

VU Research Portal

Reward-based motor adaptation can generalize across actions

van der Kooij, Katinka; Smeets, Jeroen B.J.

published in

Journal of Experimental Psychology: Learning Memory and Cognition
2019

DOI (link to publisher)

[10.1037/xlm0000573](https://doi.org/10.1037/xlm0000573)

document version

Publisher's PDF, also known as Version of record

document license

Article 25fa Dutch Copyright Act

[Link to publication in VU Research Portal](#)

citation for published version (APA)

van der Kooij, K., & Smeets, J. B. J. (2019). Reward-based motor adaptation can generalize across actions. *Journal of Experimental Psychology: Learning Memory and Cognition*, 45(1), 71-81.
<https://doi.org/10.1037/xlm0000573>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Reward-Based Motor Adaptation Can Generalize Across Actions

Katinka van der Kooij and Jeroen B. J. Smeets
Vrije Universiteit Amsterdam

Recently it has been shown that rewarded variability can be used to adapt visuomotor behavior. However, its relevance seems limited because adaptation to binary rewards has been demonstrated only when the same movement is repeated throughout the experiment. We therefore investigated whether the adaptation is action-specific and whether the amount of exploration depends on spatial complexity. Participants pointed to 3-D visual targets without seeing their hand and could use only binary reward feedback to adapt their movements. We varied the number of target positions and the number of dimensions the feedback was based on. Because the feedback was based on a 5-cm rightward shifted hand position, adaptation was needed for good performance. The participants started naïve to the perturbation. If actions were made toward a single target position and the feedback was based on the lateral component of their response only, participants adapted completely within 200 trials. Having more than 1 target position or more than 1 dimension of performance resulted in considerably less adaptation but did not affect the exploration. Thus, reward-based adaptation can generalize across actions but is reduced by spatial complexity, whereas exploration is not affected by spatial complexity.

Keywords: motor learning, reward, exploration, visuomotor, adaptation

Supplemental materials: <http://dx.doi.org/10.1037/xlm0000573.supp>

When one performs a goal-directed movement toward a visual target, the brain needs to transform the target position into a motor command. This visuomotor transformation cannot be fixed, because the body, the environment, and the tools one uses change. For this, one can rely on visuomotor adaptation: updating the transformation based on feedback about the movement (see Bastian, 2008; Huberdeau, Krakauer, & Haith, 2015; Krakauer & Mazzoni, 2011; Lackner & DiZio, 2005, for reviews). It has long been known that visuomotor adaptation relies on learning from feedback about error: differences between the movement outcome and the predicted (or intended) outcome (Baddeley, Ingram, & Miall, 2003; Benson, Anguera, & Seidler, 2011; Burge, Ernst, & Banks, 2008; Hinder, Riek, Tresilian, de Rugy, & Carson, 2010; Krakauer, 2009; van Beers, 2009). More recently, it has been shown that feedback about success (reward) may be sufficient to adapt the movement plans (Cashaback, McGregor, Mohatarem, & Gribble, 2017; Izawa & Shadmehr, 2011; Therrien, Wolpert, & Bastian, 2016). The scope of the contribution of reward-based processes to visuomotor adaptation is unclear, however, because it has primarily been demonstrated in a very specific paradigm.

How the brain processes feedback (reward and error) in visuomotor adaptation is generally studied in a center-out reaching paradigm in which participants make planar reaching movements, while performance feedback is based on a hand position that is rotated around the center (e.g., Cohen, 1967; Fernández-Ruiz, Díaz, Aguilar, & Hall-Haro, 2004; Mazzoni & Krakauer, 2006; Redding & Wallace, 1988). Experiments using this paradigm have shown that error-based and reward-based motor adaptation have complementary qualities (Huberdeau et al., 2015; Krakauer & Mazzoni, 2011; Manley, Dayan, & Diedrichsen, 2014; Wolpert, Diedrichsen, & Flanagan, 2011). Adaptation to errors is fast, smooth, and transient (Burge et al., 2008; Hinder et al., 2010; Smith, Ghazizadeh, & Shadmehr, 2006; van der Kooij, Overvliet, & Smeets, 2016), although explicit adaptation is more variable than is implicit adaptation (Benson et al., 2011; Mazzoni & Krakauer, 2006; Taylor, Krakauer, & Ivry, 2014). Reward-based adaptation is slow, erratic, and persistent (Hasson, Manczurowsky, & Yen, 2015; Izawa & Shadmehr, 2011; Nikooyan & Ahmed, 2015; Therrien et al., 2016). The scope of the reward-based adaptation is unclear because it has mainly been demonstrated in single-target tasks and because the exploratory variability it depends on hasn't been studied thoroughly in the context of visuomotor adaptation.

The scope of reward-based adaptation may be more limited than that of error-based adaptation because it requires many repetitions of the same movement plan, whereas error-based adaptation does not. In other words, reward-based adaptation may be *action-specific*, whereas error-based adaptation may be *action-independent*. Error-based adaptation is qualitatively quite similar when moving to a single target position (Izawa & Shadmehr, 2011), to eight target positions (Hinder et al., 2010), or to a different target position on each trial (van der Kooij, Brenner, van Beers, Schot, & Smeets, 2013). Binary reward feedback, in con-

This article was published Online First April 26, 2018.

Katinka van der Kooij and Jeroen B. J. Smeets, Department of Human Movement Sciences, Vrije Universiteit Amsterdam.

We thank Jasper van de Lagemaat for testing participants. The research was funded by Stichting Technologie en Wetenschap (STW) Open Technologie Programma (OTP) Grant 12,668.

The content is solely the responsibility of the authors and does not necessarily reflect the official views of Stichting Technologie en Wetenschap.

Correspondence concerning this article should be addressed to Katinka van der Kooij, Department of Human Movement Sciences, Vrije Universiteit Amsterdam, van der Boechorststraat 9, 1081 BT Amsterdam, the Netherlands. E-mail: k.vander.kooij@vu.nl

trast, does not lead to adaptation when participants move to a different target position on each trial (van der Kooij & Overvliet, 2016). Also, the generalization of adaptation depends on the type of feedback. Error-based adaptation has been found to generalize to untrained target positions (Izawa & Shadmehr, 2011), to the opposite limb (Choe & Welch, 1974; Taub & Goldberg, 1973; Taylor, Wojaczynski, & Ivry, 2011), and even from walking to pointing (Morton & Bastian, 2004). Adaptation to binary reward, on the other hand, has been found to transfer poorly to untrained target positions (Izawa & Shadmehr, 2011). However, adaptation to a more informative reward (one that depends on the size of the error) has been found to generalize to nearby target positions (Nikooyan & Ahmed, 2015). More evidence that reward-based adaptation is action-specific is provided by an elegant study by Huang, Haith, Mazzoni, and Krakauer (2011). They investigated the hypothesis that faster readaptation to the same perturbation (savings) is due to reward-based adaptation. They showed that savings depended on the repetition of target position in a hand-centered coordinate frame, which suggested that the adaptation that produced savings was tied to the trained movement plan. To assess the extent to which error-based and reward-based adaptation are action-specific or action-independent, Izawa and Shadmehr (2011) tested whether the adaptation affects proprioceptive estimates of hand position. They found that error-based adaptation affects proprioceptive estimates of hand position, whereas reward-based adaptation does not, again suggesting that the reward-based adaptation is action-specific.

Although it is tempting to interpret reward-based adaptation as inherently action-specific, it is not by definition the case. One can regard reward-based adaptation as a combination of exploration and learning from rewarded exploration (Therrien et al., 2016). This description in principle allows for learning general properties of an environment, as long as these general properties are explored. In this view, the different levels of adaptation that have been reported for single-target position paradigms (Izawa & Shadmehr, 2011; Therrien et al., 2016) and a multitarget position paradigm (van der Kooij & Overvliet, 2016) may have been caused by participants' exploring less efficiently in the multitarget case. The amount and direction of exploration has, for instance, been found to depend on the task (Wu, Miyamoto, Gonzalez Castro, Ölveczky, & Smith, 2014) and on the statistical properties of the feedback (Pekny, Izawa, & Shadmehr, 2015; Sidarta, Vahdat, Bernardi, & Ostry, 2016; Vaswani et al., 2015). In experiments that employ more than one target position, exploration may be less efficient than in single-target experiments because the search problem becomes more complex with additional target positions. There may be a position-specific cause for the reward, a general cause, or a combination of the two. It has, for instance, been found that awareness of the presence of a perturbation is associated with greater exploration along the task-relevant dimension (Manley et al., 2014), suggesting that participants explore more efficiently when they know where to search. If exploration is less efficient due to the complexity of moving to multiple targets, other forms of complexity should also hamper exploration and thereby reward-based adaptation. For instance, performing a 3-D task rather than a 1-D single-target position ($1-D_1$) task may reduce reward-based adaptation.

