

# S

## Surprise



Wendy Ross<sup>1</sup> and Margaret Webb<sup>2</sup>

<sup>1</sup>Department of Psychology, London Metropolitan University, London, UK

<sup>2</sup>University of Melbourne, Melbourne, VIC, Australia

### Abstract

An unexpected or inexplicable event with significance to the self that is accompanied by a strong emotional marker (which can be positive, negative, or mixed in valence).

### Keywords

Surprise · Insight · Serendipity · Error

## Surprise and the Possibility Space

Possibility refers to that which is not actual. It invites us to consider “what could be,” “what is to come,” “what could have been,” and “what is not / will never be.” In interaction with the world (sometimes only our internal world), possibilities can lead us to “what was unexpected.” These unexpected events, when they are of subjective significance, are marked by the epistemic emotion of surprise.

We can all recognize the feeling of surprise, typically when we encounter something unexpected but which we instinctively believe even if

we are curious about its origins. Take for example, the serendipitous surprise of bumping into a dear friend in a country that you did not expect them to be in. A further surprising event, to take a more thoroughly researched example, might be scientists receiving reports of unexpected side effects of drugs which indicate new unanticipated ways that the drug works on the body (Rocca et al. 2019). These reports are not expected, but they trigger curiosity and exploration. Surprise can be not only a marker of this expected violation, but also a marker of unexpected understanding (insight). In the case of the latter, surprise marks both the recognition of prior ignorance and the discovery of new knowledge.

Beyond these broad phenomena such as serendipity, insight, and creativity, surprise underpins and explains aspects of human cognition across several different levels and, we will argue, is fundamental to understanding how we interact with the unpredictable and unexpected nature of the possible (Glăveanu 2021). We go further to suggest that increased attention to the role of surprise in thinking and behavior can support our understanding of how cognition unfolds in the face of uncertainty, allowing us to flesh out and adjudicate between different models at different cognitive levels. In this respect, it is “surprising” that this complex phenomenon has often been relegated to a minor role (Nevo and Erev 2012). We shall end by considering the importance of surprise to the complex interactional process of serendipity and suggest that the importance of

unplanned discoveries in human discovery requires us to understand surprise as more than an epiphenomenon of otherwise clear processes.

### Levels of Surprise

At a cognitive and neural level, surprise is theorized to underpin many of the most fundamental components of learning and memory (Alexander and Brown 2019). It has been posited as a meta-cognitive signal to update mental models to incorporate new, unexpected information (Munnich and Ranney 2019). At the affective level, surprise is accompanied by reliable phenomenological markers that can be positively or negatively valenced. As a basic human emotion (Ekman 1992), surprise is universally recognizable, though inherently relational and idiosyncratic.

At the neural-computational level, the term by which surprise is theorized as fundamental to human cognition is *prediction error* (Schultz et al. 1997). Prediction error is the difference between the expected event and the outcome. If there is a discrepancy, then a prediction error occurs and flags requirements for learning. For example, if the prediction is of business as usual, and a change occurs (for example, an “appetitive event,” perhaps a shot of melted chocolate for humans such as these authors), dopaminergic neurons fire. If there is a consistent stimulus that can predict the shot of melted chocolate (for example, a particular sound), then the sound itself (if unexpected) can trigger the dopamine response (with the response having shifted through *error correction* from the reaction to the shot of melted chocolate to the sound). If the sound is made and no shot of melted chocolate is forthcoming, there is a decrease in firing of dopaminergic neurons. This variation in firing of dopamine is proposed to provide the basis of learning and conflict detection (Alexander and Brown 2019), which themselves underlie a vast array of other cognitive processes. Surprise as prediction error is a highly robust finding and has been used to model (as noted above) many and varied cognitive processes.

