 68–76.

Rodriguez Buritica, J.M., Eppinger, B., Schuck, N.W., Heekeren, H.R., Li, S.-C., 2016. Electrophysiological correlates of observational learning in children. Developmental Science, 19(5), 699–709.

Ullsperger, M., Fischer, A.G., Nigbur, R., Endrass, T., 2014. Neural mechanisms and temporal dynamics of performance monitoring. Trends Cogn. Sci. 18, 259–267.

van den Bos, W., Cohen, M.X., Kahnt, T., Crone, E.A., 2012. Striatum–Medial Prefrontal Cortex Connectivity Predicts Developmental Changes in Reinforcement Learning. Cereb. Cortex 22, 1247–1255.

van Dinteren, R., Arns, M., Jongsma, M.L.A., Kessels, R.P.C., 2014. P300 Development across the Lifespan: A Systematic Review and Meta-Analysis. PLoS ONE 9, e87347.

Windell, O., 2008. Vector Snowflake Application. Retrieved September 7, 2012, from http://www.evilmadscientist.com/2008/vector-snowflake-application/.

Wu, Y., Zhou, X., 2009. The P300 and reward valence, magnitude, and expectancy in outcome evaluation. Brain Res. 1286, 114–122.

Yu, R., Zhou, X., 2006. Brain responses to outcomes of one’s own and other’s performance in a gambling task. Neuroreport 17, 1747–1751.

SUPPLEMENTARY INFORMATION

Model-generated choices of "model players"

The choices of the "model players" presented to the participants were generated using a Q learning algorithm (Burke et al., 2010; Sutton and Barto, 1998). Specifically, after a reward *r*, the value *Q* of action *a* in the next trial was calculated according to the delta updating rule:

$$Q_a(t+1) = Q_a(t) + \alpha\left[r(t) - Q_a(t)\right]$$

where *α* is the learning rate, *r(t)* the reward obtained after performing action *a* and *t* indexes the current trial. The probability of performing action *a* was computed using a softmax function (O'Doherty, 2004).

$$P(a) = \frac{e^{Q_a(t)/\beta}}{\sum_{c \in A} e^{Q_c(t)/\beta}}$$

where *P*(*a*) is the probability of choosing action *a*, *A* is the set of all possible actions and *β* is the temperature parameter that controls the competition between possible choices. The computer-controlled behavior of the model players was associated with the same percentage of probabilistic positive or negative outcomes (80% gains for the good, 20% for the bad choice), like the participants experienced during the individual learning conditions. To ensure comparability between conditions (see Figure 1C), the mean of the rewards obtained by the model were constrained to small deviations from each condition's true mean with a 2.5% maximum deviation (between 77.5% and 82.5% upon choosing the good option and between 17.5% and 22.5% upon choosing the bad option). The Q-values were set to zero at the beginning of the task and continuously updated on subsequent trials. The β and α parameters were estimated based on data of 30 subjects (acquired in a prior pilot testing) from 10^5-simulated runs with the same model.

References

Burke, C.J., Tobler, P.N., Baddeley, M., Schultz, W., 2010. Neural mechanisms of observational learning. Proc. Natl. Acad. Sci. 107, 14431–14436.

O'Doherty, J.P., 2004. Reward representations and reward-related learning in the human brain: Insights from neuroimaging. Curr. Opin. Neurobiol. 14, 769–776.

Sutton, R.S., Barto, A.G., 1998. Introduction to Reinforcement Learning, 1st ed. MIT Press, Cambridge, MA, USA.

Title:

Adolescents are sensitive to peer influence, but only for so long

Julia M. Rodriguez Buritica[1,2] , Hauke R. Heekeren[1], and Wouter van den Bos[3]

[1]Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany; [2]Center of Lifespan Psychology, Max-Planck-Institute for Human Development, Berlin, Germany; [3]Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

Address for correspondence:

Julia M. Rodriguez Buritica, buritica@zedat.fu-berlin.de,

phone: +49 (0) 30 838 – 55749

Department of Education Sciences and Psychology, Freie Universität,

Habelschwerdter Allee 45, 14195 Berlin, Germany

Abstract

Advice taking helps to quickly acquire knowledge and make decisions. After the initial advice, however, one may start gathering first hand experience and explore alternatives. Across development, advice seems to be weighted differently against own experience. To capture developmental differences in how advice, experience and exploration influence learning, we used a 4-armed bandit task including an initial good peer advice in 8-10, 13-15 and 18-22-year olds. We show that although adolescents are highly sensitive to peer advice, they are also more exploratory like children. Whereas adults stayed with the advice over the task, only adolescents explored another good alternative. Our social learning model resolved the apparently conflicting findings of adolescents being either more or less sensitive to peer influence. Here we show that adolescents are indeed initially easily swayed to follow peer advice, but also faster exploring alternatives and put more weight on own experience compared to adults.

Keywords: development, advice-taking, social reinforcement learning, learning from experience, exploration, decision making

1. Introduction

When faced with uncertain situations, advice taking helps us to quickly acquire knowledge during learning and decision making (Biele, Rieskamp, Krugel, & Heekeren, 2011). After the initial advice one may start gathering first hand experience. For instance, when trying a new restaurant, you will decide to order the dish recommended by a friend who went there before. However, after a while you may wonder about the other dishes, particularly if the advised one is not so spectacular, and explore more of the menu. These new experiences may confirm that the friends' advice was good or not.

The integration of social information, such as advice, is very relevant from a developmental perspective given that it represents an important source for cognitive and socio-emotional development (Meltzoff, Waismeyer, & Gopnik, 2012; Nielsen & Tomaselli, 2010). Especially during adolescence social information becomes increasingly influential on learning and decision-making processes (Blakemore & Mills, 2014; Jones et al., 2014). This greater sensitivity to peer advice can have negative effects, as in the case of peer influence on risk-taking (Chein, Albert, O'Brien, Uckert, & Steinberg, 2011; Steinberg, 2008). But it may also have positive effects, for example a recent study showed that one's own experience during learning may overrule (bad) advice (i.e. for a specific choice option within a learning task) faster in adolescents compared to adults. (Decker, Lourenco, Doll, & Hartley, 2015). This may be advantageous when the social information was of low quality. Indeed, Decker and colleagues (2015) reported that inaccurate task instructions influenced adults more than younger age groups. On the other hand, learning from experience also shows significant developmental differences: For instance, children and adolescents may not learn to select optimal choice options due to their sensitivity to negative feedback (van den Bos, Cohen, Kahnt, & Crone, 2012; van Duijvenvoorde, Zanolie, Rombouts, Raijmakers, & Crone, 2008). Importantly, this relative sensitivity to negative feedback of younger age groups may also contribute to developmental differences in responses to bad advice.