In this study, we used a 3-D visuomotor adaptation paradigm to assess how reward-based adaptation and exploration depend on

two forms of spatial complexity: moving to more than one target position and performing a one-dimensional or three-dimensional task. To test whether suggesting an overall property to explore enhances exploration and adaptation, we first assessed adaptation to feedback based on a perturbed hand position in a naïve phase in which participants were not informed about the presence of a perturbation, and subsequently participants were informed that there was a perturbation and repeated the adaptation.

Method

Participants

In total, 65 participants who declared to be healthy (32 female, 59 right-handed, ages 24.5 years \pm 6.1) took part in the study. They were randomly assigned to one of four groups in which we tested the influences of the number of target positions and the number of dimensions the feedback was based on: a three-dimensional distance in space or a one-dimensional lateral distance. We call this number the "task dimensionality." There were three one-dimensional groups that pointed to one, three, or six target positions ($1-D_1$: eight male, eight female, mean age = 24.1; $1-D_3$: 10 male, six female, mean age = 24.8; and $1-D_6$: 10 male, six female, mean age = 25.6). In addition, there was a three-dimensional group that pointed to a single-target position ($3-D_1$: five male, 12 female, mean age = 23.6). Stereovision was assessed with the Stereo Fly Test (Western Ophthalmics, Lynwood, WA). Based on this test, three participants (one in the $3-D_1$ group and two in the $1-D_6$ group) with a stereo acuity $>100''$ were excluded from participation. The methods were approved by the local ethical committee and adhered to the Declaration of Helsinki (World Medical Association, 2013). Participants provided written informed consent before participating.

Materials

We used an HTC Vive for visual display (resolution 2160×1200 for each eye; frame rate 90 Hz) and motion tracking (90 Hz, <1 -mm precision), and the experiment was programmed using the Unity 3-D game engine. Participants held the Vive controller (a handle and a ring with markers that allowed tracking their movements; see Figure 1C). When referring to the position of the controller, we generally refer to the point where the ring touched the handle.

Procedure

When participants entered the room, they could view a scoreboard displaying the scores of other participants. The scoreboard was used to motivate participants to pay attention to the reward score that would be provided later in the game. Participants first performed the Stereo Fly Test, with which stereovision was assessed. After that, eye distance was measured with a ruler, and the distance between the lenses of the head mounted display (HMD) was adjusted accordingly. Participants were instructed to place the thumb of their dominant hand on an inactive button on the controller that served as a reference. They were told that a target fly ($5 \times 2 \times 5$ cm, including the wings) would appear in front of them and that their task was to "catch" the fly by touching it with their

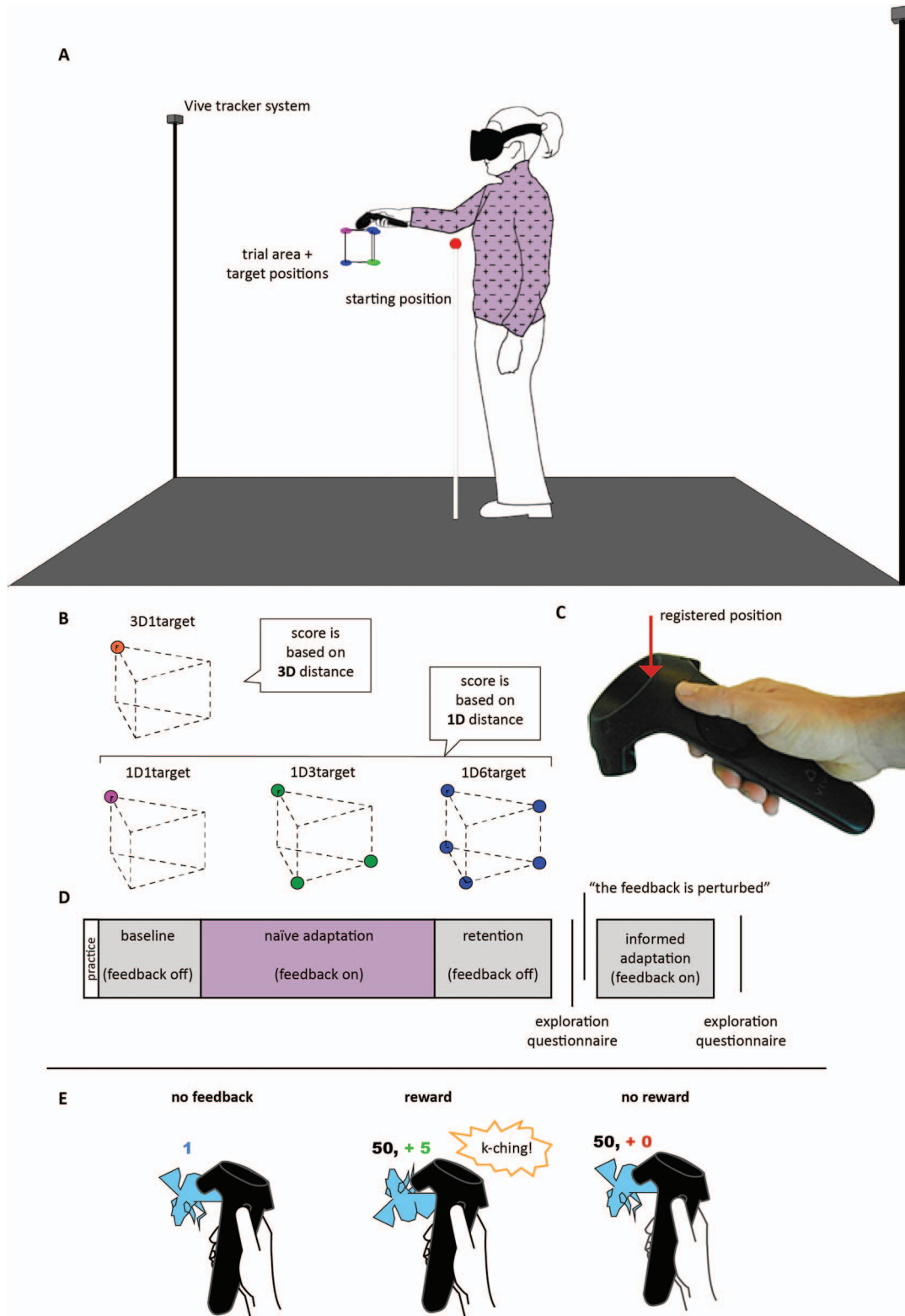


Figure 1. Methods used in the study. Panel A: Experimental set-up with HTC Vive, virtual starting position and targets. Panel B: Target positions that were used in the different groups. Panel C: Handheld controller with the registered position indicated. Panel D: Time course of the four experimental phases (baseline, naïve adaptation, retention, informed adaptation) and questionnaires (data shown the Supplementary material; Figure S1 and Figure S2). Panel E: Feedback types used in the phases without feedback ('no feedback') and in the phases with feedback ('reward', 'no reward'). See the online article for the color version of this figure.

thumb and holding their hand still for a moment before returning the controller to the start position. Participants were informed that they would perform about 200 trials, starting with a baseline phase in which they received no performance feedback, and that after that they would be able to score points and would finish with another phase without performance feedback. Participants in the 3-D₁ group were told that they would score points based on the 3-D distance between the thumb and fly, whereas participants in the other three groups were told that they could focus on the lateral dimension because scores were exclusively based on the lateral component of their end points. This provided information about the dimension to explore but left the participant naïve with respect to the presence of a perturbation. To introduce a constant bias that participants could adapt to, we based feedback on a 5-cm rightward-shifted center of the controller's ring (comparable to a 5° azimuth rotation around the cyclopean eye).

Once the participants had put on the HMD, they were asked to position themselves behind a virtual pole (height = 1.5 m) with a red knob, which was the starting position for their pointing movements (see Figure 1A). They were asked to look straight ahead, and after the experimenter pressed a button, the HMD's vertical position relative to the ground was recorded as the participants' height. The starting position was at 75% of the participants' height, and a prism-shaped volume (14-cm equilateral triangle as the base and 14 cm high) in which targets could appear was centered 40 cm straight in front of the starting position (see Figure 1B). The six corners of this volume could serve as target positions.

In the 3-D₁ and 1-D₁ groups, the target was always presented at the far highest point of the prism. In the 1-D₃ group, two additional targets were presented: one at the near-left corner and one at the near-right corner. In the 1-D₆ group, all six corners of the prism were used. The targets were presented in blocks of six trials. For the 1-D₃ and 1-D₁ groups, the target position was always the same, whereas for the 1-D₃ and 1-D₆ groups, there were no immediate repetitions of target position. For the 1-D₃ group, the set of three target positions was used twice in a pseudorandom order, and for the 1-D₆ group, the six target positions were used once in a pseudorandom order.

To start the first trial, participants were instructed to touch the starting position with the controller. Then the first target fly appeared with the trial number (cumulative across phases) indicated above it. At this moment, participants could initiate a movement with the controller toward the fly. An end point was registered when the velocity of the controller dropped below 3 cm/s. The detection of an end point was signaled to participants with the trial number's changing color, and depending on the phase (described in the next paragraph), feedback was provided. After that, participants returned the controller behind the starting position. Once the controller was behind the starting position, a new target appeared. An average trial took about 280 ms.

