Cognition 182 (2019) 213–219

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

The kinematics that you do not expect: Integrating prior information and kinematics to understand intentions

Atesh Koul[a], Marco Soriano[b,a], Barbara Tversky[c], Cristina Becchio[a,b], Andrea Cavallo[b,a,*]

[a] C'MoN, Cognition, Motion and Neuroscience Unit, Fondazione Istituto Italiano di Tecnologia, Genova, Italy
[b] Department of Psychology, University of Torino, Torino, Italy
[c] Department of Psychology, Stanford University, Stanford, CA, USA

ARTICLE INFO

Keywords:
Prior expectation
Kinematics
Intention
Action observation
Drift diffusion model

ABSTRACT

Expectations facilitate perception of expected stimuli but may hinder perception of unexpected alternatives. Here, we consider how prior expectations about others' intentions are integrated with visual kinematics over time in detecting the intention of an observed motor act (grasp-to-pour vs. grasp-to-drink). Using rigorous psychophysics methods, we find that the processes of ascribing intentions to others are well described by drift diffusion models in which evidence from observed movements is accumulated over time until a decision threshold is reached. Testing of competing models revealed that when kinematics contained no discriminative intention information, prior expectations predicted the intention choice of the observer. When kinematics contained intention information, kinematics predicted the intention choice. These findings provide evidence for a diffusion process in which the influence of expectations is modulated by movement informativeness and informative kinematics can override initial expectations.

1. Introduction

If you see someone across the street waving an arm, they could be hailing a taxi or swatting a fly. How do you decide which one it is?

One proposal is that the brain uses contextual cues to form expectations about the intention of the observed action (Clark, 2016; Iacoboni, Molnar-Szakacs, Gallese, Buccino, & Mazziotta, 2005; Kilner, 2011; Summerfield & Egner, 2009). For example, seeing a yellow car may lead you to expect that the person is trying to flag a taxi. According to this view, the process of ascribing intentions to others begins even before the start of the movement with a prior prediction of the goal or intention of the to-be observed action (Kilner, 2011).

As soon as the other person starts moving, however, visually observable kinematics become another source of information to disambiguate intention. Observers watching an agent reaching to grasp an object can judge properties of the object from the early kinematics of the movement (Ansuini et al., 2016; Podda, Ansuini, Vastano, Cavallo, & Becchio, 2017). Moreover, grasping kinematics transmit information about what the agent will do next with the object, forming the basis for ascribing intention (Ansuini, Cavallo, Bertone, & Becchio, 2015). Combining quantitative behavioral methods with modelling, Cavallo, Koul, Ansuini, Capozzi, and Becchio (2016) recently demonstrated that intention discrimination covaries with movement kinematics on a trial-by-trial basis. These latter results support a readout model in which, in absence of contextual information, intention choice is directly related to the unfolding parameters of the viewed movement (Becchio, Koul, Ansuini, Bertone, & Cavallo, 2018). What remains to be investigated is how contextual and kinematic information are integrated over time. Contextual cues bias expectations about what is likely to occur, thereby facilitating efficient processing of expected stimuli (Wurm & Schubotz, 2017). However, they can also be misleading. For example, despite the approach of a taxi, the person may be still attempting to swat a fly. How does the brain combine contextual information with the stream of kinematic information to predict a response? Is kinematic evidence powerful enough to overcome the initial, context-driven expectation?

A useful framework for addressing these questions is provided by the *drift diffusion model* (*DDM*). The DDM is a sequential sampling model that regards the decision process as the accumulation of sensory information over time until a decision boundary threshold for choice is reached (Bogacz, 2007; Gold & Shadlen, 2007; Ratcliff & McKoon, 2008; Ratcliff, Smith, Brown, & McKoon, 2016; Wagenmakers, 2009). In two experiments, we used the DDM to separate the influence of kinematics from the influence of context on the decision process. Participants observed a hand reaching for a bottle, either to pour or to drink, and were asked to decide on the intention of the observed movement. In both experiments, movements were occluded at the time of contact of

* Corresponding author at: Department of Psychology, University of Torino, Via Po, 14, 10123 Torino, Italy.
E-mail address: andrea.cavallo@unito.it (A. Cavallo).

https://doi.org/10.1016/j.cognition.2018.10.006
Received 14 May 2018; Received in revised form 9 October 2018; Accepted 9 October 2018
Available online 19 October 2018

the fingers with the bottle. In the first, we varied the informativeness of kinematic parameters of the observed reach-to-grasp movement to show that in the absence of contextual information, the rate of approach to the decision boundary is determined by the amount of intention-related information conveyed by movement kinematics (Experiment 1). In the second experiment, we further varied the likelihood of the upcoming movement (to drink or to pour) to show how the contextual cues bias intention perception both prior to and during evidence accumulation (Experiment 2).