In sum, recent studies have started to outline developmental differences in social influence (Lourenco et al., 2015; Monahan, Steinberg, & Cauffman, 2009; Steinberg, 2008; Steinberg & Monahan, 2007; Sumter, Bokhorst, Steinberg, & Westenberg, 2009). These studies suggest that there are important developmental differences in how advice and experience are weighted during learning and raised important questions: For instance, some of these studies suggest adolescents are most sensitive to advice (e.g. Steinberg & Monahan, 2007), whereas others suggest they are more sensitive to experience (e.g. Decker et al., 2015). In addition, developmental changes in using inaccurate social information can be due to differences in weighing advice and/or the negative feedback associated with following that advice. To resolve some of these outstanding issues we use a task where participants are provided with good advice (Biele, Rieskamp, & Gonzalez, 2009), in combination with computational modeling. The latter allows us to separate the effects of advice and feedback valence on learning. In addition, the model allows us to measure the short- vs. long-term effects of advice on behavior.

In the current study we used a reinforcement-learning task (modified after Biele et al., 2011), where participants had to choose one out of four card decks, which were associated with gains and losses (see Figure 1 & S1). Participants should gain as much points as possible by choosing more beneficial decks (associated with higher expected values). Unbeknownst to the participants, two of the four decks were associated with higher expected positive values ("good decks") than the other two ("bad decks"). At the beginning of the experiment participants received a good advice. Crucially, only one of the "good decks" was advised (see Methods for further details). Thus, the preference for the advised deck over the other good deck is a straight forward indicator for the effect of advice.

Based on previous work (Biele et al., 2011), we predicted that all age groups should benefit from advice taking during learning. Adolescents, however, should show the strongest sensitivity to peer influence (see van Hoorn, van Dijk, Meuwese, Rieffe, & Crone, 2016 for review), at least in the beginning of the task before learning by experience takes over (Decker et al., 2015). Throughout learning children and adolescents should show more exploration than adults (Decker et al., 2015). As a result, the two younger groups should more often discover that there is another (equally) good deck in the experiment. Children, should encounter the other positive deck less often than adolescents due to their greater difficulties to use negative feedback for leaning (van Duijvenvoorde et al., 2008). Reinforcement learning (RL) models can capture these developmental differences during learning (van den Bos, Cohen, Kahnt, & Crone, 2012) and have been useful to further describe also social learning mechanisms during advice taking in adults (Behrens, Hunt, & Rushworth, 2009; Biele et al., 2009). Previous studies with adults (Biele et al., 2009, 2011) using RL models suggest that the advised option is associated with a bonus when choosing it, which has a long term positive effect on the valuation of that option. Furthermore, these studies suggest that advice may lead to a higher initial propensity for one of the decks, which has a short-term effect and then changes quickly based on experience. Here we apply and extend these models to further capture developmental differences in how advice, experience and exploration influence learning and decision making. These techniques enable us to get a more mechanistic understanding of developmental dynamics of social influence (van den Bos & Eppinger, 2016).

2. Methods

2.1. Participants

The effective sample (see supplemental for justification of sample size) of the study consisted of 25 adults between 18-22 years of age (13 female, mean age = 20.32, *SD* = 1.15), 24 adolescents between 13-15 years of age (12 female, mean age = 13.71, *SD* = .75), and 24 children between 8-10 years of age (10 female, mean age = 9.08, *SD* = .83). Data of one adult were excluded due to techniqual problems during data collection and data of one adolescent were excluded from further analyses due to many missing responses (2 SD from the group mean). All participants had normal or corrected-to-normal vision and no neurological or psychological disorders. Prior to the experiment we obtained informed consent from the participants and their parents (in case of children). The study was approved by the Ethics Committee of the Max-Planck-Institute for Human Development, Berlin.

Subjects participated in one behavioral group session (together with same-aged participants) in which we assessed psychometric covariate measures and performance in the advice task. Participants received a compensation of 15 Euro.

General cognitive abilites of the sample were described using *I.* Numbers task (Gold, Carpenter, Randolph, Goldberg, & Weinberger, 1997) as a marker for working memory (WM) and *II.* a short version oft the CFT (Weiß, 2006) as a marker for fluid intelligence. Children had lower WM scores compared to adolescents and adults [children: $M = 7.52$ ($SD = 1.71$); adolescents: $M = 10.23$ ($SD = 1.71$); adults: $M = 10.66$ ($SD = 2.18$); $F(2, 70) = 19.77$, $p < .001$, $\eta_p^2 = .36$)], wheras adolescents and adults did not differ in their WM scores ($t(48) = .87$, $p = .39$). Fluid intelligence scores increased with age [(children: $M = 9.56$ ($SD = 2.68$), adolescents: $M = 12.29$ ($SD = 2.25$), adults: $M = 13.48$

(SD = 1.48); $F(2, 70) = 20.57$, $p < .001$, $\eta_p^2 = .37$)]. These age differences are consitent with findings from larger population-based lifespan samples (Li et al., 2004).

2.2. Experimental Desgin

In the learning task (modified after Biele et al., 2011 and programmed in PsychoPy; Peirce, 2007) participants were asked to choose one out of four card decks (see Figure 1). The goal of the task was to maximize cumultative rewards. Participants could select on card deck within a maximal response time window of 4 sec and received feedback (displayed for 1 sec) afterwards. After a short fixation cross (displayed for 1 sec) a new trial out of a total of 210 trials started. Unbeknownst to the participants, the four decks constisted of two more beneficial „good“ decks and two less beneficial „bad“ decks (see Figure 1B). Prior to the experiment participants recieved once an advice for one of the good decks at the beginning of the task (see Figure 1A). Thus, the preference for the advised deck over the other good deck, would be a clear indicatior for advice taking. Participants were told that another peer (out of a previous session) gave an adivce after he/she played the task. Unbeknownst to the participants the advice was controlled by the experimenter and was always a good advice. Participants were debriefed about the cover story after the experiment.