After a practice block of six target presentations in which veridical visual feedback of the controller was shown continuously to let participants get acquainted with the task, there were four phases without direct vision of the controller (see Figure 1D): a *baseline phase* (seven blocks), a *naïve adaptation phase* (16 blocks), a *retention phase* (eight blocks), and an *informed adaptation phase* (eight blocks). During the baseline phase, no performance feedback was provided. After the baseline phase, a text was displayed encouraging the participant to start scoring points and a

cumulative score initiated at zero was displayed next to the trial number. With this, the naïve adaptation phase started.

In the naïve adaptation phase, binary performance feedback was provided based on the end-point error. For participants in the 3-D₁ group, the end-point error was the distance between the end point of the controller and the center of the target fly, whereas for the other groups, the end-point error was the lateral distance between the end point of the controller and the center of the target fly. We used a shaping paradigm (Skinner, 1938) to ensure that participants with different levels of performance would receive a similar amount of reward. Trials were rewarded when the end-point error was smaller than a moving average of the previous five trials or smaller than 30% of the median error in the baseline phase. The score was calculated in the adaptation phases exclusively. In the multiple-target groups, we used a single running average for all target positions. The reward consisted of an animation of the fly dying, five scored points, the trial number coloring green, and a rewarding sound. When a trial was not rewarded, the fly disappeared without any sound and the trial number turned red.

The end of the naïve adaptation phase was indicated with a text that announced that the retention phase started, and the cumulative score was removed from the scene. In the retention phase, again no performance feedback was provided. After that, participants viewed their score and took off the headset to complete a short questionnaire, which we administered to get an idea of the exploration strategies employed by participants.

After the questionnaire was completed, we informed participants that the feedback had been based on a horizontally shifted hand position and that they could use this information to improve their score. After that, they put on the HMD again, and the informed adaptation phase started. This phase was identical to the naïve adaptation phase. When they had finished, the experimenter wrote their score on a sticky note and added it to the scoreboard. The experiment finished with another administration of the short exploration questionnaire.

Data Analysis

All data (including statistics) were processed and analyzed using custom-written MatLab R2017a software (MatLab, 2017). The basis of all analyses was the end-point error: the difference vector between the end point of the controller and the position of the target. Based on the end-point error, we calculated two parameters: the adaptation index and the exploration factor, as explained next.

To calculate the adaptation index, we first subtracted the bias at target position t in the baseline phase (the mean end-point error in the baseline phase for that target position) to remove natural biases (van der Kooij et al., 2013). The lateral component of the resulting value (e_x) was normalized by the lateral difference between the position-specific baseline bias at target position t and the 5-cm lateral perturbation:

$$\text{adaptation index} = \frac{e_x^t}{\text{perturbation} - \text{bias}_x^t}$$

This way, an adaptation index of zero would indicate no adaptation and an adaptation index of one would indicate complete adaptation.

We defined the *asymptotic adaptation* as the mean adaptation index in the last block of a phase. For the 3-D₁ group, we also analyzed the 3-D asymptotic adaptation, which was based on the 3-D adaptation index. The 3-D adaptation index was calculated using the same procedure as for the adaptation index. However, the 3-D adaptation index was based on the amplitude of the 3-D end-point error. The 3-D asymptotic adaptation was the mean 3-D adaptation index in the last block of a phase.

To separate exploration from other sources of variability in the data, we reasoned that in the baseline phase, participants had no reason to explore. Therefore, the variability in this phase would be a result of planning noise, execution noise, and sensory noise, not active exploration (which we refer to as “noise”). We therefore used the variability in this phase as an estimate of noise. Exploration in search of reward during the adaptation phases adds variability, so we used the increase in variability in the adaptation phases as a proxy for exploration. We could not use the standard deviation as a measure of variability because the expected adaptation to the perturbed feedback would contribute to this measure. We therefore used the amplitude of the change in the 3-D end-point error between two subsequent trials (Δ) as our measure of variability. For uniformly and normally distributed errors, Δ is about 2.2 times the standard deviation in one dimension. The contribution of exploration to the variability was calculated in two steps. We first subtracted the contribution of noise (the median Δ in the baseline phase). Second, we divided the resulting value by the noise and obtained the *exploration factor*. $\Delta\eta_x$ was the mean exploration factor in an adaptation phase. $\Delta\eta_x$ was the lateral component of the exploration factor, $\Delta\eta_y$ the vertical component, and $\Delta\eta_z$ the depth component. We expressed the exploration as a fraction of the noise because reward-based adaptation depends on a balance between noise and exploration rather than on the exploration per se (Therrien et al., 2016). A difference with the Therrien et al. (2016) approach is that we determined the amount of active exploration aimed at obtaining reward, whereas Therrien et al. defined exploration as all motor variability that the system has access to which may or may not contain additional components.

Because outliers are difficult to distinguish from exploration, no outliers were removed; all data were analyzed nonparametrically such that the analyses were robust for outliers. We used Wilcoxon’s rank sum tests for between-groups comparisons and Wilcoxon’s sign rank tests for within-group comparisons. Statistical tests were aimed at testing predictions concerning the action-specificity of adaptation and exploration. We tested three predictions that assessed how adaptation depends on the number of target positions and the spatial complexity.

Adaptation and spatial complexity.

Number of target positions. First, based on the idea that reward-based adaptation is action-specific, we predicted that the asymptotic adaptation depends negatively on the number of target positions (invoking actions in different directions). The influence of the number of target positions on the asymptotic adaptation was tested by entering the asymptotic adaptation in the naïve adaptation phase of the 1-D₁, 1-D₃, and 1-D₆ groups in a Kruskal–Wallis test. Post hoc comparisons compared the asymptotic adaptation between the individual groups using Wilcoxon’s rank sum tests that tested whether the asymptotic adaptation was always greater in the group that pointed to fewer target positions ($1-D_1 > 1-D_3$; $1-D_1 > 1-D_6$; $1-D_3 > 1-D_6$).

Dimensionality. Second, if spatial complexity, rather than the number of target positions per se, affects adaptation, one would expect the adaptation to also be reduced by another form of complexity. We therefore predicted that the asymptotic adaptation depends negatively on the task dimensionality. This prediction was tested by examining whether the asymptotic adaptation in the naïve adaptation phase was greater in the 1-D₁ group compared to the 3-D₁ group using a Wilcoxon’s rank sum test. Because the reward criterion in the 3-D₁ group was based on the 3-D error, which was generally higher than the lateral error, we additionally tested whether the adaptation asymptote in the naïve adaptation phase of the 1-D₁ group was higher than the 3-D adaptation asymptote in the naïve adaptation phase of the 3-D₁ group using a Wilcoxon’s rank sum test.

Adaptation and information. Third, based on the idea that informing participants about the presence of a perturbation increases the efficiency of exploration, we predicted that the asymptotic adaptation increases in the informed adaptation phase. This prediction was tested by comparing the asymptotic adaptation in the naïve and informed phase using a Wilcoxon signed-ranks test. We also tested, in addition to these predictions, whether the asymptotic adaptation decreased in the retention phase using a Wilcoxon signed-ranks test.

The analyses of exploration first tested whether there was a significant exploration factor ($\Delta\eta_x > 0$) using a Wilcoxon signed-ranks test. After that, we tested how the exploration factor depends on three factors: spatial complexity, information about the presence of a perturbation, and the reward history (Pekny et al., 2015; Sidarta et al., 2016).

Exploration and spatial complexity.

Number of target positions. To assess whether the exploration factor depended on the spatial complexity, we first tested whether it depended on the number of target positions, by performing a Kruskal–Wallis test on the $\Delta\eta_x$ in the naïve adaptation phase of the participants in the 1-D₁, 1-D₃, and 1-D₆ groups.

Dimensionality. Whether task dimensionality affects the lateral component of the exploration factor was assessed by comparing the mean $\Delta\eta_x$ in the naïve adaptation phase of the 1-D₁ and 1-D₃ groups using a Wilcoxon’s rank sum test. To examine whether the exploration factor is larger when a dimension is task-relevant, we used the fact that the vertical and depth dimensions were task-relevant for the 3-D₁ group but not for the 1-D₁ group. Using Wilcoxon’s rank sum tests on $\Delta\eta_y$ and $\Delta\eta_z$, we tested whether the mean vertical and depth components of the exploration factor in the naïve adaptation phase were larger for the participants in the 3-D₁ group than for the participants in the 1-D₁ group.

Exploration and information. To test how information about the presence of a perturbation affects the exploration factor, we compared the exploration factor for the naïve adaptation phase with that for the informed adaptation phase. Whether the lateral exploration increased in the informed adaptation phase was tested by comparing the participant’s mean $\Delta\eta_x$ in the naïve adaptation phase to the mean lateral factor $\Delta\eta_x$ in the informed adaptation phase using a Wilcoxon signed-ranks test. In addition, we tested whether the mean vertical and depth exploration factors ($\Delta\eta_y$ and $\Delta\eta_z$, respectively) differed between the naïve and informed adaptation phases using a Wilcoxon signed-ranks test.