2. Experiment 1

The DDM assumes that dichotomous decisions are based on the accumulation of evidence. The accumulation begins at a starting point (denoted z), that reflects baseline evidence for each alternative. When no prior information is available, the starting point z lies equidistant from the two boundaries (a), and the rate of evidence accumulation (drift rate, denoted v) is determined solely by the strength of the observed signal. In the absence of contextual cues, information conveyed by movement kinematics should set the drift rate. One way to investigate the influence of kinematic information on readout is to provide selective vision to different periods of the observed movement (e.g. Chambon, et al., 2011). However, because the time available for visual information processing is also varied across progressive occlusion points, this paradigm makes it impossible to disentangle effects due to actual information readout and the effects due to variations in the length of the viewing period (Farrow, Abernethy, & Jackson, 2005). To avoid this confusion, in Experiment 1, we took advantage of the readout model developed by Cavallo et al. (2016) to manipulate the amount of intention-related information transmitted by movement kinematics while keeping the viewing period constant.

2.1. Materials and methods

2.1.1. Participants

Twenty right-handed participants (10 females, mean age 22 years, range 19–27) took part in the experiment. The sample size was determined in advance by power analysis using effect sizes observed in a pilot study (n = 6). The sample size was calculated to detect a Cohen's d of 0.7 with alpha set at 0.05 (one-sided), and power set at 0.90 using G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007).

All participants had normal or corrected to normal vision and no history of psychological or neurological disorders. Data from one participant were discarded due to technical issues in data collection. Thus, we report data from 19 participants (10 females, mean age = 22.16 years, range = 19–27). The study was approved by local ethical committee (Comitato Etico Regione Liguria) and was carried out in accordance with the principles of the revised Helsinki Declaration (World Medical Association General Assembly, 2008). All the participants provided written informed consent.

2.1.2. Stimuli generation

Video stimuli depicted a hand reaching towards and grasping a bottle either with the intent to pour (grasp-to-pour) or to drink (grasp-to-drink). Details of how these movements were acquired are available from Cavallo et al. (2016). Below we provide a brief summary of the methods relevant to this study.

Kinematics and video recording. A near-infrared motion capture system with 9 cameras (frame rate, 100 Hz; Vicon System) was used to track the hand kinematics of a distinct set of 17 naïve participants reaching towards and grasping a bottle with the intent either to pour some water into a small glass (grasp-to-pour; N = 256) or to drink water from the bottle (grasp-to-drink; N = 256). After data collection, each trial was individually inspected for correct marker identification, and then run through a low-pass Butterworth filter with a 6 Hz cut-off. Kinematics parameters of interest (N = 16; see Supplementary Table 1) were computed throughout the reach-to-grasp phase of the movement (based on reach onset and grasp offset) at intervals of 10% of the normalized movement time. Movements were also filmed from a lateral viewpoint using a digital video camera (Sony Handy Cam 3-D, 25 frames/sec) placed at about 120 cm from participant's hand starting position with the camera view angle directed perpendicularly to the agent's midline.

2.1.3. Stimuli selection

To manipulate the amount of intention-specifying information transmitted by movements while keeping the viewing period constant, we applied the Classification and Regression Tree (CaRT) model developed by Cavallo et al. (2016). Based on kinematic parameters of a given movement, this readout model provides an estimate of the probability of the predicted choice. In the present study, we combined this information with the actual movement intention to select, for each intention, two subsets of movements: (i) high-informative movements (N = 50, 25 grasp-to-pour and 25 grasp-to-drink), i.e., movements for which the predicted probability of correct choice was high (mean predicted accuracy: 0.70); (ii) low-informative movements (N = 50, 25 grasp-to-pour and 25 grasp-to-drink), i.e., movements for which the predicted probability of correct choice was at chance level (mean predicted accuracy: 0.50). The corresponding videos were used as stimuli for the experiment (see Supplementary Videos 1–4). All video clips were occluded at the time the fingers contacted the object using Adobe Premiere Pro CS6 (.mp4 format, disabled audio, 25 frames/s, resolution 1280 × 800 pixel). Thus, each video clip started with the actual reach onset, and ended at grasp offset, with the duration of the videos varying according to the actual duration of the movements (range = 760–1440 ms). The duration of the video clips did not differ between intentions (grasp-to-pour: mean ± SEM = 1.05 ± 0.019 s; grasp-to-drink: mean ± SEM = 1.02 ± 0.019 s; $t_{(98)} = 1.25$, $p = 0.21$).

2.1.4. Task and procedure

Participants sat in front of a 24-inch computer monitor (1280 × 800 resolution, 75 Hz) at a viewing distance of 50 cm. Task structure conformed to a one-interval forced choice task with symmetric binary choice (to drink vs. to pour). Each trial began with a screen (1500 ms) informing the participant about the button press for a specific intention. For half of the participants, the Italian word 'beve' (to drink) on the left prompted a button press with index finger on the left button of a wireless keyboard touchpad, while the word 'versa' (to pour) on the right prompted a button press with middle finger on the touchpad right button. The position of the two words was counterbalanced across participants. This screen was followed by a green fixation cross (+) at the center of the monitor for 1500 ms. Then, a video-clip showing the reach-to-grasp phase of the action was presented. To ensure that movement sequences could be temporally attended, i.e., that participants had enough time to focus on the hand before movement start, 9 (corresponding to 360 ms), 11 (440 ms), or 13 (520 ms) static frames were randomly added at the beginning of each video clip. These static frames depicted the initial hand posture as displayed in the first frame of the to-be-observed video-clip. Participants were instructed to respond drink or pour as accurately and quickly as possible either during the video, or within a maximum of 3000 ms after the video ended. No feedback was provided to participants at any stage of the experiment.

