Mirror self recognition as a product of forward models; implications for delusions of body image and visual neglect ☆

Sean O'Connor

568A Jericho Road, Whittier, NC 28789, United States

ABSTRACT

Forward models allow individuals to learn to predict the sensory consequences of their own behavior. Social forward models have been proposed as an extension of forward models, allowing individuals to learn to predict the response of another to the individual's own behavior. This article proposes similarly that an individual who treats their reflection as another may learn to predict the behavior of their reflection, offering a new perspective on mirror self-recognition and a potential framework through which to investigate visual delusions. Specifically this article investigates this framework by considering four body image delusions; mirrored-self misidentification, body integrity disorder (BID), phantom limbs, and gender dysphoria, and two delusions associated with visual neglect; somatoparaphrenia and mirror agnosia.

Introduction

Forward models have been proposed as a mechanism for the central nervous system to differentiate between sensory input from the environment and sensory input from self-generated behavior [1]. Pickering & Garrod [2] use an example of moving a hand to a target to demonstrate how a motor act is proposed to consist of two parallel processes simultaneously triggered by a motor command:

"First, it (the motor command) causes the action implementer to generate the act, which in turn leads the perceptual implementer to construct a percept of the experience of moving the hand….Second, it sends an efference copy of the action command to cause the forward action model to generate the predicted act of moving the hand….The predicted act then causes the forward perceptual model to construct a predicted percept of the experience of moving the hand."

These two processes allow the individual to differentiate between sensory input from self-generated behavior and sensory input from the environment by comparing the predicted sensory experience with the actual sensory experience. When the prediction generated by the forward model = perceptual experience (also referred to as P = E) a phenomenon known as sensory attenuation occurs, dampening attention to the sensory experience and allowing the individual to know that the experience is a result of their own behavior [3]. A common example used to display this sensory attenuation is that an individual cannot tickle himself. According to motor theory this is because the predicted experience of the efference copy for the tickling motor act matches the experience of tickling; thus, P = E and the tickling sensation is canceled [2].

When the predicted sensory experience and actual sensory experience do not match (also referred to as P =/= E), sensory neurons fire, drawing the attention of the individual to this prediction error [3]. This prediction error corrects the forward model such that the next action can be modified. A simple example to demonstrate this is that if a forward model is used to predict where an individual's hand will be during a grasp, and the hand is a little further to the right than predicted, this prediction error is noted and the next time the individual attempted the same grasp, he would grasp a little further to the right [2].

Thus, forward models can be used for predicting the perceptual outcomes of an action and granting agency of the sensory experience by attenuating the sensory effects of self-motion. Forward models have further been extended beyond simple motor control to social control. As Wolpert, Doya, and Kawato [4] write,

"We can consider a similar forward or predictive model for social interaction. In this case another person's response to my motor commands or communicative behaviour is modelled. Again, discrepancies between anticipated and actual behaviour can be used to refine such a model. Therefore, by monitoring one's own action and the response of others it is possible to learn a predictive model of the likely behaviour of someone in response to our actions."

This article extends this idea by arguing that if an individual treats their reflection in the mirror as another and learns to predict the response of this reflection to their own actions, this "reflection prediction" would allow mirror self-recognition and could help to further our understanding of certain visual delusions related to the body image and caused by visual neglect. This is because, when in front of a mirror, an individual would now be predicting the image of their reflection; thus, $P = E$ and sensory attenuation should occur, dampening the attention of the individual to the reflection except when $P = / = E$.

Mirror self recognition

Mirror self-recognition is the ability of an individual to recognize their reflection in the mirror as belonging to themselves. In general, animals respond to their reflection in a mirror in one of three ways: treating the reflection as a conspecific indefinitely, treating the reflection as a conspecific before learning to ignore the reflection, and exhibiting self-directed behavior, which is considered experimental evidence that a species is capable of mirror self-recognition [5]. Although this self-directed behavior has proven difficult to test for and understand, it's existence in a species is widely used as a measurement for self-awareness [6]. The "mirror test", developed by Gordon Gallup Jr., is the classic experiment to determine whether a species is capable of self-directed behavior [5]. This involves monitoring the reaction of animals who have had an area of the body they do not normally see marked; in Gallup's original experiment, this marking was accomplished with an odorless red dye [5]. If this marking elicits self-directed behavior in the mirror (i.e. attempting to remove the dot), the species is said to be capable of mirror self-recognition; for example, Gallup used this test to conclude that chimpanzees were capable of mirror self-recognition and monkeys were not, as chimpanzees exhibited self-directed behavior and monkeys did not [5].

The view of mirror self-recognition as a product of forward models offers a new perspective to understand the three general responses of a species to a mirror;

1) Those who indefinitely treat the reflection as another do not learn to predict their reflection
2) Those who treat the reflection as another before learning to ignore it do predict their reflection; the resulting prediction allows $P = E$, allowing agency of the reflection while dampening the individual's attention to the reflection.
3) Those who display self-directed behavior are displaying prediction error correction, as the mark will cause $P = / = E$, and removing the mark will cause $P = E$.