Exploration and reward. To test whether lateral exploration increases after a nonrewarded trial (Pekny et al., 2015; Sidarta et

al., 2016), we compared the participants' mean lateral exploration factor following nonrewarded trials with the mean lateral exploration factor following rewarded trials in the naïve adaptation phase and in the informed adaptation phase using Wilcoxon signed-ranks test.

Results

Adaptation Decreases With Spatial Complexity and Increases With Information

As expected, the adaptation index increased with block number (see Figure 2A). The increase depended on the experimental group: The most adaptation occurred for the 1-D₁ group. We planned three statistical analyses to compare the amount of adaptation. Our first planned statistical analysis (comparing the three one-dimensional groups) confirmed the prediction that the adaptation asymptote depended on the number of target positions, $\chi^2(45, N = 46) = 9.9, p = .007, \eta^2 = .22$ (see Figure 2B). Post hoc comparisons showed that the adaptation asymptote was higher in the 1-D₁ group compared to the adaptation asymptote in the 1-D₃ group ($Mdns = .95$ and $.31$, respectively; $z = 2.97, p = .003, r = .52$). The adaptation asymptote in the 1-D₃ group ($Mdn = .34$) did not significantly differ from the adaptation asymptote in the 1-D₆ group ($z = -1.16, p = .25, r = .20$) and also did not differ from the adaptation asymptote in the 1-D₆ group ($z = 1.51, p = .13, r = .27$).

Our second planned statistical analysis (concerning the dependency on task dimensionality) confirmed the second prediction, that the asymptotic adaptation would be higher in the 1-D₁ group compared to the 3-D₁ group ($Mdns = .95$ and $.07$, respectively; $z = 3.15, p < .001, r = .55$). The adaptation asymptote in the 1-D₁ group was also higher than the 3-D adaptation asymptote in the 3-D₁ group ($Mdns = .95$ and $.19$, respectively; $z = 4.13, p < .001, r = .72$), indicating that less adaptation in the 3-D₁ group was not due to testing the lateral adaptation, whereas the reward criterion was based on a 3-D distance. Thus, the adaptation depended negatively on the number of target positions and on the task dimensionality.

Our third planned statistical analysis (comparing the informed and retention phases with the naïve adaptation phase) showed that the asymptotic adaptation did not decrease from the naïve adaptation phase to the retention phase ($Mdns = .44$ and $.28$, respectively; $z = -.24, p = .81, r = .02$) and increased from the naïve adaptation phase to the informed adaptation phase ($Mdns = .44$ and $.89$, respectively; $z = 2.09, p = .036, r = .18$; see Figure 2B). In other words, the retention was good (in line with the reported persistence of reward-based adaptation), and the information about the perturbation improved adaptation, in line with our prediction. Moreover, in the informed adaptation phase, the adaptation did not depend on the number of target positions, $\chi^2(45, N = 46) = .72, p = .69, \eta^2 = .02$, or the dimensionality ($z = -1.02, p = .30, r = .17$).

We additionally checked whether participants in the one-dimensional groups with more than one target did not simplify their task by responding to the lateral position of the targets only. For the 1-D₃ and 1-D₆ groups, we correlated the controllers' end points and target positions in the naïve adaptation phase. We found

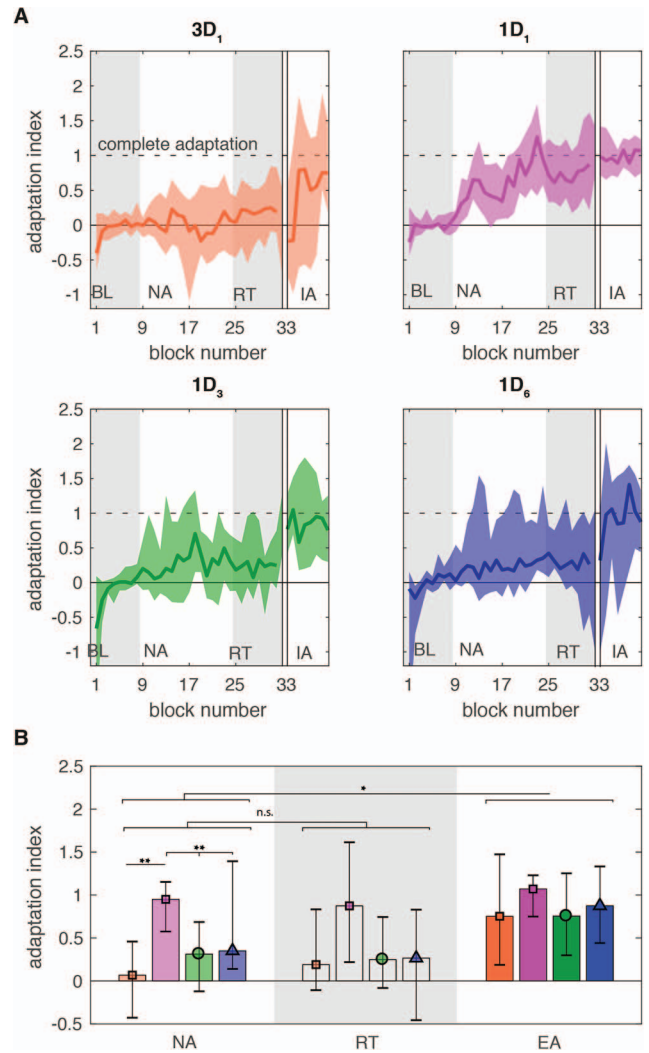


Figure 2. Adaptation. Overview of the median adaptation with interquartile range. Panel A: Adaptation index as a function of block number for the four different one- and three-dimensional groups—3-D₁ (one target position), 1-D₁ (one target position), 1-D₃ (three target positions), and 1-D₆ (six target positions)—in the different phases (BL = baseline; NA = naïve adaptation; reaction time [RT] = retention; IA = informed adaptation). Shaded areas indicate the interquartile range, light-gray vertical sections indicate episodes without feedback. Dashed lines indicate the level at which adaptation would be complete whereas solid lines indicate the level at which there would be no adaptation. Panel B: Asymptotic adaptation in the different phases. Comparisons that we tested statistically are indicated with a horizontal bar. Error bars indicate the interquartile range, the vertical shaded area indicates an episode without performance feedback. * $p < .05$. ** $p < .01$. See the online article for the color version of this figure.

that the end points and target positions in both the 1-D₃ and 1-D₆ target groups were significantly correlated in all dimensions ($p < .01$). For the 1-D₃ target group, the Pearson's R for the correlation between the end points and targets were $.83, .98$, and $.28$ for the lateral, vertical, and depth components, respectively. For the 1-D₆ target group, these values were $.83, .96$, and $.41$,

respectively. This shows that participants followed the target positions in all dimensions.

Exploration Does Not Depend on Spatial Complexity

We found that in the naïve adaptation phase, there was clear exploration ($z = 6.6, p < .001, r = .83$). We planned three comparisons on the exploration concerning spatial complexity, information about the presence of a perturbation, and the reward history.

Although Figure 3A suggests that the exploration is larger in the 1-D₃ group compared to the other groups, this apparent effect of the number of target positions on the lateral exploration factor $\Delta\eta_x$ in the naïve adaptation phase was not significant, $\chi^2(45, N = 46) = 4.19, p = .12, \eta^2 = .09$. The lateral exploration factor in the naïve adaptation phase also did not differ between the 3-D₁ and 1-D₁ groups ($z = .09, p = .93, r = .02$), which indicated that the exploration factor did not depend on task dimensionality. Whether exploration is larger when a dimension is task-relevant was assessed by comparing the vertical exploration factor $\Delta\eta_y$ and depth exploration factor $\Delta\eta_z$ for the 3-D₁ group, for whom these dimensions were task-relevant, with the 1-D₁ group, for whom these dimensions were task-irrelevant (see Figure 3A, right section). It is interesting that neither the vertical exploration factor nor the depth exploration factor was greater in the 3-D₁ group compared to the 1-D₁ group ($z = .95, p = .34, r = .17$, and $z = -0.31, p = .76, r = .05$, respectively). Moreover, the exploration factor was very similar for the task-relevant and task-irrelevant directions.

Exploration Increases With Information and Reward Absence

Information about the presence of lateral perturbation influenced the exploration factor (see Figure 3B). The perturbation-relevant lateral exploration factor significantly increased ($z = 2.26, p = .009, r = .23$; see Figure 3B, left section), whereas the exploration factor for the perturbation irrelevant dimensions decreased ($\Delta\eta_y: z = -4.45, p < .001; \Delta\eta_z: z = -3.59, p < .001, r = .32$; see Figure 3B, right section).

The history of reward also influenced the exploration (see Figure 3C, left section). In the naïve adaptation phase, the lateral exploration factor following nonrewarded trials was larger than that following rewarded trials ($z = 4.39, p < .001, r = .39$). This finding was replicated in the informed adaptation phase ($z = 4.74, p < .001, r = .42$; see Figure 3C, right section).