There were four blocks of 100 trials each. Each block randomly mixed high-informative (N = 50) and low-informative movements (N = 50). The entire experiment lasted approximately 60 min. Stimuli presentation, timing, and randomization procedures were controlled using E-prime (version 2.0.10.242).

2.1.5. Drift diffusion modeling

The proportion of correct responses and reaction times (RTs) for correct and incorrect trials are shown in Fig. 1.

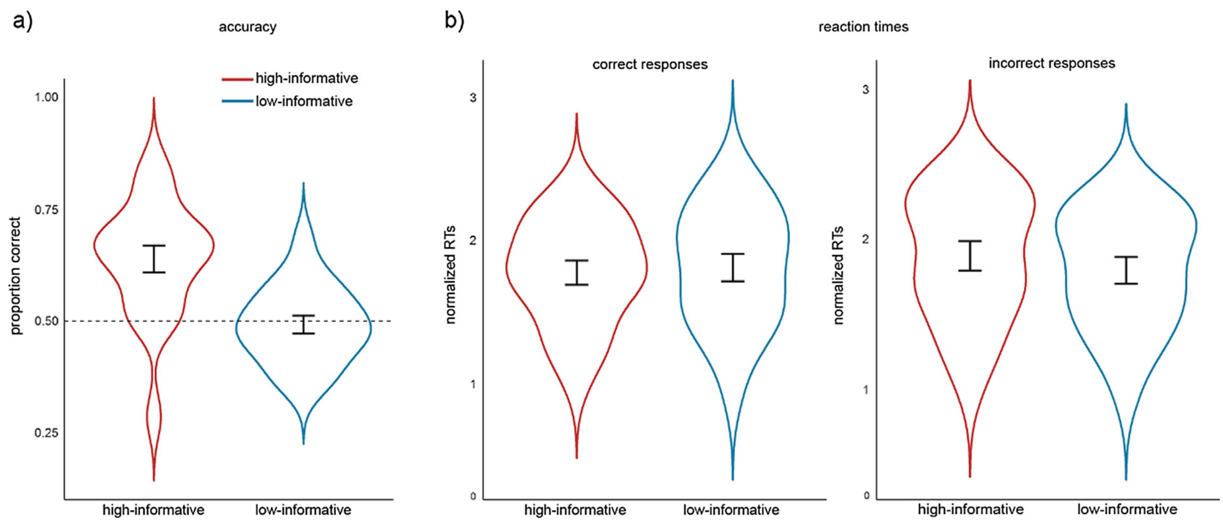

Fig. 1. Proportion of correct responses and reaction times in Experiment 1. Violin plots depicting distributions of proportion correct (a) and reaction times (b) for high- and low-informative movements. Error bars represent standard error of mean.

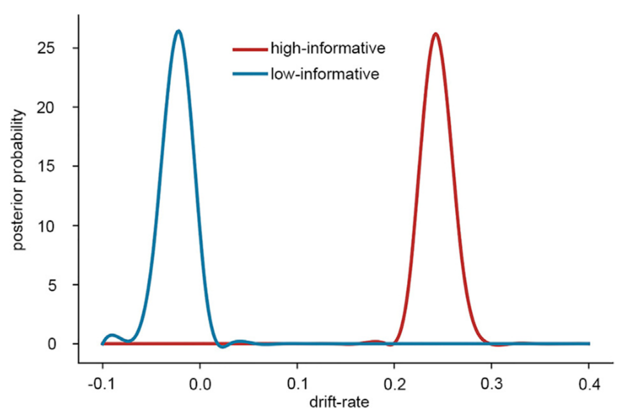

Fig. 2. Drift rate posterior probability distributions for high- and low-informative movements in Experiment 1. Drift rates were significantly larger for high-informative movements compared to low-informative movements ($P_{p|D}$ [high-informative > low-informative] = 1.0).

We fit each participant's choices and RT distributions (for both correct and incorrect trials) using a DDM. Estimation of the DDM parameters was performed using the python-based hierarchical DDM (HDDM) toolbox (Wiecki, Sofer, & Frank, 2013). This toolbox estimates the parameter estimates in a Bayesian framework and additionally allows for quantification of estimation uncertainty in the form of the posterior distribution. The posterior distributions are approximated accurately using the Markov chain Monte-Carlo (MCMC) sampling methods. The DDM parameters are modelled as distributed according to a normal (or truncated normal, depending on the bounds of parameter intervals) distribution centred around the group mean with group variance. A total of 5000 samples were drawn for each model where the first 200 were discarded (burn-in) as the initial samples are likely to be unreliable due to selection of random starting point.