Nevertheless, how the feeling of surprise enters *awareness* is unclear. We do not stroll through the world constantly surprised that our predictions are correct or incorrect; we learn

through the difference between predicted and actual interactions with the world, but each interaction is not overwhelmingly memorable (Fazio and Marsh 2009). Yet moments of surprise *are* memorable and often accompanied by a strong affective component (Ekman and Friesen 1971; Reisenzein and Studtmann 2007). Some models (such as the Alexander and Brown model noted above) propose multiple levels of error detection. In this case, it may be that sufficiently high levels of error detection also trigger the feeling of surprise as we know it, as well as sufficiently high levels of activation for consciously linking that feeling with the surprising event. It may also be that different magnitudes of prediction error (across multiple circuits) could offer an explanation (Hélie and Sun 2010) in their model of insight, and propose a similar mechanism by which an insight is reached. In either case, it may be that surprise can indicate the point at which relying on an unconscious updating of the current knowledge state surpasses tolerable bounds.

Surprise as a feeling is therefore an important component of what motivates an individual to explore the world and knowledge space. As Clarke (2018, p. 521) suggests: “We humans often seem to actively seek out surprising events, deliberately harvesting novel and exciting streams of sensory stimulation.” This perspective points not only to the importance of surprise in human cognition and behavior but also to our role in generating it through epistemic “foraging.” In this way, rather than casting the human as a passive recipient of the surprising moment, surprise can be seen as part of a process of the flow of actions and events. It is something we actively create through continual explorative action (Austin 1979). In short, surprise in all its forms marks the experience of prodding at an unpredictable world.

Theories of prediction error and surprisal underlie the proposal of active inference (for a recent overview, see Friston et al. 2017), a view of cognition which is increasing in popularity (Hohwy 2013). This theory proposes that cognition is motivated by a desire to minimize prediction error by reducing uncertainty. While surprise

is inevitable in an uncertain world, a strict approach to prediction error minimization proposes that an organism will seek to *minimize* surprise through actions which decrease the difference between mental models and the external stimuli, and minimize surprisal in all its forms. There is no clear motivation to increase prediction error because this is cognitively inefficient. We see that we have to account for this *desire* for phenomenological surprise as well as cataloguing moments of neural surprisal. The novelty-seeking nature of phenomena such as creativity then becomes essential for such models to explain (to avoid the so-called Darkened Room scenario by which an agent ends up in the most predictable environment – a dark room; see Gallagher in prep). Clark (2018) addresses the role of surprise seeking in theories of active inference and ends up accounting for it by recognizing the importance of complex social and cultural environments, what he calls “exotic statistical baths” (p. 532). These environmental structures scaffold epistemic foraging rather than it being something which can be explained by a purely internal model. Ultimately, then the feeling of surprise and its metacognitive and epistemic function cannot be accounted for on a neural level alone and requires us to understand cognition as affective and extended.

## The Epistemic Role of Surprise

Surprise refers to a feeling of sharp, sudden, and unexpected expansion of the possibility space (see ► “the Adjacent Possible”, this volume). Importantly, this expansion is one which results in an updating of the current knowledge state rather than a rejection of the intrusion (Steinbock 2018). Expressing surprise is not the same as expressing skepticism; we are surprised by something which is evidentially true but which we have not considered before. This is clearly illustrated when we say “That would surprise me if it *were* true” where the level of surprise is related to the truth of the object. We are not surprised by something we do not believe; surprise requires us to grant the truth of the proposition which surprises

us. It is this rapid recognition of new knowledge that accompanies the confidence of an insight experience. Surprise is therefore epistemically important: an epistemic feeling.