Discussion

The aim of this study was to test how reward-based adaptation and exploration are affected by two forms of spatial complexity: the number of target positions and the task dimensionality (the number of dimensions that the feedback was based on). In addition, we tested how information about the presence of a perturbation influences reward-based adaptation.

Adaptation Depends on Spatial Complexity, but Exploration Does Not

When participants moved to a single target position and received performance feedback based on a single dimension (known

by the participant), they adapted completely within the 96 trials of the naïve adaptation phase and retained almost everything they had learned during 48 trials without performance feedback (see Figure 2). The amount of adaptation depended on the number of target positions. When they were naïve about the perturbation, the groups that pointed to three target positions adapted less (about one third) than did the group that pointed to a single target location. The group that pointed to six target locations on average adapted as much as did the group with three target locations, but for this group, the difference with the group that pointed to a single target position was not statistically significant. When (known by the participant) the reward depended on the 3-D error rather than on only the lateral error, there was also less adaptation compared to when participants performed a one-dimensional task. When informed about the nature of the perturbation, adaptation increased.

For the exploration, in contrast to the adaptation, there was no statistical evidence that it depended on the spatial complexity: The lateral exploration did not significantly depend on the number of target positions or on the task dimensionality (see Figure 3A). Informing the participants that we had introduced a lateral perturbation had a clear effect on exploration: The perturbation-relevant exploration increased, whereas the exploration in the perturbation-irrelevant dimensions decreased (see Figure 3B). As has been previously reported in the literature (Pekny et al., 2015; Sidarta et al., 2016), the exploration was larger following nonrewarded trials than following rewarded trials (see Figure 3C).

Thus, our main finding is that reward-based adaptation depends on spatial complexity, whereas exploration does not. The finding that naïve adaptation was possible only when participants moved to a single target is consistent with reports in the literature that reward-based adaptation occurs only when participants repeatedly move to the same target position (Huang et al., 2011; Izawa & Shadmehr, 2011) and is also consistent with the finding that there was no reward-based adaptation in a 3-D task in which participants moved more than one target position (van der Kooij & Overvliet, 2016).

Our study extends knowledge on reward-based adaptation in two ways. First, our findings show that visuomotor adaptation was also not possible when participants moved to a single target in a three-dimensional task. Second, we showed that providing explicit information about the presence of a perturbation made reward-based adaptation possible even when participants moved to multiple targets. This is consistent with findings that explicit instructions to counter a perturbation allow participants to learn a more complex perturbation using error-based information (Bedford, 1993). One could argue that because the naïve adaptation phase was performed before the informed adaptation phase, increased adaptation in the informed phase could also have been due to learning in the previous phase, in line with faster relearning of the same adaptation, a phenomenon called *savings* (Huang et al., 2011; Krakauer, 2009; Zarah, Weston, Liang, Mazzoni, & Krakauer, 2008). This is unlikely, because there was no evidence of learning in the naïve phase for the multitarget and multidimensional groups.

The results on exploration equally replicate and extend findings in the literature. The finding that the exploration decreased when a participant had just been rewarded (see Figure 3C) replicates findings that exploration is dependent on the reward history (Pekny et al., 2015; Sidarta et al., 2016). In the context of motor adaptation, exploration has not been properly related to spatial

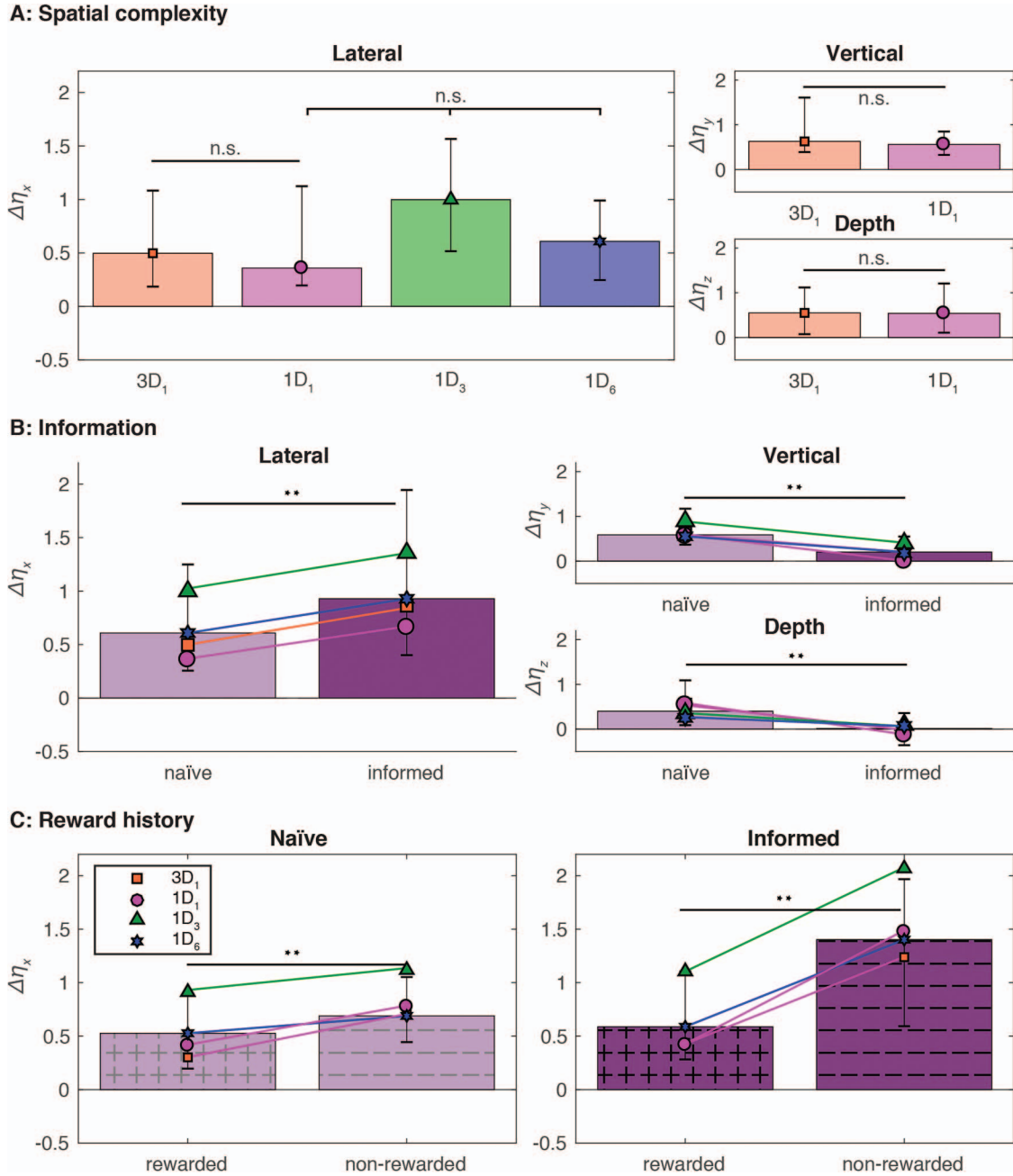


Figure 3. Exploration. Active exploration. The exploration factor $\Delta\eta$ is the mean additional variability induced by the feedback, expressed as fraction of the variability without feedback. Presented are medians with interquartile range. Panel A, left section: Task-relevant dimension. Lateral exploration factor $\Delta\eta_x$ in the naive adaptation phase for the four different groups: 3-D₁ (1 - D task, one target position), 1-D₁ (1 - D task, one target position), 1-D₃ (1 - D task, three target positions), and 1-D₆ (1 - D task, six target positions). Panel A, right section: Task relevance versus task irrelevance. Vertical exploration factor ($\Delta\eta_y$) and depth exploration ($\Delta\eta_z$) in the naive adaptation phase of the 3-D₁ group (dimensions are task-relevant) and for the 1-D₁ group (dimensions are task-irrelevant). Panel B: The perturbation-relevant lateral exploration factor increases in the informed adaptation phase (left section). The exploration factor in the perturbation-irrelevant dimensions decreased in the informed adaptation phase (right section). Panel C: The exploration factor was larger following nonrewarded trials than following rewarded trials. Error bars indicate the interquartile range. * $p < .05$. ** $p < .01$. See the online article for the color version of this figure.

complexity yet. Our study suggests that exploration does not depend on spatial complexity (see Figure 3A). Moreover, the pattern of exploration did not match the pattern of adaptation in the different groups in Figure 2B. The 1-D single-target group adapted the most, yet there was no evidence that this group explored the lateral dimension more than did other groups, because it showed the lowest exploration factor. Therefore, it is unlikely that the influence of spatial complexity on the adaptation was mediated by differences in exploration. Next we discuss what the results on adaptation and exploration mean for current theories of reward-based motor adaptation.