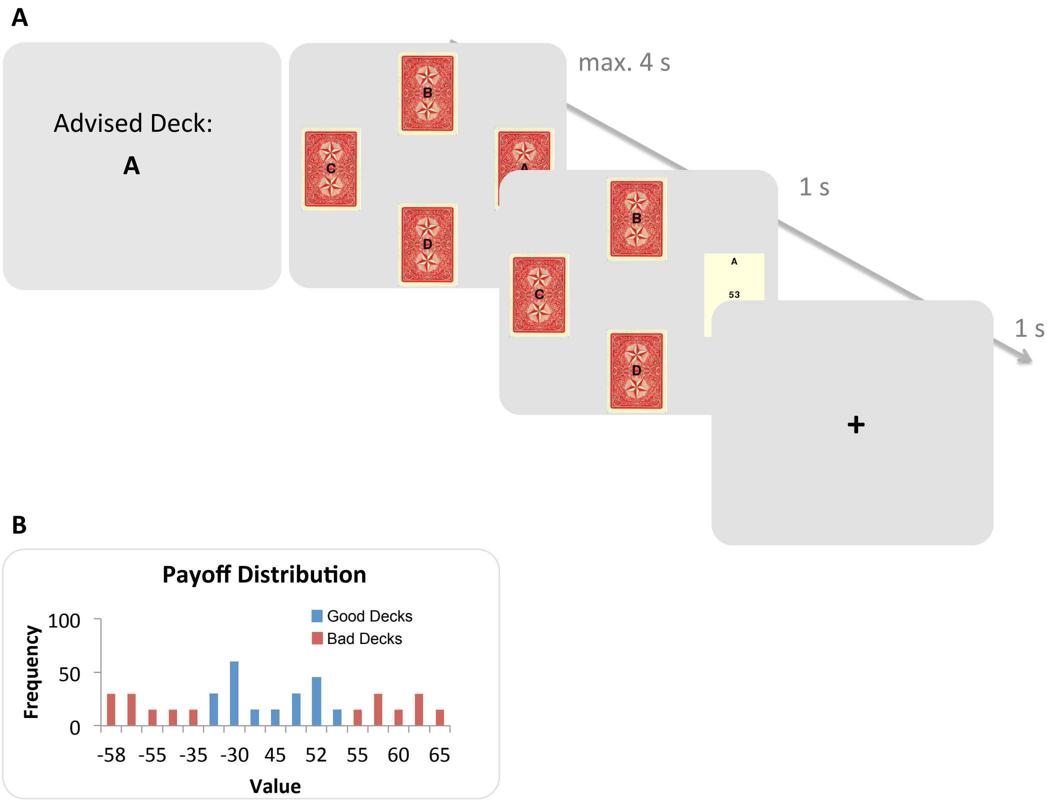

Figure 1. ***Experimental design.*** **(A)** Participants received advice prior they were asked to play a 4-armed bandit task. Every trial started with the presentation of 4 card decks, where one should be selected within max. 4 seconds. Afterward the associated feedback was presented. Before a new trial started a fixation, cross was displayed for 1 second. **(B)** Payoff Distribution separately for the good and the bad decks.

Furthermore, to investigate the effect of advice following on positive and negative feedback, each card deck (good and bad) was associated with 50% losses and 50% gains, although the bad decks were associated with higher losses and slightly higher gains than the good decks (see Figure 1B). Overall, the expected value of the good decks was on average 10, whereas the bad decks had an expected value of 2.5 on average.

2.3. Data analysis

Statistical analyses were performed using R. Responses exceeding the response deadline (4 sec) were excluded from further analyses [children: M = 7.42 trials (SD =

8.05), adolescents: M = 6.83 trials (SD = 7.56), adults: M = 4.24 trials (SD = 5.09)]. Importantly, the three age groups did not differ from each other in terms of missing responses ($F(2, 70) = 1.439$, $p = .24$, $\eta_p^2 = .04$).

Choice behavior was analyzed separately for each deck [Advised Deck (AD), Other Good Deck (GD), Bad Decks (BD)] using logistic regression with the independent predictors of age group (linear trend; adults, adolescents, children), of age group ^{2} (quadratic trend; adults, adolescents, children), trial (1:210, z-transformed), and their interactions. General cognitive abilities were analyzed using univariate ANOVAs and independent samples t-tests. Effect sizes (partial eta squared, η_p^2) are reported.

2.4. Modeling Social Influence

To further investigate the processes underlying the influence of advice on feedback-based learning in the N-armed bandit task, we fit reinforcement learning (RL) models to each participant's behavioral data (Sutton & Barto, 1998). These models have been successfully applied to describe the behavior of teenagers and children (Lourenco et al., 2015; van den Bos et al., 2012), as well as describing the influence of advice (Biele et al., 2009, 2011). The basic RL model uses the prediction error to update the beliefs associated with each choice option (e.g. Deck A, B, C, or D). The prediction error (δ_t) compares the current outcome (r_t) with the predicted outcome (w_t):

$$\delta_t = r_t - w_t(\text{chosen\_stimulus})$$

Whenever feedback is better (worse) than expected, the model will generate a positive (negative) prediction error, which is used to increase (decrease) the predicted value (decision weight, (w_t)) associated with the chosen option. The impact of the prediction errors on forming new beliefs is scaled by the learning rate ($0<\alpha<1$):

$$w(i)_{t+1} = w(i)_t + \alpha \,.\, \delta_t$$

A high learning rate indicates that new information has a much stronger impact on future behavior than less recent information.

To model trial-by-trial choices, we used the soft-max choice rule to compute the probability (*P*) of choosing one of the decks based upon one's own predictions about the outcomes of all decks on trial *t* (Montague, Hyman, & Cohen, 2004):

$$P(i)_t = \frac{e^{(\theta \cdot w(i)_t)}}{\sum_{j=1}^{N} e^{(\theta \cdot w(j)_t)}}$$

where the θ parameter is a free parameter that indicates the sensitivity of the subject to the differences in decision-weights. The lower the θ parameter, the more exploratory the choices appear. To test various models of social influence on these basic learning processes we tested and compared several learning models.

Social Influence RL models. To investigate the influence of advice on learning, we first compared how an "outcome-bonus" model, a "prior" model, and a combined "prior & outcome-bonus" model described participants' choices (Biele et al., 2011). The bonus model differs from the standard RL model by assuming that there is a constant bonus associated with choosing the recommended option. Accordingly, the updating rule was modified such that:

$$w(i)_{t+1} = w(i)_t + \alpha \,.\, (r_t + (\lambda(i) . \beta \,.\, \mu) + -w_t(\text{chosen\_stimulus}))$$

where $\lambda(i)$ is an indicator function that takes the value 1 if option i is recommended and the value 0 if option i is not recommended. The β_b parameter captures the extent to which social influence leads to an outcome bonus ($0<\beta<\infty$), and μ is the expected payoff from choosing randomly among all options (6.25).