Epistemic feelings are those feelings which motivate the exploration of the world and knowledge space (Arango-Muñoz and Michaelian 2014). They are elements of the emotional spectrum closely connected to knowledge and knowledge exploration. However, surprise differs from most epistemic feelings because it is initially not directed toward action rather it is in a reaction to actions and events. However, while it is complete in itself, it is not a short-lived phenomenon with no additional effects. The updating of the knowledge base requires action to incorporate this new knowledge which leads to exploratory feelings such as curiosity and wonder. Johnson-laird and Oatley (1992) argue that Ekman’s basic emotions each serve an important function. In the case of surprise, we suggest the feeling of surprise serves a metacognitive function, that of forcing a reorientation of epistemic activity. It is the cessation and disruption of the existing expectations which marks the threshold into consciousness. Therefore, while surprise has an epistemic element, it is a moment not of exploration (epistemic or otherwise) but rather one of stopping of action and diverting attention.

Importantly, the new knowledge marked by the feeling of surprise comes from the process of interacting with the world (Glăveanu [forthcoming](#)) lending weight to arguments that engagement is important in understanding creativity and innovation. Surprise marks the moment when external stimuli challenge internal stability; it is impossible to be surprised by something of which you are already aware (although you may become surprised by an unanticipated perspective). Instead, surprise marks a moment when the world comes into a novel focus and a person’s perspective on that world changes. It is an epistemic judgment related to prior expectations and the likelihood of an event. It brings into focus the unexpectedness of being in the world and also destabilizes ideas of unrestrained human agency and the imposition of that agency on inert material. In this way, surprise marks the realization that what could be is wildly

different from what is. It is, thus, the epistemic feeling most closely associated with the unanticipated possible.

Additionally, this requirement for interaction undergirding the feeling of surprise reminds us of the distributed nature of knowledge. The feeling of surprise is a marker of the moment of incongruence between external stimuli and a mental plan or perspective forcing it to be updated. As Gerlain (2019) notes, it reminds us of an implicit world outside the world of our current experience. When it comes to the relationship between surprise and the possible, surprise marks the moment when the unexpected, unconsidered, and seemingly impossible are recognized as possible and supports the disambiguation of the possible from the merely probable.

### Surprise as a Necessary Component of Insight

Moments of insight are universally recognized as instances of realization in which feelings of sudden confidence arise for an answer. This answer may have been actively searched for (e.g., when trying to solve an ill-defined problem in computer programming), or not actively searched for (e.g., realizing the unsought solution to an issue while showering). The definitional components of the feeling, whether accompanied by pleasure or dismay, are that of *sudden confidence* (Cushen and Wiley 2012). Moments of insight provide a disjunct between the previous representation of the problem and the current representation of the problem, with no conscious memory of the problem-solving pathway between the two.

The feeling accompanying moments of insight, that of surprising certainty in a solution, is remarkably difficult to elicit on a reliable basis, partly because, as noted above, surprise is relational. The feeling of insight is generally termed an “aha moment” in the relevant field: This is the feeling of sudden confidence in the answer (see Insight, this volume). Danek et al. (2014) identified several aspects of the feeling of insight phenomenology of which surprise was the second strongest. This confirmed previous theorizing that placed surprise as a key marker of the insight experience. For example, Gick and Lockhart (1995) suggest that the transition to a solution state which marks

the moment of insight is marked by surprise that the representation required to solve the problem is different to the one held in mind by the problem solver. Surprise is also closely related to insight radicality (Jarman 2014), that is, the difference between the previous (unhelpful) problem representation and the novel one. Jarman found that increased radicality predicted the intensity of the insight experience leading us to suggest increased surprise is key to the experience.

Despite being a cornerstone in the phenomenon of insight (as a sudden and untraceable moment of realization that the original representation of a problem was incorrect), surprise is rarely investigated as an emotional component of problem solving. However, it could be that it provides an indication of the underlying cognitive processes. There is some debate about what the cognitive processes are that underlie insight. In brief, the question is whether insightful problem solving relies on a particular special process or whether it comes from the unconscious application of the same processes which underlie all cognition (Gilhooly et al. 2015). Although it seems likely that the answer relies on a mix of both, it could be that surprise indicates which is most dominant at any one time.