Reinforcement of Successful Movement Plans

The finding that the adaptation in the informed phase did not depend on the number of target positions contrasts with the idea that reward-based motor adaptation is action-specific. Reward-based adaptation is considered action-specific because the rewards are held to reinforce the executed movement plan, not some higher order process (Dayan, Averbeck, Richmond, & Cohen, 2014; Huang et al., 2011; Izawa & Shadmehr, 2011). This idea is supported by findings that reward-based learning does not generalize across target positions (Izawa & Shadmehr, 2011) and does not occur without repetition of the same movement direction (Huang et al., 2011). Such an action-specific learning does not explain adaptation in the informed stage of our study. Therefore, at least in the informed phase, another mechanism of reward-based learning must have contributed to the adaptation. Moreover, action-specificity of learning does not explain why participants who performed a three-dimensional, single-target task were unable to adapt. In the next section, we discuss an alternative mechanism for reward-based learning.

Learning From Rewarded Exploration

Recently an alternative type of reward-based adaptation has been proposed. In this type of learning, participants update their adaptation with rewarded exploration rather than with rewarded movement plans (Therrien et al., 2016). In the model proposed in that article, the adaptation is updated following rewarded trials only, because reward indicates that one made an exploration in the right direction, whereas reward absence does not provide information about the direction of the target: One may have overshot it or explored in the wrong direction. This type of learning could allow for action-independent learning, which could explain the action-independent learning we observed in the informed stage of our experiment. However, the model also needs to be extended in two ways to explain the full pattern of data.

Spatial complexity. First, the gain of learning may depend on spatial complexity. In the naïve adaptation phase, both multidimensional and multitarget groups adapted much less than did the single target, one-dimensional group. This suggests that the lack of adaptation may have been due to spatial complexity rather than to action dependency of learning. Spatial complexity might affect the adaptation because it may obscure credit assignment and participants may weight their reward-based learning by trust in the source of errors (Chen, Mohr, & Galea, 2017). To learn optimally from rewarded exploration, participants need to determine whether the reward was caused by noise or by exploration and whether they are exploring the property that caused the errors.

For the credit assignment that determines which aspect of the movement caused the reward, it has been shown that participants can rapidly determine whether curvature or direction is rewarded (Dam, Kording, & Wei, 2013). In a spatially more complex task, participants also need to determine whether the cause of errors is specific to a spatial location or consistent across spatial locations. Determining this may require a much larger number of trials. Until credit assignment is solved, participants may use a conservative learning gain, updating the adaptation with only a portion of the successful exploration. Similarly, when performing a 3-D task, participants need to determine the dimensions that caused the errors. Information about the presence of a perturbation solved the credit assignment problem and may thereby have facilitated the reward-based learning. Another way in which spatial complexity may affect reward-based learning is by affecting perceptual noise, which has been found to reduce the gain of visuomotor adaptation (Burge et al., 2008). For instance, pointing to target positions that differed in three dimensions may have made it more difficult to discriminate exploration from noise.

Reward-dependent exploration. Second, the data show that the exploration was not constant, as assumed in the model of Therrien et al. (2016), but was larger following nonrewarded trials (as also observed by Pekny et al., 2015; Sidarta et al., 2016). Thus, participants may update the adaptation following rewarded trials, whereas they may enhance the exploration following nonrewarded trials. The benefit of exploring more when being unsuccessful is that precision can be maintained on the successful trials. The aim of motor control processes is to minimize not only biases but also variability (Harris & Wolpert, 1998; Thorp, Kording, & Mussa-Ivaldi, 2017). Exploration facilitates bias reduction but obstructs minimization of variability. One way in which the brain may balance the costs and benefits of exploration is by exploring only when necessary.

We defined *exploration* slightly differently from how Therrien et al. (2016) did, which may explain the finding that the exploration in our study was reward-dependent, whereas the exploration in the Therrien et al. (2016) study could be explained by a constant exploration. Therrien et al. (2016) defined *exploration* as all the noise that participants have access to, whereas we defined it as noise that was evoked by the prospect of obtaining reward. Therefore our study focused on active exploration, which may be more sensitive to reward-based strategies, whereas the Therrien et al. (2016) definition may also encompass accidental noise sources.

Explicit and Implicit Processes

Overall, the sensitivity of adaptation to explicit information about the presence of a perturbation and the reward-dependent active exploration seems most consistent with an explicit learning process. Such an explicit process may be a decision-making process, not a sensorimotor process. It has for instance been shown that participants' performance in a decision-making task can be used to predict performance on a reward-based visuomotor adaptation task (Chen et al., 2017). However, there was no evidence that explicit awareness of the perturbation was necessary to adapt in the naïve adaptation phase. Following the naïve and explicit stage, we administered an explorative questionnaire in which participants were asked to draw where they aimed relative to the target fly (see Figure S2 in the online supplemental materials). On

this questionnaire, only 35% of the participants in the one-dimensional single-target group indicated that they aimed to the left of the target. It is interesting that the other 65% percent of the participants adapted almost completely to the perturbation, whereas the participants who indicated that they aimed to the left tended to overshoot the adaptation by almost a 100%.

Probably, a combination of implicit and explicit process contributes to the reward-based adaptation, as has been found for error-based adaptation (Benson et al., 2011; Bond & Taylor, 2015; Mazzoni & Krakauer, 2006; Redding & Wallace, 2006; Taylor et al., 2014). An interesting question for future research is whether implicit and explicit processes of reward-based motor adaptation are fundamentally different. For instance, implicit learning may rely on reinforcement of successful movement plans (Izawa & Shadmehr, 2011), whereas explicit learning may rely on learning from rewarded exploration (Therrien et al., 2016). Alternatively, both processes may occur on a combination of implicit and explicit levels of processing.

Conclusion

Together the results show that the scope of reward-based learning in visuomotor adaptation is not limited to learning from repetition of the same movement. Adaptation in multitarget tasks may be hindered by spatial complexity, resulting in problems with credit assignment, rather than by the lack of movement repetition. Exploration did not depend on the spatial complexity but was reduced on rewarded trials and increased after suggesting a dimension to explore.