The simple "prior" model assumes that the recommendation sets an initial strong positive prior for the recommended deck but has no further influence on choices.

The initial reward expectation in the prior model is defined as:

$$w(recommended)_1 = \beta_p \,.\, \mu$$

where β_p captures the social influence on the prior expectations and μ is the expected payoff from choosing randomly among all options (6.25). Finally, the combined "prior+bonus" model assumes people have an increased prior β_p for the recommended deck, and the recommendation is also associated with a constant bonus β_b.

Next, based on previous developmental studies we expected that gains and losses are asymmetrically updated (Kahnt et al., 2009; van den Bos et al., 2012) thus we extended all the social influence models with two independent learning rates instead of one, i.e. one learning rate for positive feedback (α_{pos}) and one for negative feedback (α_{neg}). Finally, we tested versions of the social influence models that assumed that the outcome bonus associated with the recommended option would decline over time. To model this waning influence of advice over time we modulated the bonus parameter:

$$w(i)_{t+1} = (\beta_b \,.\, \mu).\left(\frac{1}{t}\right)^{\pi} + \ \delta_t)$$

where π is the free parameter that captures how quickly the effect of influence decays, with smaller values indicating less decline ($0 < \pi < \infty$).

Note that for all models except the models with a prior, the decision weight (*w*) for each option was always set to 0 at the beginning of the experiment. This reflects the assumptions that the participants had the same expectation of rewards for each option at the beginning. In addition, following Biele and colleagues (2009; 2011), all models assume that participants who received advice will always choose the recommended option in their first trial. This is implemented by setting probability of choosing the recommended deck at 100% for the first trial.

Model Fitting Procedure. The learning rates and sensitivity parameter were individually estimated by fitting the model predictions to participants' decisions. We used a robust combination of grid-search and maximum likelihood estimation using the Nelder–Mead simplex algorithm implemented in the optim function in R to estimate the model parameters for each participant. For each point on the grid the log likelihood of corresponding parameter setting was estimated. The five grid points that produced the maximum over all starting positions were selected as starting point for finding the final solution using minimization. For model selection purposes we computed the Bayesian information criterion (BIC), where lower BIC values indicate better fit (see Figure 3A).

3. Results

3.1. Choice behavior

Choice behavior was analyzed separately for each deck [Advised Deck (AD), Other Good Deck (GD), Bad Decks (BD)] using a logistic regression with the independent predictors of age group (linear trend; children, adolescents, adults), of age group ^{2} (quadratic trend; children, adolescents, adults), trial (1:210, z-transformed), and their interactions.

Advised Deck. The advised deck was chosen more often with increasing age (Age: $b1 = 0.06$, $SD = .01$, $t = 8.30$, $p < .001$; see Table S1 & Figure 2). With more time of sampling the advised deck was selected less often in adolescents (see Figure 2) compared to the other two age groups (Age^2 * Trial: $b5 = -0.01$, $SD = .004$, $t = -3.24$, $p = .001$; see supplemental Table S1). *Other Good Deck*. As can be seen in Figure 2 adolescents chose the other good deck more often compared to the other two age groups (Age^2: $b2 = 0.04$, $SD = .006$, $t = 6.04$, $p < .001$; see Table 2) and with more time of sampling (Age^2 * Trial: $b5 = 0.04$, SD $= .01$, $t = 6.04$, $p < .001$; see supplemental Table S2 & Figure 2). *Bad Decks*. Bad Decks were selected less often with increasing age

(Age: $b1$ = - 0.04, SD = .007, t = -6.34, p < .001; see Table S3) and by adolescents compared to adults and children (Age^2: $b2$ = -0.05, SD = .007, t = - 6.79, p < .001; see supplemental Table S3 & Figure 2).

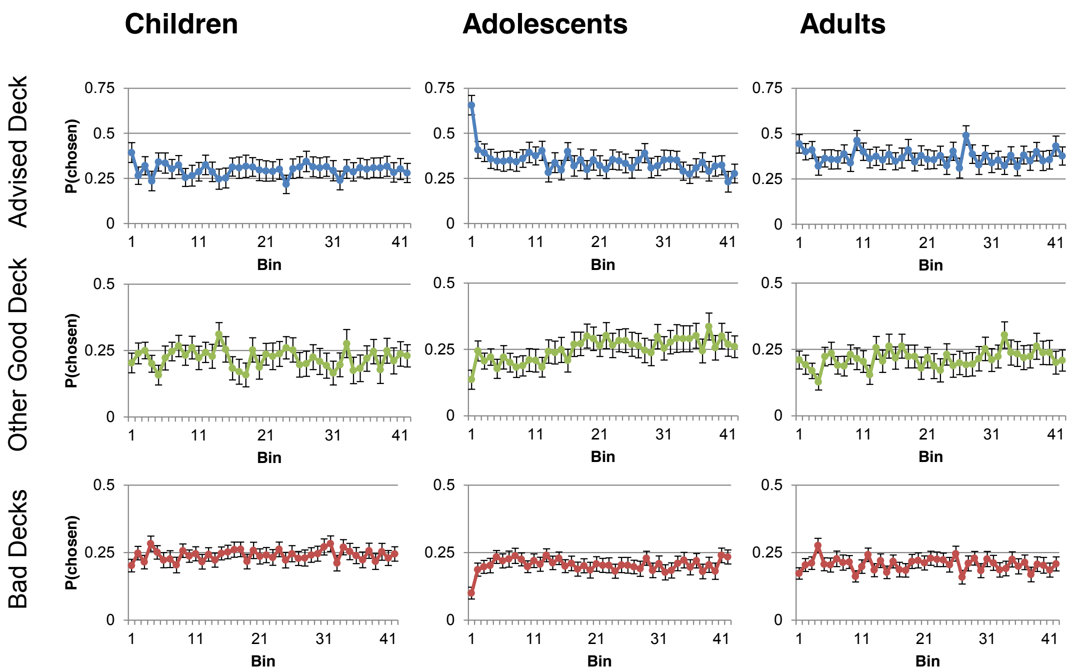

Figure 2. *Choice Behavior for the Advised Deck, Other Good Deck and Bad Decks (collapsed over the two bad decks).* Proportion chosen separately for bin (of 5 trials each) and age group.