Thus, it is the view of this article that a species that ignores its reflection is capable of mirror self-recognition and that self-directed behavior is the result of prediction error correction. This article views self-directed behavior as a trait which likely emerged in those species which had a strong enough selection pressure to correct reflection prediction errors - for example, species for which appearance has a social fitness consequence, or species which are threatened by disease spread by bugs or blood which may resemble a "mark". However, self-directed behavior does indicate awareness of the location of self and objects in space relative to a mirror, and a reflection prediction may provide a mechanism for this awareness. This makes sense intuitively, as a successful reflection prediction would require an understanding of a relationship between distance of self and objects and a reflection, for example, for an individual to predict that as they moved closer to the mirror, their own reflection grows larger, or to predict that as they move an object in front of them, the reflection moves the object in front of itself. This understanding of the spatial location of self and objects relative to a mirror as a product of a successful reflection prediction yet separate from self-directed behavior is consistent with research indicating that

both species capable of self-directed behavior and species not capable of self-directed behavior have demonstrated the ability to solve spatial problems with the use of a mirror [7].

This article has proposed that a reflection prediction may allow an individual both agency over their reflection and an understanding of the location of self and objects in space relative to a mirror. A delusion is defined as "A belief or altered reality that is persistently held despite evidence or agreement to the contrary" [8]. If a reflection prediction is required for agency over one's body image, errors in this reflection prediction may help to explain delusions that cause an individual to not feel agency over their reflection, or feel agency over a body image that is different than their reflection. This article will consider four such "body image delusions"; mirrored-self misidentification, body integrity disorder, phantom limbs, and gender dysphoria. Furthermore, visual neglect is a disorder characterized by a loss of the ability to visually pay attention to stimuli in one half of the visual field [9]. This article will consider two delusions that may result from the prediction error caused by visual neglect; somatoparaphrenia, and mirror agnosia.

Body image delusions

Mirrored-self misidentification is a cognitive disorder that causes an individual normally capable of mirror-self-recognition to mistake their own reflection for another individual that looks just like them [10]. According to this model, such a reaction would be expected if there was an error that caused $P = / = E$, causing a lack of agency over the reflection and causing the afflicted to treat their reflection as a conspecific, just as species incapable of mirror self-recognition do.

Body integrity disorder (BID) is in part characterized by the desire for amputation of a limb [11]. According to this model, such a reaction would be expected if an individual predicted their reflection does not have a limb. If this is the case, $P = / = E$, and the desire for amputation of the limb may be considered a form of prediction error correction; if the limb is amputated, $P = E$ in front of a mirror. This idea is consistent with research indicating BID is a result of a "mismatch between the mental body image and the physical body" [11].

Patients with phantom limbs continue to experience the sensations of a limb after the limb has been amputated [12]. If BID fits in this model as the result of seeing the reflection of a limb the individual no longer predicts they possess, phantom limbs may be considered the opposite; the result of predicting the reflection of a limb which they no longer actually possess. The experience of the phantom limb may thus be considered the individual feeling agency over a body image that is different than their reflection. This may help explain why mirror therapy has been successfully used to help treat the experience of phantom limbs [12]; mirror therapy replaces the image of the missing reflection, acting as prediction error correction by allowing $P = E$ in front of a mirror and causing the individual to gain agency over the reflection.

Gender dysphoria is characterized by distress caused by identifying as a member of the gender that is opposite their birth gender [13]. According to this model, this would be expected if an individual predicted that their reflection looked like a member of the gender opposite their birth gender; then, $P = / = E$ and the individual would not feel agency over their reflection. Crossdressing and sexual reconstructive surgery, two common means of coping with this disorder [13], may be considered prediction error correction, as this may allow $P = E$ in front of a mirror.

Visual neglect

Visual neglect is a disorder characterized by a loss of the ability to visually pay attention to objects in one half of the visual field [9]. If this visual neglect prevents an individual from predicting the reflection of self and objects in the neglected half of the visual field, and the individual actually saw a reflection of themselves or objects in that half of

the visual field, $P = / = E$, and a prediction error would result. Somatoparaphrenia and mirror agnosia, two delusions that may accompany visual neglect, may be viewed through this model as delusions caused by this prediction error.

For example, somatoparaphrenia is a delusion in which an individual claims that part, or in the case of individuals with visual neglect, half, of their body belongs to another person [14]. According to this framework, if an individual could not predict the reflection of half of their body, but actually experienced seeing the reflection of half of their body, $P = / = E$ and the individual would not gain agency over half of the half of their reflection that they could not predict. Somatoparaphrenia may be considered a delusion that corrects this prediction error to allow $P = E$ in front of a mirror; This is because an individual who cannot predict half of their reflection will process that half of their reflection as belonging to another, thus, by believing half of their body actually belongs to another, ironically, $P = E$ in front of a mirror and sensory attenuation may be re-achieved. This may help explain why somatoparaphrenia can be temporarily reversed with mirror therapy [14]; the delusional belief itself may substitute for the reflection prediction and allow $P = E$ in front of a mirror.