References

- Baddeley, R. J., Ingram, H. A., & Miall, R. C. (2003). System identification applied to a visuomotor task: Near-optimal human performance in a noisy changing task. *Journal of Neuroscience*, *23*, 3066–3075.
- Bastian, A. J. (2008). Understanding sensorimotor adaptation and learning for rehabilitation. *Current Opinion in Neurobiology*, *21*, 628–633. <http://dx.doi.org/10.1097/WCO.0b013e328315a293>
- Bedford, F. L. (1993). Perceptual and cognitive spatial learning. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 517–530. <http://dx.doi.org/10.1037/0096-1523.19.3.517>
- Benson, B. L., Anguera, J. A., & Seidler, R. D. (2011). A spatial explicit strategy reduces error but interferes with sensorimotor adaptation. *Journal of Neurophysiology*, *105*, 2843–2851. <http://dx.doi.org/10.1152/jn.00002.2011>
- Bond, K. M., & Taylor, J. A. (2015). Flexible explicit but rigid implicit learning in a visuomotor adaptation task. *Journal of Neurophysiology*, *113*, 3836–3849. <http://dx.doi.org/10.1152/jn.00009.2015>
- Burge, J., Ernst, M. O., & Banks, M. S. (2008). The statistical determinants of adaptation rate in human reaching. *Journal of Vision*, *8*, 20. <http://dx.doi.org/10.1167/8.4.20>
- Cashaback, J. G. A., McGregor, H. R., Mohatarem, A., & Gribble, P. L. (2017). Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLoS Computational Biology*, *13*, e1005623. <http://dx.doi.org/10.1371/journal.pcbi.1005623>
- Chen, X., Mohr, K., & Galea, J. M. (2017). Predicting explorative motor learning using decision-making and motor noise. *PLoS Computational Biology*, *13*, e1005503. <http://dx.doi.org/10.1371/journal.pcbi.1005503>
- Choe, C. S., & Welch, R. B. (1974). Variables affecting the intermanual transfer and decay of prism adaptation. *Journal of Experimental Psychology*, *102*, 1076–1084. <http://dx.doi.org/10.1037/h0036325>
- Cohen, M. M. (1967). Continuous versus terminal visual feedback in prism aftereffects. *Perceptual and Motor Skills*, *24*, 1295–1302. <http://dx.doi.org/10.2466/pms.1967.24.3c.1295>
- Dam, G., Kording, K., & Wei, K. (2013). Credit assignment during movement reinforcement learning. *PLoS ONE*, *8*, e55352. <http://dx.doi.org/10.1371/journal.pone.0055352>
- Dayan, E., Averbeck, B. B., Richmond, B. J., & Cohen, L. G. (2014). Stochastic reinforcement benefits skill acquisition. *Learning & Memory*, *21*, 140–142. <http://dx.doi.org/10.1101/lm.032417.113>
- Fernández-Ruiz, J., Díaz, R., Aguilar, C., & Hall-Haro, C. (2004). Decay of prism aftereffects under passive and active conditions. *Cognitive Brain Research*, *20*, 92–97. <http://dx.doi.org/10.1016/j.cogbrainres.2004.01.007>
- Harris, C. M., & Wolpert, D. M. (1998, August 20). Signal-dependent noise determines motor planning. *Nature*, *394*, 780–784. <http://dx.doi.org/10.1038/29528>
- Hasson, C. J., Manczurovsky, J., & Yen, S.-C. (2015). A reinforcement learning approach to gait training improves retention. *Frontiers in Human Neuroscience*, *9*, 459. <http://dx.doi.org/10.3389/fnhum.2015.00459>
- Hinder, M. R., Riek, S., Tresilian, J. R., de Rugy, A., & Carson, R. G. (2010). Real-time error detection but not error correction drives automatic visuomotor adaptation. *Experimental Brain Research*, *201*, 191–207. <http://dx.doi.org/10.1007/s00221-009-2025-9>
- Huang, V. S., Haith, A., Mazzoni, P., & Krakauer, J. W. (2011). Rethinking motor learning and savings in adaptation paradigms: Model-free memory for successful actions combines with internal models. *Neuron*, *70*, 787–801. <http://dx.doi.org/10.1016/j.neuron.2011.04.012>
- Huberdeau, D. M., Krakauer, J. W., & Haith, A. M. (2015). Dual-process decomposition in human sensorimotor adaptation. *Current Opinion in Neurobiology*, *33*, 71–77. <http://dx.doi.org/10.1016/j.conb.2015.03.003>
- Izawa, J., & Shadmehr, R. (2011). Learning from sensory and reward prediction errors during motor adaptation. *PLoS Computational Biology*, *7*, e1002012. <http://dx.doi.org/10.1371/journal.pcbi.1002012>
- Krakauer, J. W. (2009). Motor learning and consolidation: The case of visuomotor rotation. *Advances in Experimental Medicine and Biology*, *629*, 405–421. http://dx.doi.org/10.1007/978-0-387-77064-2_21
- Krakauer, J. W., & Mazzoni, P. (2011). Human sensorimotor learning: Adaptation, skill, and beyond. *Current Opinion in Neurobiology*, *21*, 636–644. <http://dx.doi.org/10.1016/j.conb.2011.06.012>
- Lackner, J. R., & DiZio, P. (2005). Motor control and learning in altered dynamic environments. *Current Opinion in Neurobiology*, *15*, 653–659. <http://dx.doi.org/10.1016/j.conb.2005.10.012>
- Manley, H., Dayan, P., & Diedrichsen, J. (2014). When money is not enough: Awareness, success, and variability in motor learning. *PLoS ONE*, *9*, e86580. <http://dx.doi.org/10.1371/journal.pone.0086580>
- MatLab. (2017). [Computer Software]. Natick, MA: Mathworks, Inc.
- Mazzoni, P., & Krakauer, J. W. (2006). An implicit plan overrides an explicit strategy during visuomotor adaptation. *Journal of Neuroscience*, *26*, 3642–3645. <http://dx.doi.org/10.1523/JNEUROSCI.5317-05.2006>
- Morton, S. M., & Bastian, A. J. (2004). Prism adaptation during walking generalizes to reaching and requires the cerebellum. *Journal of Neurophysiology*, *92*, 2497–2509. <http://dx.doi.org/10.1152/jn.00129.2004>
- Nikooyan, A. A., & Ahmed, A. A. (2015). Reward feedback accelerates motor learning. *Journal of Neurophysiology*, *113*, 633–646. <http://dx.doi.org/10.1152/jn.00032.2014>
- Pekny, S. E., Izawa, J., & Shadmehr, R. (2015). Reward-dependent modulation of movement variability. *Journal of Neuroscience*, *35*, 4015–4024. <http://dx.doi.org/10.1523/JNEUROSCI.3244-14.2015>
- Redding, G. M., & Wallace, B. (1988). Components of prism adaptation in terminal and concurrent exposure: Organization of the eye-hand coordination loop. *Perception & Psychophysics*, *44*, 59–68. <http://dx.doi.org/10.3758/BF03207476>

- Redding, G. M., & Wallace, B. (2006). Generalization of prism adaptation. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1006–1022. <http://dx.doi.org/10.1037/0096-1523.32.4.1006>
- Sidarta, A., Vahdat, S., Bernardi, N. F., & Ostry, D. J. (2016). Somatic and reinforcement-based plasticity in the initial stages of human motor learning. *Journal of Neuroscience*, *36*, 11682–11692. <http://dx.doi.org/10.1523/JNEUROSCI.1767-16.2016>
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York, NY: Appleton-Century.
- Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biology*, *4*, e179. <http://dx.doi.org/10.1371/journal.pbio.0040179>
- Taub, E., & Goldberg, L. A. (1973). Prism adaptation: Control of intermanual transfer by distribution of practice. *Science*, *180*, 755–757. <http://dx.doi.org/10.1126/science.180.4087.755>
- Taylor, J. A., Krakauer, J. W., & Ivry, R. B. (2014). Explicit and implicit contributions to learning in a sensorimotor adaptation task. *Journal of Neuroscience*, *34*, 3023–3032. <http://dx.doi.org/10.1523/JNEUROSCI.3619-13.2014>
- Taylor, J. A., Wojaczynski, G. J., & Ivry, R. B. (2011). Trial-by-trial analysis of intermanual transfer during visuomotor adaptation. *Journal of Neurophysiology*, *106*, 3157–3172. <http://dx.doi.org/10.1152/jn.01008.2010>
- Therrien, A. S., Wolpert, D. M., & Bastian, A. J. (2016). Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. *Brain: A Journal of Neurology*, *139*, 101–114. <http://dx.doi.org/10.1093/brain/awv329>
- Thorp, E. B., Kording, K. P., & Mussa-Ivaldi, F. A. (2017). Using noise to shape motor learning. *Journal of Neurophysiology*, *117*, 728–737. <http://dx.doi.org/10.1152/jn.00493.2016>
- van Beers, R. J. (2009). Motor learning is optimally tuned to the properties of motor noise. *Neuron*, *63*, 406–417. <http://dx.doi.org/10.1016/j.neuron.2009.06.025>
- van der Kooij, K., Brenner, E., van Beers, R. J., Schot, W. D., & Smeets, J. B. J. (2013). Alignment to natural and imposed mismatches between the senses. *Journal of Neurophysiology*, *109*, 1890–1899. <http://dx.doi.org/10.1152/jn.00845.2012>
- van der Kooij, K., & Overvliet, K. E. (2016). Rewarding imperfect motor performance reduces adaptive changes. *Experimental Brain Research*, *234*, 1441–1450. <http://dx.doi.org/10.1007/s00221-015-4540-1>
- van der Kooij, K., Overvliet, K. E., & Smeets, J. B. J. (2016). Temporally stable adaptation is robust, incomplete and specific. *European Journal of Neuroscience*, *44*, 2708–2715. <http://dx.doi.org/10.1111/ejn.13355>
- Vaswani, P. A., Shmuelof, L., Haith, A. M., Delnicki, R. J., Huang, V. S., Mazzoni, P., . . . Krakauer, J. W. (2015). Persistent residual errors in motor adaptation tasks: Reversion to baseline and exploratory escape. *Journal of Neuroscience*, *35*, 6969–6977. <http://dx.doi.org/10.1523/JNEUROSCI.2656-14.2015>
- Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience*, *12*, 739–751.
- World Medical Association. (2013). *WMA declaration of Helsinki: Ethical principles for medical research involving human subjects*. Retrieved from <http://www.who.int/bulletin/archives/79%284%29373.pdf>
- Wu, H. G., Miyamoto, Y. R., Gonzalez Castro, L. N., Ólveczky, B. P., & Smith, M. A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature Neuroscience*, *17*, 312–321. <http://dx.doi.org/10.1038/nn.3616>
- Zarahn, E., Weston, G. D., Liang, J., Mazzoni, P., & Krakauer, J. W. (2008). Explaining savings for visuomotor adaptation: Linear time-invariant state-space models are not sufficient. *Journal of Neurophysiology*, *100*, 2537–2548. <http://dx.doi.org/10.1152/jn.90529.2008>

Received July 3, 2017

Revision received January 19, 2018

Accepted January 19, 2018 ■

Supplemental Materials

Reward-based motor adaptation can generalize across actions

by **K. van der Kooij & J. B. J. Smeets, 2019,**

Journal of Experimental Psychology: Learning, Memory, and Cognition **45**:71-81

<http://dx.doi.org/10.1037/xlm0000573>

Supplementary Material.

To explore the strategies that participants used, we administered two short exploration questionnaires (Table 1s): one following the naïve stage and one following the informed stage. The responses showed some expected results. For instance, those participants who performed a 3D task explored the vertical dimension more than the participants who performed a 1D task and that participants who performed a 3D task reduced the vertical exploration when they had been informed that there was a lateral perturbation. The full questionnaire data are plotted in Figure S1 for the reader's reference.

