Evolution and Human Behavior 38 (2017) 325–333

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.ehbonline.org

Original Article

Individual differences in learning behaviours in humans: Asocial exploration tendency does not predict reliance on social learning

Wataru Toyokawa [a,b,*,1], Yoshimatsu Saito [b,c], Tatsuya Kameda [c]

[a] Department of Behavioural Science, Hokkaido University, N10W7, Kita-ku, Sapporo 060-0810, Japan
[b] Japan Society for the Promotion of Science, 5-3-1 Kojimachi, Chiyoda-ku, Tokyo 102-0083, Japan
[c] Department of Social Psychology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan 113-0033

ARTICLE INFO

Article history:
Initial receipt 10 June 2016
Final revision received 10 November 2016

Keywords:
Social learning
Individual differences
Reinforcement learning
Two-armed bandit
Copy-when-uncertain

ABSTRACT

A number of empirical studies have suggested that individual differences in asocial exploration tendencies in animals may be related to those in social information use. However, because the 'exploration tendency' in most previous studies has been measured without considering the information-gathering processes, it is yet hard to conclude that the animal asocial exploration strategies may be tied to social information use. Here, we studied human learning behaviour in both asocial and social two-armed bandit tasks. By fitting reinforcement learning models including asocial and/or social decision processes, we measured each individual's (1) asocial exploration tendency and (2) social information use. We found consistent individual differences in the exploration tendency in the asocial tasks. We also found substantive heterogeneity in the adopted learning strategies in the social task: Nearly one-third of participants used predominantly the copy-when-uncertain strategy, while the remaining two-thirds were most likely to have relied only on asocial learning. However, we found no significant individual association between the exploration frequency in the asocial task and the use of the social information in the social task. Our results suggest that the social learning strategies may be independent from the asocial exploration strategies in humans.

1. Introduction

To find better behavioural options in foraging, mate choice, nest search, etc., group living animals can benefit from asocial information-gathering strategy (e.g., reinforcement learning rules (Sutton & Barto, 1998; Trimmer, McNamara, Houston, & Marshall, 2012)) and from strategic use of social information (Boyd & Richerson, 1985; Laland, 2004). Although there has been much recent interest in the inter-individual variation in both asocial and social learning behaviour (Mesoudi, Chang, Dall, & Thornton, 2016; Reader, 2015), little is known about whether and (if so) how they associate with each other.

Individual differences in asocial exploration tendency might be related to different individual optimums in the exploration-exploitation trade-off. Given the limited time/energy budget, a single animal must strike the right balance between trying unfamiliar behaviours to sample information (i.e., 'exploration') versus choosing known best behaviour (i.e., 'exploitation') so as to improve the long-term net decision performance (Cohen, McClure, & Yu, 2007; Hills et al., 2014). The optimal balance of exploration-exploitation depends on the costs and benefits of information gathering, which may differ between individuals. For example, an individual with poor information processing performance may have lower benefits of exploration, an individual with shorter expected life-span may benefit less from sampling more information, while an individual experiencing a temporary volatile environment may be forced to explore so as to update its knowledge (Reader, 2015).

On the other hand, the individual variation in reliance on social information might come from the balance of cost and benefit of copying others (Mesoudi et al., 2016). For instance, an individual possessing inaccurate private information will potentially incur a large cost if relying solely on the private knowledge and hence may tend to copy others more (e.g., 'copy-when-uncertain' (Laland, 2004; Rendell et al., 2011)), an individual living in a large group may benefit more from following the majority (King & Cowlishaw, 2007), while an individual faced with a highly volatile environment may rely more on private information due to the potentially large cost from copying an out-of-date behaviour (Aoki & Feldman, 2014).

Some factors may simultaneously affect the individual differences in both domains: Environmental volatility may increase asocial exploration tendency while decreasing copying tendency. On the other hand, a common cognitive ability underlying both asocial and social learning

* Corresponding author.
E-mail addresses: wt25@st-andrews.ac.uk (W. Toyokawa), middle158725071@gmail.com (Y. Saito), tkameda@l.u-tokyo.ac.jp (T. Kameda).
[1] Present address: School of Biology, University of St Andrews, St Andrews, Fife, UK, KY16 9TH.

http://dx.doi.org/10.1016/j.evolhumbehav.2016.11.001
1090-5138/

may generate a positive correlation between them (Mesoudi et al., 2016). Indeed, the increasing body of empirical studies has shown both negative and positive correlations between the asocial exploration tendency and the social information use (but see Webster & Laland, 2015). For instance, the individual exploration propensity negatively correlates with the individual tendency of copying conspecifics in barnacle geese *Branta leucopsis* (Kurvers, Prins, van Wieren, van Oers, Nolet and Ydenberg, 2010; Kurvers, van Oers, Nolet, Jonker, van Wieren, Prins and Ydenberg, 2010) and zebra finches *Taeniopygia guttata* (Rosa, Nguyen, & Dubois, 2012), while the opposite is true in three-spined sticklebacks *Gasterosteus aculeatus* (Nomakuchi, Park, & Bell, 2009) and great tits *Parus major* (Marchetti & Drent, 2000).