We examined two variants of the DDM with different parameter constraints for drift diffusion: Model 1 (M1) assumed the same value for the drift rate for high-informative and low-informative movements, Model 2 (M2) allowed drift rate to vary between high- and low-informative movements. The deviance information criterion (DIC) was used for model comparison, where lower DIC values favor models with the highest likelihood and least number of parameters (Spiegelhalter, Best, Carlin, & van der Linde, 2002). The actual absolute value of the DIC is strongly dependent on the sample size of the experiment. A difference of 10 between model DIC scores is generally interpreted as evidence in favor of the better (i.e., lower) scoring model (Dunovan, Tremel, & Wheeler, 2014).

Bayesian hypothesis testing was performed by analyzing the posterior distribution of drift rate parameters. To determine significant differences, we calculated the proportion of posteriors in which the drift rate for high-informative movements was higher than that of low-informative movements. A difference of less than 5% in the posterior distribution overlap ($P_{p|D}$) was considered significant. Since the hierarchical estimation procedure utilized violates the independence assumption, we did not analyze subject parameter estimates in a frequentistic test.

2.2. Results

We compared two variants of the HDDM with different constraints for the drift diffusion parameter. The model fit significantly better when drift was allowed to vary as a function of movement informativeness (DIC = 28387.08 for M1, DIC = 28200.88 for M2). This indicates that kinematics are incorporated in the decision process by a change in drift rate.

Fig. 2 shows the posterior parameter estimates for high-informative and low-informative movements. Drift rates were significantly larger for high-informative movements than for low-informative movements ($P_{p|D}$ [high-informative > low-informative] = 1.0). As predicted, whereas drift accumulation showed a positive gradient for high-informative movements (P_p|D [high-informative > 0] = 1.0), for low-informative movements, drift rate was not significantly higher than 0 ($P_{p|D}$ [low-informative > 0] = 0.13). Taken together, these results indicate that in the absence of contextual information, the rate of evidence accumulation reflects the readout of the information conveyed by movement kinematics.

3. Experiment 2

Having quantified the effect of movement informativeness on the decision process, in Experiment 2 we proceeded to investigate how movement informativeness is combined with context-driven expectations to make a response. Formally, two mechanisms have been suggested for incorporating expectations about stimulus occurrence in DDM: adjusting the starting point (z) of the decision process (the 'origin model' or prior bias) or selectively increasing the rate of evidence accumulation (v) for the expected stimuli (the 'gain model' or dynamic bias; Moran, 2015; for review, see Summerfield & de Lange, 2014). By fitting alternative versions of the DDM, in Experiment 2 we tested the influence of these mechanisms both separately and in combination.

3.1. Materials and methods

3.1.1. Participants

Twenty healthy participants (10 females, mean age 22.25 years, range 19–25) who had not taken part in Experiment 1 were recruited for this study. All participants had normal or corrected to normal vision and no history of psychological or neurological disorders. The study was approved by local ethical committee (Comitato Etico Regione Liguria) and was carried out in accordance with the principles of the revised Helsinki Declaration (World Medical Association General Assembly, 2008). All the participants provided written informed consent.

3.1.2. Stimuli, task and procedures

The stimuli, task, and procedures were identical to those used in Experiment 1, except that each trial was preceded by a cue informing the participant of the probability that the upcoming stimulus would be a ‘to drink’ or ‘to pour’ movement. The cue consisted of the Italian word ‘beve 80%’ (to-drink 80%) or ‘versa 80%’ (to-pour 80%) displayed above the picture of the to-be-grasped bottle. An empty glass was displayed close to the bottle when the cue was ‘to-pour 80%’, but not when the cue was ‘to-drink 80%’. Participants were informed of the cue’s meaning prior to starting the experiment. For example, a ‘to-pour 80%’ cue indicated an 80% probability of seeing a grasp-to-pour movement and a 20% probability of seeing a grasp-to-drink movement on the upcoming trial. After 2000 ms, the cue screen disappeared and the trial began. This manipulation resulted in four conditions: high-informative movements preceded by a valid cue (valid/high-informative), high-informative movements preceded by an invalid cue (invalid/high-informative), low-informative movements preceded by a valid cue (valid/low-informative), low-informative movements preceded by an invalid cue (invalid/low-informative). As in experiment 1, participants completed 4 blocks. Each block had 100 trials with 80 valid trials (40 trials × 2 movement informativeness [high/low] and 20 invalid trials (10 trials × 2 movement informativeness [high/low]). The entire experiment lasted approximately 70 min.

3.1.3. Drift diffusion modeling

The proportion of correct responses and reaction times (RTs) for correct and incorrect trials per condition are shown in Fig. 3.