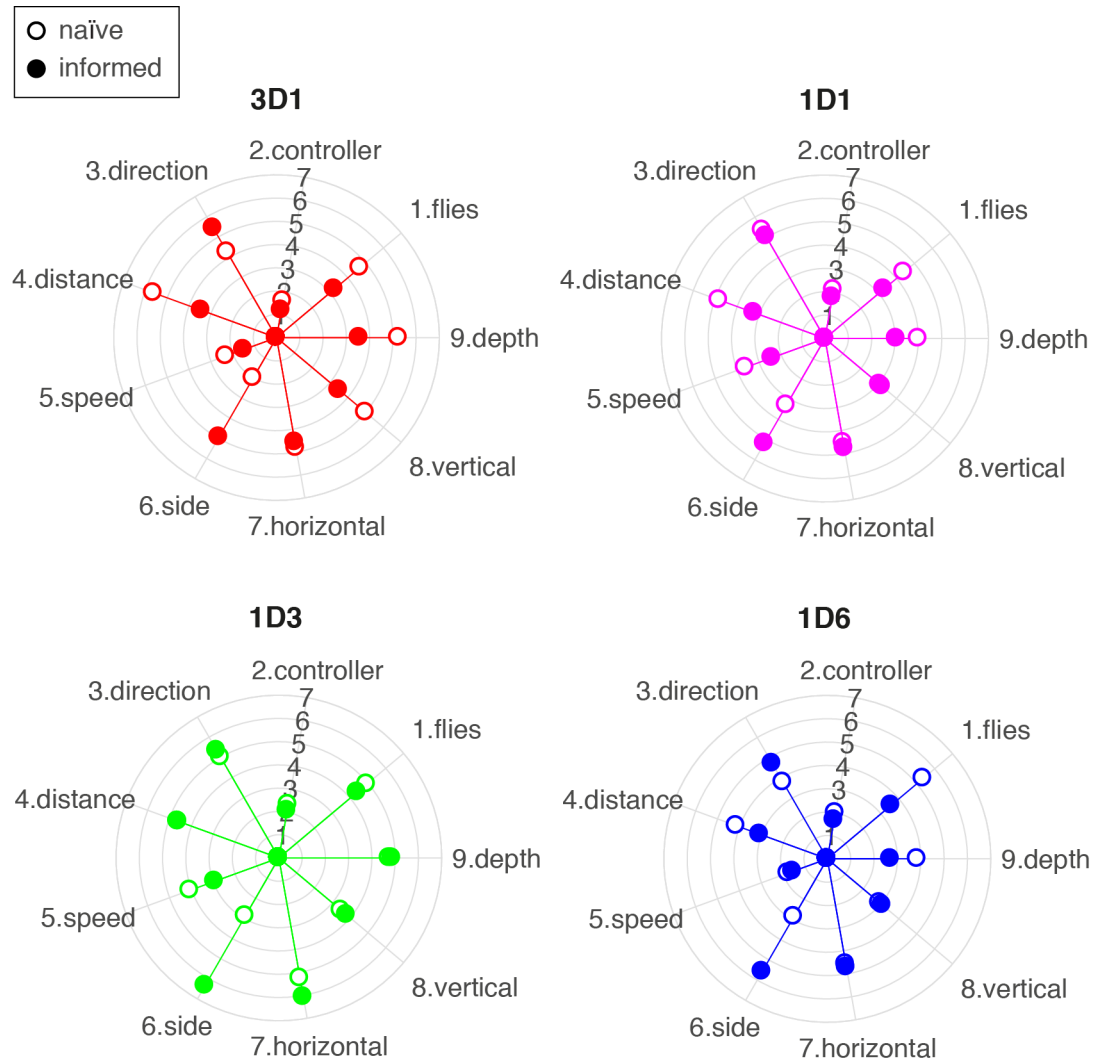


Figure S1. Exploration questionnaire. Plotted are mean responses for the different groups on the different 7-point Likert scale items. Responses following the naïve stage are plotted with open symbols whereas the responses following the informed stage are plotted with filled symbols. The full questionnaire items are provided below.

Table S1. *Exploration questionnaire*

<i>Indicate for the following statements the extent to which you agree.</i>							
	Not at all		A little			Very much	
	1	2	3	4	5	6	7
1. Some flies were more difficult to catch than others							
I have tried to catch more flies by varying in.....							
	Not at all		A little			Very much	
2. The way I held the controller	1	2	3	4	5	6	7
3. The direction in which I moved my hand	1	2	3	4	5	6	7
4. The distance I moved my hand	1	2	3	4	5	6	7
5. The speed of my hands' movement	1	2	3	4	5	6	7
6. The side at which I hit the fly	1	2	3	4	5	6	7
7. The horizontal position of my hand's endpoint	1	2	3	4	5	6	7
8. The vertical position of my hand's endpoint	1	2	3	4	5	6	7
9. The depth position of my hand's endpoint	1	2	3	4	5	6	7
Please "draw below where you aimed relative to the fly to score points:"							
Top view				Frontal view			
0				0			

Participants who indicated in their drawing in the last question that they aimed to the left of the fly were scored as aware, others were scored as unaware. Figure S2.A shows that following the naïve stage, in the 1D1 group a greater fraction of participants was classified as aware than in the other groups. Following the informed adaptation phase, the fraction of participants that was classified as aware increased in all groups. Figure 2.B shows that participants who were classified as unaware adapted less than participants who were classified as aware. However, the participants that were classified as unaware following the naïve adaptation stage did adapt.

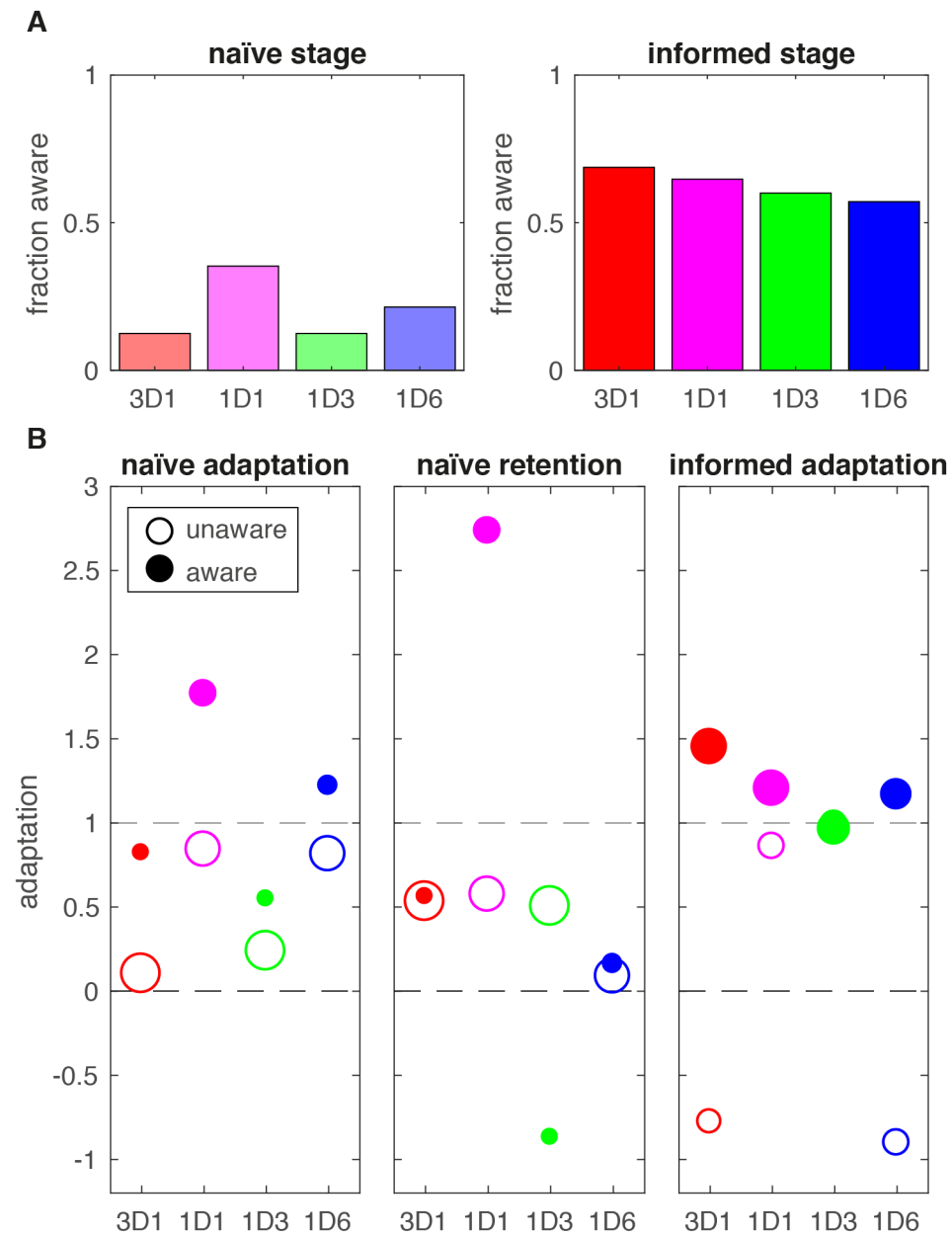


Figure S2. Awareness of the perturbation. A. Fraction of participants that was scored as 'aware' following the naïve and informed adaptation stage. B. Adaptation index for the unaware (open symbols) and aware (filled symbols) participants in the naïve adaptation, naïve retention and informed adaptation phase. The diameter of the symbols reflects the number of the participants.