Mirror agnosia is another delusion that occurs in individuals with visual neglect. This disorder characterized by a misunderstanding of the locations of objects in space relative to a mirror for objects located in an area of the visual field that an individual cannot process due to visual neglect [9]. An individual with this disorder who sees the reflection of an object from the visually neglected area will not reach for the object in the visually neglected area, they will try to reach through the mirror [9]. According to this framework, this would be expected if the visual neglect caused an individual to not predict the reflection of an object from the damaged half of the visual field. If the individual actually saw the reflection of an object from this half of the visual field, $P = / = E$ and a prediction error would result. The delusion that the object is through the mirror, rather than in the neglected visual field, is very rational from this perspective – for example, if a healthy individual saw the reflection of an object in a mirror that they did not see outside of the mirror, this author would expect much the same reaction – the healthy individual may try to reach through the mirror. This would be difficult to test outside of a virtual reality environment, yet, is consistent with the idea proposed earlier that a successful reflection prediction may allow the localization of objects in space relative to a mirror.

While translating this view of body image delusions and delusions from visual neglect as the result of prediction error correction in the prediction of a reflection is, to this author's knowledge, novel, the view of delusions in general as the result of prediction error is not. For example, Phillip Corlett concludes:

> "Single factor, prediction error account gives a parsimonious account of delusions that generates novel predictions about how best to treat delusions and incorporates numerous biological, clinical and phenomenological data regarding delusions." [15]

Conclusion

This article has argued that a forward model which allowed an individual to learn to predict the behavior of their reflection could help to further the scientific understanding of mirror self-recognition and provides a potential framework or investigating delusions related to the body image and those caused by visual neglect. While no experiment was run to support the claims of this article, the author believes that a "reflection prediction" offers enough explanatory power for mirror self-recognition and the delusions discussed above to warrant further study. Future work should include theoretical and experimental extensions and incorporating this model into existing social, neurocognitive, and evolutionary models.

Declaration of Competing Interest

No financial conflict of interest to disclose.

References

[1] Crapse TB, Sommer MA. Corollary discharge across the animal kingdom. Nat Rev Neurosci 2008;9(8):587–600. https://doi.org/10.1038/nrn2457.

[2] Pickering MJ, Garrod S. An integrated theory of language production and comprehension. Behav Brain Sci 2013;36(4):329–47. https://doi.org/10.1017/S0140525X12001495.

[3] Schröger E, Kotz SA, SanMiguel I. Prediction and attention in current research on perception and action. Brain Res 2015;1162.

[4] Wolpert DM, Doya K, Kawato M. A unifying computational framework for motor control and social interaction. Philos Trans R Soc Biol Sci 2003;358(1431):593–602. https://doi.org/10.1098/rstb.2002.1238.

[5] Gallup Jr. GG. Chimpanzees: self recognition. Science 1970;167(3914):86–7. https://doi.org/10.1126/science.167.3914.86.

[6] Prior H, Schwarz A, Güntürkün O. Mirror-induced behavior in the magpie (Pica pica): evidence of self-recognition. Publ Lib Sci Biol 2008;6(8):e202https://doi.org/10.1371/journal.pbio.0006020.

[7] Povinelli DJ. Failure to find self-recognition in Asian elephants (Elephas maximus) in contract to their use of mirror cues to discover hidden food. J Comp Psychol 1989;103(2):122–31. https://doi.org/10.1037/0735-7036.103.2.122.

[8] Delusion. In Merriam-Webster.com Retrieved June 3, 2019, from 2011https://www.merriam-webster.com/dictionary.

[9] Ramachandran VS, Altschuler EL, Hillyer S. Mirror agnosia. Proc R Soc Lond B 1997;264:645–7.

[10] Breen N, Caine D, Coltheart M. Mirrored-self misidentification: two cases of focal onset dementia. Neurocase 2001;7(3):239–54. https://doi.org/10.1093/neucas/7.3.239.

[11] Blom RM, Hennekam RC, Denys D. Body integrity identity disorder. PLoS One 2012;7(4):e34702.

[12] Ramachandran VS, Brang D, McGeoch P. Size reduction using mirror visual feedback (MVF) reduces phantom pain. Neurocase 2009;15(5):357–60.

[13] Adams Stephen. Bridging the divide. Medicus 2017;57(3):18.

[14] Fotopoulou A, Jenkinson PM, Tsakiris M, Haggard P, Rudd A, Kopelman MD. Mirror-view reverses somatoparaphrenia: dissociation between first- and third-person perspectives on body ownership. Neuropsychologia 2011;49(14):3946–55. https://doi.org/10.1016/j.neuropsychologia.2011.10.011.

[15] Corlett P. Delusions and prediction error. In: Bortolotti L, editor. Delusions in context. Cham: Palgrave Macmillan; 2018.