Expected Value. Finally, we ran a logistic regression where we analyzed the expected value (EV) of each choice per trial (i.e., $EV_{advised}$ = 10, $EV_{good\_alternative}$ = 10 and $EV_{bad\_alternatives}$ = 2.5) with the independent predictors of age group (linear trend; adults, adolescents, children), of age group ^{2} (quadratic trend; adults, adolescents, children), trial (1:210, z-transformed), and their interactions. Interestingly, the expected value was higher with increasing age (Age: $b1$ = 0.33, SD = .054, t = 6.34, p < .001; see Table 4) and for adolescents compared to adults and children (Age^2: $b1$ = 0.36, SD = .053, t = 6.79, p < .001). No further interactions with the variable trial reached significance (p's >.325)

3.2. Modeling Social Influence

The model comparison yielded several important findings. First, as expected, the model comparisons show that all social models are better than the baseline RL learning model. Second, consistent with previous findings (Bos et al., 2012), the dual learning rate models are generally better at capturing behavior than the single learning rate models. Third, the decay parameter does not improve model fit. Finally, social influence is best described by two separate (uncorrelated, $r_{pearson}$ =. 01, p = .99) processes: (1) an increased prior, and (2) a constant bonus for the recommended option. For further quality control we have simulated the behavior of the "prior + bonus, dual RL" model (see supplemental and Figure S2). For each age group we simulated 100 agents using the medians of the parameter fits for that group. The results of these simulations indicate that the model is adequately able to capture the qualitative difference in learning behavior seen in each age group (e.g. the initial peak in adolescence; see supplemental Figure S2). Finally, Kolmogorov-Smirnov tests indicated that the distributions of all of the model parameters deviate significantly from a normal distribution (all p's < .001). For subsequent analyses all parameters were transformed to approximate normality via the Box-Cox transformation.

Parameter Estimates – Prior + Bonus dual RL. To better understand the processes that underlie age difference in task behavior we further investigated age differences in parameter estimates (see Supplement, Table S4). We did not find any significant age trend for the learning rate associated with gains (α_{win}), but in line with previous studies (Bos et al., 2012) we found that the children had a higher learning rate for negative outcomes compared to adolescents and adults (α_{loss} : $\beta_{emerging}$= .28, t = 2.08, p = 0.041), indicating they are most sensitive to negative feedback (see Figure 3B). In addition, there was a significant linear increase in the temperature function (θ : β_{linear}=

.05, t = 3.03, p = 0.003), indicating that the older participants behaved less exploratory than the younger participants. With regards to social influence we found that the prior, but not the bonus, was showing age related differences. More specifically, adolescents' prior expectation based on recommendation was significantly larger than that of both children and adults (β_p : $\beta_{quadratic}$ = .14, t = 2.3, p = 0.025).

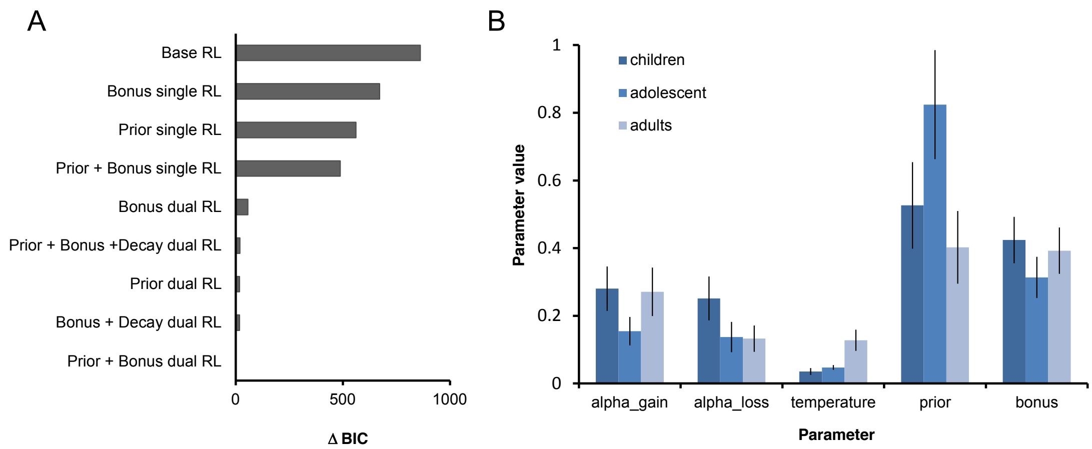

Figure 3. (A) *BIC's for model comparison*. The relative difference in BIC values for each model compared to the model with the lowest over BIC value. The Bayes factor for comparing the best (lowest BIC) and second best model is 6494, which indicates the best fitting model is very strongly favored over all other models tested. Note that this model also is the winning model if we perform these comparisons on the level of age groups separately. **(B)** *Parameter estimates for the prior + bonus dual RL model.* Estimates separately for the three age groups and the two learning rates (alpha_gain, alpha_loss), temperature, prior and bonus.

3.3. Exploratory Simulations

Next, we explored how these group level strategies would fare in different environments. More specifically, we explored an environment in which there was another deck that was better than the recommended deck, and one in which it was worse. Given that the adolescents seemed to be the only subjects who learned that there was another good deck in the original experiment, we expected their strategy to be specifically advantageous in the environment where there was a better alternative present.