However, the term 'exploration' has not been explicitly defined as information-gathering behaviour in the previous literature, and might have been confounded with other personality traits (reviewed in Réale, Reader, Sol, McDougall, & Dingemanse, 2007). Broadly speaking, more active, neophilic, or bolder individuals tend to be labelled as 'explorative' while more inactive, neophobic, or shyer individuals tend to be labelled as 'unexplorative' (Réale et al., 2007). However, it was untested whether individuals labelled as 'explorative' actually gather information more during the learning process compared to those labelled as 'non-explorative' (Carere & Locurto, 2011; Groothuis & Carere, 2005; Koolhaas et al., 1999). Therefore, it remains unclear whether the individual differences in asocial information-gathering strategy might associate with those in social information use.

In this study, we focused on human learning behaviour in a two-armed bandit (2AB) problem, and tested whether the individual differences in asocial exploration tendency predicted the reliance on social learning. Because the possible individual correlation between asocial exploration and social learning could be either positive or negative (Mesoudi et al., 2016), we did not make any specific prediction about the direction of the correlation.

The 2AB is the most basic test-bed problem of reinforcement learning (Sutton & Barto, 1998). Therefore, we were able to fit a computational reinforcement learning model to each participant's decision data so as to estimate individual information-gathering processes (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Keasar, Rashkovich, Cohen, & Shmida, 2002; O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003; Racey, Young, Garlick, Pham, & Blaisdell, 2011; Toyokawa, Kim, & Kameda, 2014). In the task, individuals have two choice options, but at the outset they do not have exact knowledge of which option is more profitable (Fig. 1a). However, they can update their knowledge of the options through the experiences of earning rewards. Fitting the learning model, we can infer the knowledge-updating process for each participant, so as to categorise each decision made by each participant into either exploitation (i.e., choosing the option with higher estimated reward value as of that round; see the Material and methods section) or exploration (i.e., choosing the other option with lower estimated reward value). The 'exploration' measured in this study, therefore, directly relates to information-gathering behaviour during the reinforcement learning.

In addition to the asocial situation where participants engaged in the 2AB task alone (hereafter, 'solitary task'), participants also played the 2AB task in a pairwise situation ('paired task') in which they were able to observe the other participant's choice (but not the peer's earned payoff) displayed on the monitor. To examine whether the participants adopted social learning strategy in the paired task, we fitted several asocial- and social-learning models to each participant's decision data, and then selected the most likely learning model individually. Also, we analysed each participant's gaze movement measured by an eye-tracker in order to confirm the participant's information use during the task. Finally, we examined whether the exploration tendency in the solitary task (i.e., asocial exploration) predicted the use of social information in the paired task.

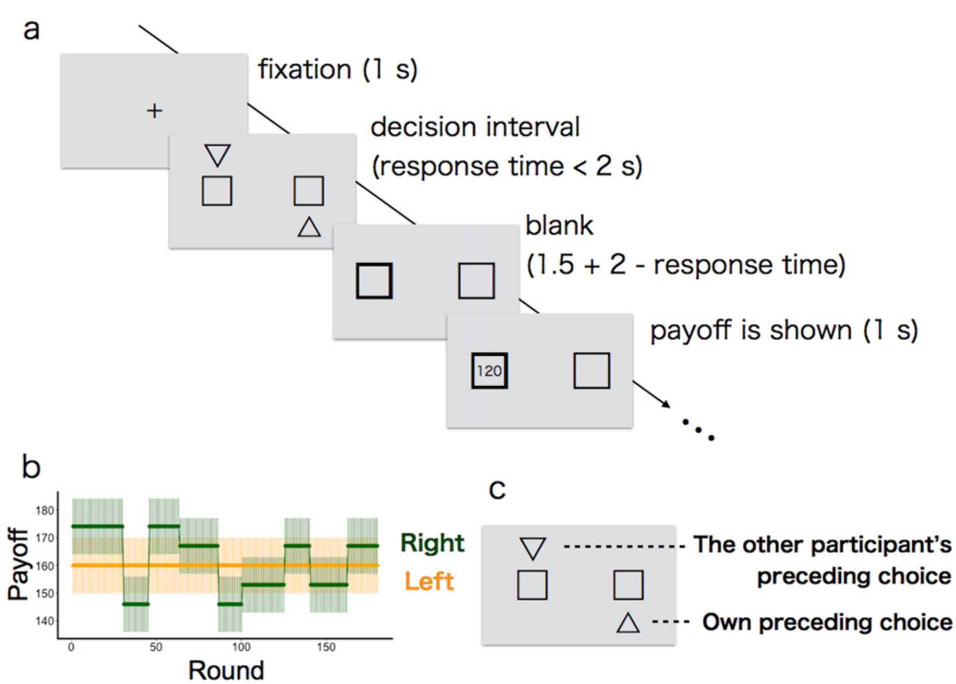

Fig. 1. The restless two-armed bandit task. (a) Illustration of the time line within a round. After fixation with a crossbar for 1 s, two slots (boxes) were presented. The participant had to choose one within 2 s (decision interval). When choosing one, the frame of the chosen option turned to be bold. Up to 3.5 s later, the number of payoff points earned was revealed to the participant (120 points in this example). After a further second, the next round started with a crossbar. (b) Example of mean payoffs for each option in paired task (bold lines). The payoff received for a particular choice is drawn from a Gaussian noise around each mean (shaded areas show 1 S.D. of this noise). Note that the most profitable slot ('optimal option') was switched several times due to the volatility; only one box was volatile and the other box's mean was fixed. The left-right location of the volatile box was counterbalanced across pairs in the paired task (the right box was volatile in this example). (c) Example of social information in the paired task. The other participant's choice in the preceding round was marked by a downward triangle (left in this example) while their own choice in the preceding round was marked by an upward triangle (right in this example).