In order to solve a problem with insight, a problem solver must have an unpredicted shift in understanding. The nature of this shift is currently unclear – does it represent a shift in the problem representation or a shift in the understanding of the way to solve the problem. In this way, surprise is obviously critical to theories which propose sudden representational change and non-monotonic problem solving (Danek 2018) because it marks a sharp change in the representation from previous iterations. In contrast, current theories of incremental progress (such as outlined by MacGregor et al. 2001) fail to address the feeling of surprise, but it may be that surprise in this model occurs at a different point, that is, the point at which the problem solver realizes that the approach they were using to solve the problem will not work, so-called criterion failure. In this case, surprise is not generated by the solution but by the moment that it becomes apparent that the confidence in the previous solution attempt is

misplaced. A fuller appreciation of the epistemic role of surprise could scaffold further attempts to adjudicate between the two theories.

### Surprise as a Threshold Marker: The Case of Serendipity

We have all had moments of serendipity – a moment of good fortune which seems completely suited to us. Serendipity refers to the interaction of an accident and the personal “sagacity” to exploit that accident. The etymological roots of accident imply something entering from outside the system. It comes from the Latin *accidere* meaning “to fall down, impinge on, be heard, happen” (Merriam-Webster), and serendipity is firmly tied to the “act of noticing” (Rubin et al. 2011). At the heart of the notion of a serendipitous accident is that it comes from outside the existing cognitive system, involves a disruption, and that this disruption is significant enough to be noticed. These aspects can be collapsed into one: A serendipitous accident is surprising. In this way, much like for insight, surprise is at the heart of the moment of serendipity.

As with insight, surprise also gives us an interesting perspective on the mechanisms which may underlie serendipitous cognition. In this case, we are returned to the notion of a threshold which is marked by the feeling of surprise. Serendipity is a disruptive phenomenon requiring the agent to turn away from one path to follow another prompted by something unexpected. Serendipity is necessarily sparked by an accident (Makri and Blandford 2012; McCay-Peet and Toms 2015; Merton and Barber 2004). However, human activity takes place against a continual flow of accidents so the occurrence of an accident cannot be enough to spark the sense of serendipity, rather an accident is something which acts as a “game changer” (Arfini et al. 2018), and it is this which requires the element of surprise. It is not yet clear what makes an accident become important enough to be selected and enacted (see Ross, 2022), but it may be an investigation into the nature of surprise can lend support to attempt to understand these complexities.

In other words, because serendipity is something which is consciously experienced, the

accident must be noticed consciously and assessed as interesting enough to be exploited. Models of surprise as a metacognitive signal can give us information on how that assessment takes place. Take, for example, the model proposed by Munnich and Ranney (2019): They offer three responses to learning new surprising information: accommodating it, treating it as an anomaly, and rejecting it. Each of these consists of three aspects – the existing belief network, the event, and the action which is undertaken. We reject an idea when it is too difficult to explain within the context of our own belief network. As Arfini et al. (2018, p. 5) write:

Fleming’s “Oh!” reaction was when he managed to frame and understand the antibiotic effect of a mold. He did not enter his laboratory to find a moldy culture singing the chorus of Mamma mia! that would have sparked another kind of reaction.

Things which are too distant from our belief network, where the radicality is too great, are not drawn upon to update it. The same is true of serendipitous accidents; those which fall outside the realm of what we see as possible are not followed up.

Similarly, there is a resource cost to exploiting accidents, and option D in the model encompasses the idea that an interesting anomaly is noticed, but it is not considered interesting enough to invest the time and cost. Barber and Fox (1958) tell the story of two scientists who are surprised by the same thing, but one has the resources and time to follow it up and the other does not. The realization that surprise-driven curiosity has a cost is crucial to understanding why some accidents are exploited and others are not.