We compared four hypothetical variants of the DDM for the intention choice behavior with different constraints for the starting point z and the drift rate v parameters. Model 1 (M1) included a single estimate for both the starting point and the drift rate. Model 2 (M2) included a single estimate for the starting point but allowed the drift rate to vary between high- and low-informative movements. Model 3 (M3) assumed a variable starting point as a function of cue and a variable drift rate as function of movement informativeness. The fourth model, M4, assumed a variable starting point as a function of cue. In addition, it assumed the drift rate to vary as a function of both cue validity and movement informativeness.

Table 1
Experiment 2. Deviance information criterion (DIC) for competing diffusion models of increasing complexity.

Model	Model parameter dependencies	DIC values
M1	None	30648.12
M2	‘v’ depends on movement informativeness	30445.57
M3	‘v’ depends on movement informativeness ‘z’ depends on cue	30441.98
M4	‘v’ depends on movement informativeness and cue validity ‘z’ depends on cue	30182.37

3.2. Results

Table 1 shows DIC values as a function of model complexity. Based on estimated DIC values for each model, the model fit better when the starting point varied as a function of cue and the drift rate was allowed to vary as a function of both cue validity and movement informativeness (M4). Because DIC score for M4 was lower than that of all the other models by at least 10 (Burnham & Anderson, 2002), this indicates that M4 provided a sufficiently better fit to the data than competing models to justify its greater complexity.

Examining posterior probability estimates of starting point, confirmed a significant shift in the origin of the accumulation process as a function of cue ($P_{p|D}$ [cue ‘to drink’ < 0.5] < 0.001 and $P_{p|D}$ [cue ‘to pour’ > 0.5] < 0.001). This indicates that when one intention (e.g., ‘to drink) was more probable than the other intention (e.g., ‘to pour’) as signaled by the contextual cue, the starting point of the decision process (z) shifted towards the ‘to-drink’ boundary and vice versa. Additionally, drift rates of both high-informative and low-informative movements were faster following a valid cue than an invalid cue ($P_{p|D}$ [valid/high-informative > invalid/high-informative] = 1.0; $P_{p|D}$ [valid/low-informative > invalid/low-informative] = 1.0), with valid/high-informative drift rates being faster than valid/low-informative drift rates ($P_{p|D}$ [valid/high-informative > valid/low-informative] = 1.0). This indicates that uptake of intention-specifying information was faster for expected (validly cued) compared to unexpected (invalidly cued) movements.

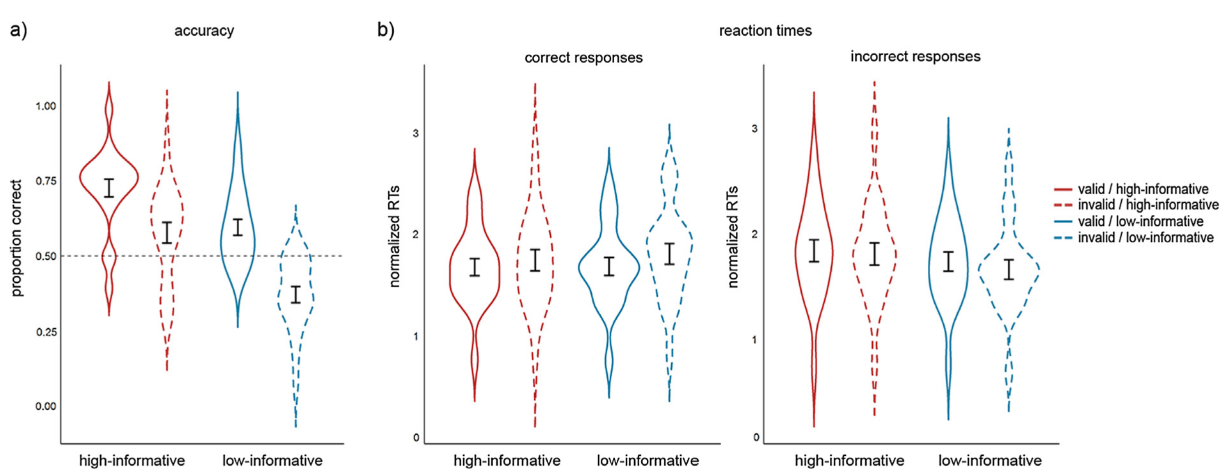

Fig. 3. Proportion of correct responses and reaction times in Experiment 2. Violin plots depicting distributions of proportion correct (a) and reaction times (b) for high- and low-informative movements following a valid (solid line) or invalid cue (dashed line). Error bars represent standard error of mean.