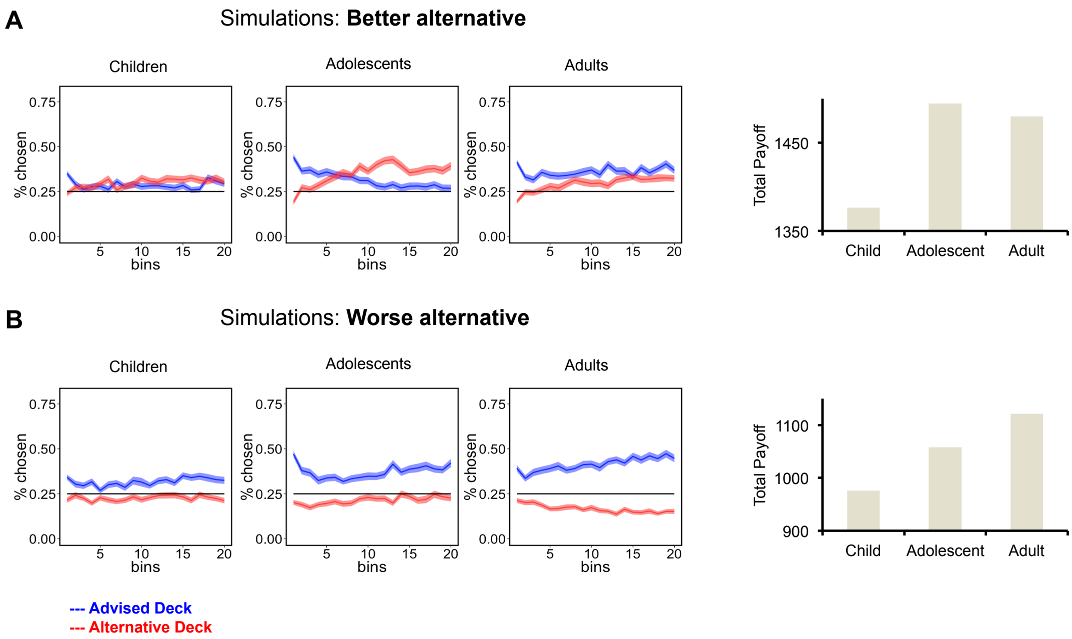

Figure 4. Simulated data using the best fitting model and the median parameter values for each group. The simulations are the result of 100 iterations and 210 trials. Shaded areas represent standard error. **(A)** *Simulation: Better alternative.* Simulations in context of when the other good deck was better than the advised deck ($EV_{advised}$ = 10; $EV_{alternative}$= 12.5) the blue line represents choices from the advised deck, red line represents the better alternative. Total payoff per age group for better alternative. Payoff is calculated by multiplying the % of choices per deck times the expected value (times number of trials (N=210)). When there is a better alternative adolescents have a slight advantage over the adults. **(B)** *Simulation: Worse alternative.* Simulations in which the advised deck was better than all other decks. Again, the blue line is the advised deck ($EV_{advised}$ = 10), red line represents the worse alternative ($EV_{alternative}$= 7.5). Total payoff per age group for worse alternative. Payoff is calculated by multiplying the % of choices per deck times the expected value (times number of trials (N=210)).

Indeed, as expected the simulations indicate that the adolescents are quicker at switching from the recommended to the optimal deck (see Figure 4A). However, this was not the case when the alternative was worse (see Figure 4B). In that case the exploratory behavior of the children and adolescents is costly (see Figure 4).

4. Discussion

In this age-comparative study we used a 4-armed bandit task (Biele et al., 2011) including an initial peer advice in children (8-10 year old), adolescents (13-15 year old) and young adults (18-22 year old). As expected all age groups followed the advice in the beginning of the task, however after a few rounds the behavior of the different age groups started to diverge. The most salient developmental differences suggest that (1) adolescents are initially the most sensitive to advice, (2) adults are most consistently following the advice, (3) younger participants' behavior is more exploratory and (4) children's sensitivity to negative feedback hampers their learning. Computational modeling helped to further describe these developmental differences and extended them by (5) showing higher exploration rates prior to adulthood and (6) confirming children's higher sensitivity to negative feedback. Moreover, developmental differences could be captured and simulated within a (7) social influence model: Our *prior + bonus dual RL model* gives first insights in possible underlying mechanism in advice-taking and exploring alternatives across development: First, our social learning model integrates different effects of advice as well as experience. Moreover, our social learning model is able to resolve the apparently conflicting findings of adolescents being more sensitive to peer influence- they are more willing to take risks in the presence of peers- but on the other hand explore and "re-evaluate" alternatives to the advice more quickly, compared to adults. Second, the model makes the interesting prediction that the adolescents' behavior may be optimal in certain environments. These novel findings are discussed in more detail below.

In line with previous developmental findings (Chein et al., 2011), adolescents showed the highest initial sensitivity to peer advice compared to children and adults. This suggests that social influence most strongly impacts adolescents' initial expectations (i.e. their priors). Thus our findings support the view that adolescence may be a developmental period with a particularly high sensitivity to social influence (Blakemore & Mills, 2014; Jones et al., 2014; van Hoorn et al., 2016). However, a recent study by Decker et al. 2015 also suggest that this influence does not have a long lasting effect. That is, teenagers may be easily swayed to try out something when it is suggested by their peers, such as skipping classes, but when this is not positively reinforced, they will also likely be the first age group to stop pursuing the suggested behavior. Adults on the other hand showed a more consistent influence of advice over time, which is adaptive in this experiment, but the data of Decker et al. 2015 suggest they will also do so when the advice is not good. Taken together these findings suggest a more nuanced view on developmental differences in advice taking suggesting that adolescence might be not only a unique period associated with higher peer-influence on behavior, but also with a healthy reliance on personal experiences.

In line with previous developmental studies (Decker et al., 2015) children and adolescents showed more exploratory behavior (see Figure 3B). Higher exploration rates have been linked to the protracted maturation of prefrontal cognitive control functions (Decker et al., 2015; Thompson-Schill, Ramscar, & Chrysikou, 2009). Furthermore, exploratory behavior in itself has many positive aspects, particularly in dynamic and unknown environments, and has been suggested to be an important adaptation in human development (Thompson-Schill et al., 2009). Indeed, in the current task the explorative behavior of the adolescents may have resulted in a benefit That is, adolescents increasingly selected the other good deck with learning and chose the bad decks less often than adults and children (see Figure 2). Interestingly, adolescents also showed the highest expected earnings compared to the other two age groups. This suggests that their exploration may have led them to find out more about the expected value of each of the decks and ended up using a strategy that led to higher earnings. Our simulations further support this hypothesis. That is, if the other good deck was associated with even higher expected values than the advised deck, adolescent's higher exploration behavior would lead to higher learning rates compared to adults and higher total payoffs within the task (see Figure 4). These findings highlight the importance of taking into account the structure of the environment when making normative statements about certain types of behavior. Future studies, using different learning environments, are needed to further explore the possible harms and benefits of adolescent vs. adult learning strategies.

Finally, although children also showed increased explorative behavior this did not result in choosing the other good deck more often (see 2). This might be the result of children's greater sensitivity to negative feedback compared to adolescents and adults (van Duijvenvoorde et al., 2008). Our modeling results support this view by showing higher learning rates for losses in children compared to the other two age groups (see Figure 3B). Previous studies suggested that children's performance decreases as the probability of negative feedback increases (Eppinger, Mock, & Kray, 2009). Thus, children's difficulties to use negative feedback for learning should be particularly salient in the current task, where each card deck was associated with 50% losses and 50% gains (although they differ in their magnitude; see Figure S1). In sum, although (negative) experience weighs more than advice for children, they are not able to benefit from their experience to the same degree as adolescents in this learning environment.