In other words, serendipitous surprise relies on the event being sufficiently close to what the system views as possible; even if the knowledge is not there before, it requires the deviation from an existing theory rather than uncontextualized knowledge (Yaqub 2018). This is because serendipity requires the enacting of the unplanned event, and the boundaries of the system must be flexible and dynamic in order to assimilate this new knowledge. This is a quasi-liminal knowledge state which requires both expertise (to recognize and assimilate) and ignorance

(to experience surprise). An explanation for the role of serendipity could be related to the threshold required to experience surprise.

## Surprise and The Possible

As we noted in our opening paragraph, possibility refers to that which is not actual. Part of the difficulty in conceptualizing the use of other possible worlds is that those worlds are often rooted in our existing experiences. Possibility thinking is best conceived of as a process, a continuous redefining and moving forward through adjacent possibles to generate novel possibilities. Surprise provides an emotional identifier of when the possible has become internally actualized, as a marker of how we can interact with hitherto unanticipated domains: the marker of the actualized and expanded possible.

## References

- Alexander, W. H., & Brown, J. W. (2019). The role of the anterior cingulate cortex in prediction error and signaling surprise. *Topics in Cognitive Science*, *11*(1), 119–135. <https://doi.org/10.1111/tops.12307>.
- Arango-Muñoz, S., & Michaelian, K. (2014). Epistemic feelings, epistemic emotions: Review and introduction to the focus section. *Philosophical Inquiries*, *2*(1), 97–122. [10/ggd6zt](https://doi.org/10/ggd6zt)
- Arfini, S., Bertolotti, T., & Magnani, L. (2018). The antinomies of serendipity: How to cognitively frame serendipity for scientific discoveries. *Topoi*. <https://doi.org/10.1007/s11245-018-9571-3>.
- Austin, J. H. (1979). The varieties of chance in scientific research. *Medical Hypotheses*, *5*(7), 737–741. [https://doi.org/10.1016/0306-9877\(79\)90035-5](https://doi.org/10.1016/0306-9877(79)90035-5).
- Barber, B., & Fox, R. C. (1958). The case of the floppy-eared rabbits: An instance of serendipity gained and serendipity lost. *American Journal of Sociology*, *64*(2), 128–136. <https://doi.org/10.1086/222420>.
- Clark, A. (2018). A nice surprise? Predictive processing and the active pursuit of novelty. *Phenomenology and the Cognitive Sciences*, *17*(3), 521–534. <https://doi.org/10.1007/s11097-017-9525-z>.
- Cushen, P. J., & Wiley, J. (2012). Cues to solution, restructuring patterns, and reports of insight in creative problem solving. *Consciousness and Cognition*, *21*(3), 1166–1175. <https://doi.org/10.1016/j.concog.2012.03.013>.
- Danek, A. H. (2018). Magic tricks, sudden restructuring and the aha! Experience. In F. Vallée-Tourangeau (Ed.), *Insight: On the origin of new ideas* (pp. 51–79). Routledge.
- Danek, A. H., Fraps, T., von Maller, A., Grothe, B., & Ollinger, M. (2014). It's a kind of magic: What self-reports can reveal about the phenomenology of insight problem solving. *Frontiers in Psychology*, *5*. <https://doi.org/10.3389/fpsyg.2014.01408>.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3–4), 169–200. <https://doi.org/10.1080/02699939208411068>
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124–129. <https://doi.org/10.1037/h0030377>.
- Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review*, *16*(1), 88–92. <https://doi.org/10.3758/PBR.16.1.88>.
- Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: A process theory. *Neural Computation*, *29*(1), 1–49. [https://doi.org/10.1162/NECO\\_a\\_00912](https://doi.org/10.1162/NECO_a_00912)
- Gerlain, A. (2019). How implicit is surprise? In N. Depraz & A. Celle (Eds.), *Surprise at the intersection of phenomenology and linguistics* (Vol. 11, pp. 77–91). John Benjamins Publishing Company. <https://doi.org/10.1075/ceb.11>.
- Gick, M. L., & Lockhart, R. S. (1995). Cognitive and affective components of insight. In R. J. Sternberg & J. Davidson (Eds.), *The nature of insight* (pp. 197–228). MIT Press.
- Gilhooly, K. J., Ball, L. J., & Macchi, L. (2015). Insight and creative thinking processes: Routine and special. *Thinking & Reasoning*, *21*(1), 1–4. <https://doi.org/10.1080/13546783.2014.966758>.
- Glăveanu, V. P. (2021). *The possible: A sociocultural theory*. Oxford University Press.
- Glăveanu, V. P. (forthcoming). What's inside the prepared mind. In W. Ross & S. Copeland (Eds.), *The art of serendipity*. Palgrave.
- Hélie, S., & Sun, R. (2010). Incubation, insight, and creative problem solving: A unified theory and a connectionist model. *Psychological Review*, *117*(3), 994–1024. <https://doi.org/10.1037/a0019532>.
- Hohwy, J. (2013). *The predictive mind* (1st ed.). Oxford University Press.
- Jarman, M. S. (2014). Quantifying the qualitative: Measuring the insight experience. *Creativity Research Journal*, *26*(3), 276–288. <https://doi.org/10.1080/10400419.2014.929405>.
- Johnson-laird, P. N., & Oatley, K. (1992). Basic emotions, rationality, and folk theory. *Cognition and Emotion*, *6*(3–4), 201–223. <https://doi.org/10.1080/02699939208411069>.
- MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2001). Information processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning*,