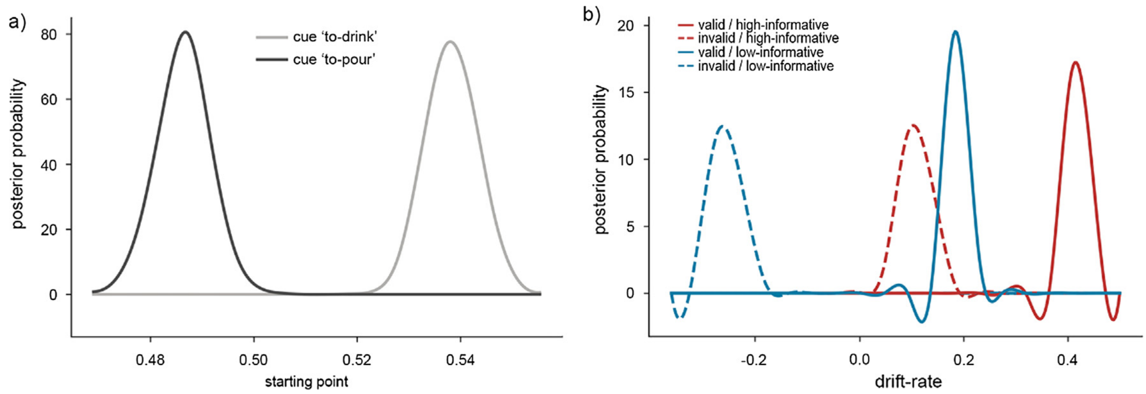

Fig. 4. Starting point (a) and drift rate (b) posterior probability distributions in Experiment 2. When one intention (e.g., 'to drink) was more probable than the other intention (e.g., 'to pour') as signaled by a contextual cue, the starting point of the decision process shifted towards the cued boundary (Panel a). The drift rates of both high-informative (red curves) and low-informative (blue curves) movements were faster following a valid cue (solid line) than an invalid cue (dashed lines), with valid/high-informative drift rates being faster than valid/low-informative drift rates. All drift rates had positive values except invalid/low-informative. This indicates that on all but invalid/low-informative movements the process reached the correct response boundary (Panel b). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

As illustrated in Fig. 4, for low-informative movements, valid and invalid drift rates had opposite signs. This indicates that on valid cue trials observers were driven towards the correct response boundary. For invalid cue trials, however, they were driven towards the incorrect response boundary. For high-informative movements, drift rates had positive values under both valid and invalid cue conditions ($P_{p|D}$ [valid/low-informative > 0] $= 1.0$ and $P_{p|D}$ [invalid/high-informative > 0] $= 1.0$). This indicates that regardless of cue validity the process reached the correct response boundary, though more slowly for invalid cues than for valid cues.

4. General discussion

If you see someone across the street waving an arm, are they hailing a taxi or swatting a fly? How do we decide?

Our results show that, in the absence of contextual influences, the readout of kinematic information drives the decision. Observed movements vary in the amount of intention-specifying information they transmit (Becchio et al., 2018; Cavallo et al., 2016). Our DDM analysis in Experiment 1 shows that, in the absence of contextual cues, the rate of evidence accumulation reflects the uptake of such information, and is significantly higher for high-informative movements than low-informative movements. This complements earlier work (Cavallo et al., 2016) by showing that variations in movement informativeness map directly onto the amount of information per time unit that is absorbed.

A second finding is that expectations bias intention attribution both prior to and during evidence accumulation. Accumulation-to-bound models provide two possible explanations for the effects of expectations: (i) expectations lead to shifts in the starting point of evidence accumulation (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Diederich & Busemeyer, 2006; Edwards, 1965; Laming, 1968; Link & Heath, 1975; Ratcliff, 1985; Voss, Rothermund, & Voss, 2004; Wagenmakers, Ratcliff, Gomez, & Mckoon, 2007); or (ii) expectations change the rate with which evidence is accumulated (Ashby, 1983; Diederich & Busemeyer, 2006; Ratcliff, 1985). Our findings (Experiment 2) suggest that both these mechanisms influence the ascription of intention. Prior to evidence accumulation, expectations bias participants to make one response over the other by displacing the starting point towards the cued intention (e.g., to pour). During evidence accumulation, a dynamic mechanism ensures that this initial bias translates into a gain in processing expected kinematics.

Consistent with the Bayesian perspective that prior information should hold most sway over decisions when sensory evidence is imprecise or ambiguous (Summerfield & De Lange, 2014), we found that the influence of expectations on evidence accumulation was greater for low-informative movements than for high-informative movements. Notably, participants exposed to invalidly cued high-informative movements were nonetheless able to identify the unexpected intention. A third implication of our results is that informative kinematics can counter invalid contextual expectations.

These findings have consequences for theories casting action understanding as hierarchical predictive coding (Clark, 2013, 2016; Friston, Mattout, & Kilner, 2011; Kilner, Friston, & Frith, 2007; Zacks, Speer, Swallow, Braver, & Reynolds, 2007). Under predictive coding, evidence accumulation is best conceived as an evolving probability distribution over multiple possible causes (that is, a generative model) of sensation. The generative model starts with a prior prediction of the intention of an observed action. Given this prior, the model generates a prediction of the kinematics of the action. By comparing the predicted sensory evidence with the actual sensory evidence, the system can assess the likelihood of the intention. If the prediction is correct, we are able to infer the intention of the observed action (Kilner et al., 2007).