5. Conclusion

Taken together, our findings show that peer-advice guides learning from one's own experience, especially in adults, although adolescents show the highest initial sensitivity to peer advice. Crucially, higher exploration rates enable adolescents to discover other opportunities. Thus, our results extend previous findings by showing that their more explorative behavior can be – depending on the environmental structure – even more beneficial than the more adult-like learning strategies. This raises interesting questions regarding which possible features of the everyday environment of adolescents affords such exploratory behavior and highlights the need to understand the structure of the environment in which development takes place.

Acknowledgments

We would like to thank Valerie Keller, Raphael Schultz and Jann Wäscher for help during data acquisition, Lukas Nagel for programming the task and as well as our participants for their contribution to the study. This research was in part funded by the Freie Universität Berlin, and the Open Research Area grant (ORA ID176) awarded to Wouter van den Bos. Julia Rodriguez Buritica's doctoral fellowship was supported by the Berlin School of Mind and Brain and Freie Universität Berlin. We are grateful to the Center of Lifespan Psychology at the Max Planck Institute for Human Development for making data collections possible.

References

Behrens, T. E. J., Hunt, L. T., & Rushworth, M. F. S. (2009). The Computation of Social Behavior. *Science*, *324*(5931), 1160–1164.

Biele, G., Rieskamp, J., & Gonzalez, R. (2009). Computational Models for the Combination of Advice and Individual Learning. *Cognitive Science*, *33*(2), 206–242.

Biele, G., Rieskamp, J., Krugel, L. K., & Heekeren, H. R. (2011). The Neural Basis of Following Advice. *PLOS Biol*, *9*(6), e1001089.

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

Chein, J., Albert, D., O'Brien, L., Uckert, K., & Steinberg, L. (2011). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry: Peer influence on risk taking. *Developmental Science*, *14*(2), F1–F10.

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.

Eppinger, B., Mock, B., & Kray, J. (2009). Developmental differences in learning and error processing: Evidence from ERPs. *Psychophysiology*, *46*(5), 1043–1053.

Gold, J. M., Carpenter, C., Randolph, C., Goldberg, T. E., & Weinberger, D. R. (1997). Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Archives of General Psychiatry*, *54*(2), 159–165.

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.

Kahnt, T., Park, S. Q., Cohen, M. X., Beck, A., Heinz, A., & Wrase, J. (2009). Dorsal Striatal–midbrain Connectivity in Humans Predicts How Reinforcements Are Used to Guide Decisions. *Journal of Cognitive Neuroscience*, *21*(7), 1332–1345.

Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the Couplings Among Intellectual Abilities and Constituent Cognitive Processes Across the Life Span. *Psychological Science*, *15*(3), 155–163.

Lourenco, F. S., Decker, J. H., Pedersen, G. A., Dellarco, D. V., Casey, B. J., & Hartley, C. A. (2015). Consider the source: Adolescents and adults similarly follow older adult advice more than peer advice. *PloS One*, *10*(6), e0128047.

Meltzoff, A. N., Waismeyer, A., & Gopnik, A. (2012). Learning about causes from people: Observational causal learning in 24-month-old infants. *Developmental Psychology*, *48*(5), 1215–1228.

Monahan, K. C., Steinberg, L., & Cauffman, E. (2009). Affiliation with antisocial peers, susceptibility to peer influence, and antisocial behavior during the transition to adulthood. *Developmental Psychology*, *45*(6), 1520–1530.

Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, *431*(7010), 760–767.

Nielsen, M., & Tomaselli, K. (2010). Overimitation in Kalahari Bushman Children and the Origins of Human Cultural Cognition. *Psychological Science*.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, *28*(1), 78–106.

Steinberg, L., & Monahan, K. C. (2007). Age differences in resistance to peer influence. *Developmental Psychology*, *43*(6), 1531–1543.

Sumter, S. R., Bokhorst, C. L., Steinberg, L., & Westenberg, P. M. (2009). The developmental pattern of resistance to peer influence in adolescence: Will the teenager ever be able to resist? *Journal of Adolescence*, *32*(4), 1009–1021.

Sutton, R. S., & Barto, A. G. (1998). *Introduction to Reinforcement Learning* (1st ed.). Cambridge, MA, USA: MIT Press.

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition Without Control When a Little Frontal Lobe Goes a Long Way. *Current Directions in Psychological Science*, *18*(5), 259–263.

van den Bos, W., Cohen, M. X., Kahnt, T., & Crone, E. A. (2012). Striatum–Medial Prefrontal Cortex Connectivity Predicts Developmental Changes in Reinforcement Learning. *Cerebral Cortex*, *22*(6), 1247–1255.

van den Bos, W., & Eppinger, B. (2016). Developing developmental cognitive neuroscience: From agenda setting to hypothesis testing. *Developmental Cognitive Neuroscience*, *17*, 138–144.

van Duijvenvoorde, A. C. K. van, 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. *The Journal of Neuroscience*, *28*(38), 9495–9503.

van Hoorn, J., van Dijk, E., Meuwese, R., Rieffe, C., & Crone, E. A. (2016). Peer Influence on Prosocial Behavior in Adolescence. *Journal of Research on Adolescence*, *26*(1), 90–100.

Weiß, R. (2006). CFT 20-R. Grundintelligenztest Skala 2. Manual. Göttingen: Hogrefe Verlag.

Supplementary

1. Sample Size

To determine the number of participants for each of the age groups we have based ourselves on the study that most closely resembles the current design. In van den Bos et al., 2012 we have used a very similar task and also a very similar reinforcement learning model to analyze the data (except there was no advise). To estimate the number of subjects we need to find age differences in parameter estimates we have focused on the two effects of the learning rates reported (i.e. as alpha's) in that paper. First there was a significant effect of age on the negative learning rate (i.e. alpha loss) and a marginal significant effect on the positive learning rate (i.e. alpha gain) ($F_{2,67}$ = 9.87, $P < 0.001$ and $F_{2,67} = 2.73$, $P = 0.06$ respectively). This translates in Cohen's f of .53 and .27 (large and medium effect sizes). These effect sizes result in power estimates of .61 and .46. We used the R pwr toolbox to calculate the minimum number of subjects needed to find an effect at alpha of .05. These analyses suggest that we need between 21 subjects (per group) to for the medium effect size and 29 subjects per group for the large effect size. This, taken together with the observation that other similar studies (is this true, you could check the Decker studies, e.g., Decker et al., 2015) use similar group sizes, we decided to select 25 number of subjects for each group.