- Memory, and Cognition*, 27(1), 176–201. <https://doi.org/10.1037//0278-7393.27.1.176>.
- Makri, S., & Blandford, A. (2012). Coming across information serendipitously – Part 1: A process model. *Journal of Documentation*, 68, 684–705. <https://doi.org/10.1108/00220411211256030>.
- McCay-Peet, L., & Toms, E. G. (2015). Investigating serendipity: How it unfolds and what may influence it. *Journal of the Association for Information Science and Technology*, 66, 1463–1476. <https://doi.org/10.1002/asi.23273>.
- Merton, R., & Barber, E. (2004). *The travels and adventures of serendipity: A study in sociological semantics and the sociology of science*. Princeton University Press.
- Munnich, E., & Ranney, M. A. (2019). Learning from surprise: Harnessing a metacognitive surprise signal to build and adapt belief networks. *Topics in Cognitive Science*, 11(1):164–177. <https://doi.org/10.1111/tops.12397>.
- Nevo, I., & Erev, I. (2012). On surprise, change, and the effect of recent outcomes. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00024>.
- Reisenzein, R., & Studtmann, M. (2007). On the expression and experience of surprise: No evidence for facial feedback, but evidence for a reverse self-inference effect. *Emotion*, 7(3), 612–627. <https://doi.org/10.1037/1528-3542.7.3.612>.
- Rocca, E., Copeland, S., & Ralph Edwards, I. (2019). Pharmacovigilance as scientific discovery: An argument for trans-disciplinarity. *Drug Safety*. <https://doi.org/10.1007/s40264-019-00826-1>.
- Ross, W. (2022). Heteroscalar serendipity and the importance of accidents. In W. Ross & S. Copeland (Eds.), *The Art of Serendipity*. Palgrave MacMillan.
- Rubin, V. L., Burkell, J., & Quan-Haase, A. (2011). Facets of serendipity in everyday chance encounters: A grounded theory approach to blog analysis. *Information Research*, 16. Available at <http://InformationR.net/ir/16-3/paper488.html>
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599. <https://doi.org/10.1126/science.275.5306.1593>.
- Steinbock, A. J. (2018). Surprise as emotion: Between startle and humility. In N. Depraz & A. J. Steinbock (Eds.), *Surprise: An emotion?* (pp. 3–21). Berlin Heidelberg: Springer.
- Yaqub, O. (2018). Serendipity: Towards a taxonomy and a theory. *Research Policy*, 47, 169–179. <https://doi.org/10.1016/j.respol.2017.10.007>.