But what if the prediction is incorrect? Predictive coding theories typically make the simplifying assumption that "kinematics are not transparently present in motor sequences alone, since there is no unique mapping between such sequences and the intentions behind them" (Clark, 2016). Under this assumption, the intention that is inferred will always "depend upon the prior information received from a contextual level" (Kilner et al., 2007). It follows that if the prediction is incorrect, the incorrectly predicted intention should be inferred.

Our results show that this analysis corresponds to the limited case of unexpected (i.e., invalidly cued) low-informative movements. For high-informative movements, the error term generated by the comparison between predicted and observed kinematics is large enough to override the initial incorrect prediction. 'What is present' thus prevails over 'what is probable'. This challenges the common view that kinematic information "under-constraints the space of candidate cause" and therefore requires prior expectations for intention understanding (e.g. Chambon et al., 2011).

Taken together, our findings give evidence for a model in which the influence of expectations is modulated by movement informativeness (Fig. 5). It would be interesting to explore extensions of this model from decision tasks with only two alternatives to real-world decisions with multiple alternatives. Another important future direction will be to examine situations in which kinematic information and contextual information change not only from trial to trial (as in the present task) but also within a trial. For instance, during the unfolding of a given action, the spatial arrangement of the observed scenario might change, or

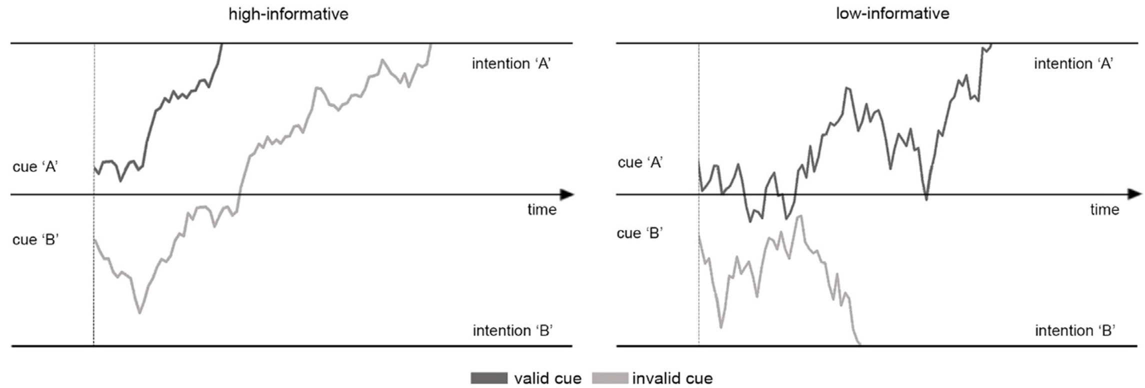

Fig. 5. Schematic representation of drift diffusion model of the decision process as a function of movement informativeness and expectations. For high-informative movements, the process reaches the correct response boundary regardless of the expected intention. For low-informative movements, the processes reaches the correct response boundaries only when the correct intention is expected.

additional contextual information might be provided. Understanding how human observers integrate information present at different time points may lead to a better comprehension of the dynamics of evidence accumulation over time. Finally, it would be important to consider situations in which "behavior affords behavior" and the perceived action is directly tied to one's own action or not (Tversky, 2018). Visual motion direction decisions have been shown to be biased by the cost of the action that is used to report the decision (Hagura, Haggard, & Diedrichsen, 2017). We would predict that the cost of and the possibility for our own actions also partly define how we ascribe intentions to others (Sommerville & Woodward, 2005; Sommerville, Woodward, & Needham, 2005).

Acknowledgments

The authors would like to thank Laura Manno for her help with the data acquisition. AC was funded by Fondazione Compagnia di San Paolo, Excellent Young PI Grant (CSTO167915). CB was supported by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement no. 312919.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2018.10.006.

References

Ansuini, C., Cavallo, A., Bertone, C., & Becchio, C. (2015). Intentions in the brain. *The Neuroscientist, 21*(2), 126–135. https://doi.org/10.1177/1073858414533827.

Ansuini, C., Cavallo, A., Koul, A., D'Ausilio, A., Taverna, L., & Becchio, C. (2016). Grasping others' movements: Rapid discrimination of object size from observed hand movements. *Journal of Experimental Psychology: Human Perception and Performance, 42*(7), 918–929. https://doi.org/10.1037/xhp0000169.

Ashby, F. G. (1983). A biased random walk model for two choice reaction times. *Journal of Mathematical Psychology, 27*(3), 277–297. https://doi.org/10.1016/0022-2496(83)90011-1.

Becchio, C., Koul, A., Ansuini, C., Bertone, C., & Cavallo, A. (2018). Seeing mental states: An experimental strategy for measuring the observability of other minds. *Physics of Life Reviews, 24*(12), 67–80. https://doi.org/10.1016/j.plrev.2017.10.002.

Bogacz, R. (2007). Optimal decision-making theories: Linking neurobiology with behaviour. *Trends in Cognitive Sciences, 11*(3), 118–125. https://doi.org/10.1016/j.tics.2006.12.006.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review, 113*(4), 700–765. https://doi.org/10.1037/0033-295X.113.4.700.