2. Choice behavior

In Table S1 to S3 the exact results of the logistic regressions for the choice behavior with respect to the advised deck (see Table S1), the other good deck (see Table S2) and the bad decks [i.e. collapsed over both bad decks (see Table S3), with the same payoff distributions (see Figure S1)] are reported:

Table S1. Summary of logistic regression for *choice behavior AD = intercept +* B_1*Age-Group +* B_2 *Age-Group* ^{2} *+* B_3*Trial +* B_4 *Age-Group * Trial +* B_5 *Age-Group* ^{2} ** Trial*

	Estimate	S.E.	*t* value	*p*
Intercept	0.338	0.004	87.325	<.001 ***
Age	0.055	0.007	8.299	<.001 ***
Age^2	0.011	0.007	1.751	0.08
Trial	-0.013	0.004	-3.235	0.001 **
Age * Trial	-0.005	0.007	-0.752	0.452
Age^2 * Trial	-0.019	0.007	-2.889	0.004 **

Table S2. Summary of logistic regression for *choice behavior GD = intercept +* B_1*Age-Group +* B_2 *Age-Group* ^{2} *+* B_3*Trial +* B_4 *Age-Group * Trial +* B_5 *Age-Group* ^{2} ** Trial*

	Estimate	S.E.	*t* value	*p*
Intercept	0.229	0.003	66.665	<.001 ***
Age	-0.011	0.006	-1.867	0.062
Age^2	0.036	0.006	6.037	<.001 ***
Trial	0.013	0.003	3.782	<.001 ***
Age * Trial	0.009	0.006	1.594	0.111
Age^2 * Trial	0.026	0.006	4.415	<.001 ***

Table S3. Summary of logistic regression for *choice behavior BD = intercept +* B_1*Age-Group +* B_2 *Age-Group* ^{2} *+* B_3*Trial +* B_4 *Age-Group * Trial +* B_5 *Age-Group* ^{2} ** Trial*

	Estimate	S.E.	*t* value	*p*
Intercept	0.433	0.004	106.88	<.001 ***
Age	-0.044	0.007	-6.342	<.001 ***
Age^2	-0.048	0.007	-6.791	<.001 ***
Trial	-0.0004	0.004	-0.118	0.906
Age * Trial	-0.004	0.007	-0.633	0.527
Age^2 * Trial	-0.007	0.007	-0.985	0.325

3. Simulations

For quality control we performed simulations using the medians of the parameter estimates for each age group (see Figure S2), using the best fitting model (prior + bonus, dual RL). For each age group we simulated 100 agents. These agents were presented with the original Decks and the pseudo-random ordering of feedback that was used in the behavioral experiment.

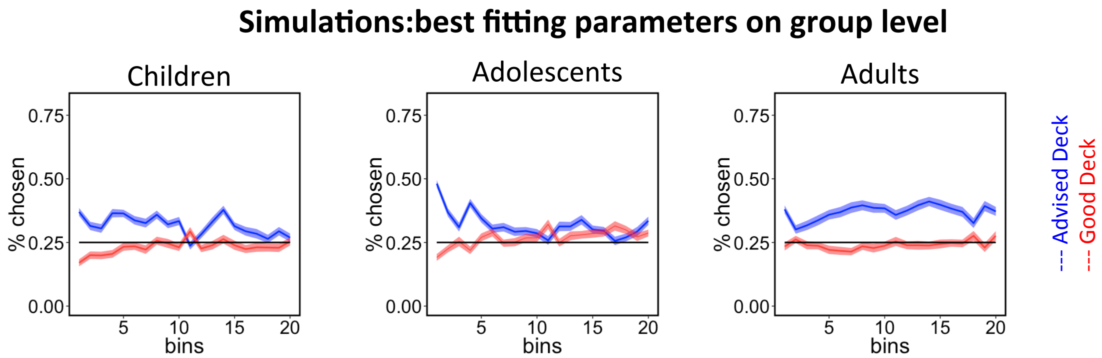

Figure S2. Simulated data using the best fitting model and the median parameter values for each group. The simulations are the result of 100 iterations and 210 trials. Shaded areas represent standard error, blue line is the advised deck, red line the other good deck.

The simulations indicate that the best fitting model is able to capture the main behavioral difference between age groups; the initial peak of the advice effect in adolescence, adolescents choosing the other good deck more than the other age groups, and the overall increase in performance with age (see Figure S2).

5. Parameter Estimates – Prior + Bonus dual RL

Table S4 shows age differences in the in parameter estimates of the “prior + bonus dual RL” model.

Table S4. Age differences in parameter estimates.

	Age	Age^2	Age^3
Parameter			
α_{gain}	-.002	.019	.006
α_{loss}	.139	.158	.289*
theta	.054**	-.001	.079*
prior	-.049	.144*	.002
bonus	-.029	.110	-.101

Note:	* p<0.05; ** p<0.01; *** p<0.001

6. Recovery Analyses.

To further explore the validity of the reinforcement learning models and the model selection procedure we performed model recovery analyses. For these analyses we selected the winning model (dual RL + prior + bonus) and extracted the median parameter values of the fitting results of the total subject pool. These parameters, that are the most representative, were then used to generate behavior as described in the simulation section. We have simulated the behavior for 100 subjects and then used all dual learning rate reinforcement models to fit those 100 simulated subjects. As expected the generative model (dual RL + prior + bonus) showed the best fit (see Table S5).

Table S5. BIC's for simulated data.

Dual RL models	BIC
Bonus + prior	25047
Prior	25231
Bonus	25246
Bonus + prior + decay	25950
Bonus + decay	26449

Furthermore, the extract the parameter fits for the 100 randomly simulated datasets show the same patterns, and the populations medians are all within the 95% confidence interval (see Figure S3). Together these findings support the validity of the model specification and selection.

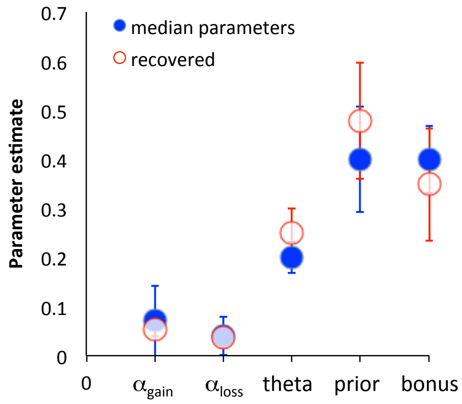

Figure S3. Parameter estimates for the median and recoverd parameters. Error bars represent 95% confidence interval.