Burnham, K. P., & Anderson, D. R. (2002). Model selection and multimodel inference: A practical information-theoretic approach. *Ecological Modelling, 172*. https://doi.org/10.1016/j.ecolmodel.2003.11.004.

Cavallo, A., Koul, A., Ansuini, C., Capozzi, F., & Becchio, C. (2016). Decoding intentions from movement kinematics. *Scientific Reports, 6*(November), 37036. https://doi.org/10.1038/srep37036.

Chambon, V., Domenech, P., Pacherie, E., Koechlin, E., Baraduc, P., & Farrer, C. (2011). What are they up to? The role of sensory evidence and prior knowledge in action understanding. *PLoS ONE, 6*(2), https://doi.org/10.1371/journal.pone.0017133.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences, 36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477.

Clark, A. (2016). *Surfing Uncertainty*. Oxford University Presshttps://doi.org/10.1093/acprof:oso/9780190217013.001.0001.

Diederich, A., & Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: Bound-change, drift-rate-change, or two-stage-processing hypothesis. *Perception & Psychophysics, 68*(2), 194–207. https://doi.org/10.3758/BF03193669.

Dunovan, K. E., Tremel, J. J., & Wheeler, M. E. (2014). Prior probability and feature predictability interactively bias perceptual decisions. *Neuropsychologia, 61*(1), 210–221. https://doi.org/10.1016/j.neuropsychologia.2014.06.024.

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology, 2*(2), 312–329. https://doi.org/10.1016/0022-2496(65)90007-6.

Farrow, D., Abernethy, B., & Jackson, R. C. (2005). Probing expert anticipation with the temporal occlusion paradigm: Experimental investigations of some methodological issues. *Motor Control, 9*(3), 332–351. https://doi.org/10.1123/mcj.9.3.330.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146.

Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics, 104*(1–2), 137–160. https://doi.org/10.1007/s00422-011-0424-z.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30*(1), 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038.

Hagura, N., Haggard, P., & Diedrichsen, J. (2017). Perceptual decisions are biased by the cost to act. *ELife, 6*, e18422. https://doi.org/10.7554/eLife. 18422.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., & Mazziotta, J. C. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology, 3*(3), 0529–0535. https://doi.org/10.1371/journal.pbio.0030079.

Kilner, J. M. (2011). More than one pathway to action understanding. *Trends in Cognitive Sciences, 15*(8), 352–357. https://doi.org/10.1016/j.tics.2011.06.005.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processing, 8*(3), 159–166. https://doi.org/10.1007/s10339-007-0170-2.

Laming, D. R. J. (1968). *Information theory of choice-reaction times.* Oxford, England: Academic Press.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika, 40*(1), 77–105. https://doi.org/10.1007/BF02291481.

Moran, R. (2015). Optimal decision making in heterogeneous and biased environments. *Psychonomic Bulletin & Review, 22*(1), 38–53. https://doi.org/10.3758/s13423-014-0669-3.

Podda, J., Ansuini, C., Vastano, R., Cavallo, A., & Becchio, C. (2017). The heaviness of invisible objects: Predictive weight judgments from observed real and pantomimed grasps. *Cognition, 168*, 140–145. https://doi.org/10.1016/j.cognition.2017.06.023.

Ratcliff, R. (1985). Theoretical interpretations of the speed and accuracy of positive and negative responses. *Psychological Review, 92*(2), 212–225. https://doi.org/10.1037/0033-295X.92.2.212.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences, 20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007.

Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: The relation between action processing and action production in infancy.

Cognition, 95(1), 1–30. https://doi.org/10.1016/j.cognition.2003.12.004.
Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition, 96*(1), 1–11. https://doi.org/10.1016/j.cognition.2004.07.004.
Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B (Statistical Methodology), 64*(4), 583–639. https://doi.org/10.1111/1467-9868.00353.
Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience, 15*(11), 745–756. https://doi.org/10.1038/nrn3838.
Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences, 13*(9), 403–409. https://doi.org/10.1016/j.tics.2009.06.003.
Tversky, B. (2018). Sometimes we can see some mental states. *Physics of Life Reviews, 24*, 88–90. https://doi.org/10.1016/j.plrev.2017.11.017.
Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32*(7), 1206–1220. https://doi.org/10.3758/BF03196893.
Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology, 21*(5), 641–671. https://doi.org/10.1080/09541440802205067.
Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & Mckoon, G. (2007). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language, 58*(1), 140–159. https://doi.org/10.1016/j.jml.2007.04.006.
Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the drift-diffusion model in python. *Frontiers in Neuroinformatics, 7*(August), 14. https://doi.org/10.3389/fninf.2013.00014.
Wurm, M. F., & Schubotz, R. I. (2017). What's she doing in the kitchen? Context helps when actions are hard to recognize. *Psychonomic Bulletin & Review, 24*(2), 503–509. https://doi.org/10.3758/s13423-016-1108-4.
Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin, 133*(2), 273–293. https://doi.org/10.1037/0033-2909.133.2.273.