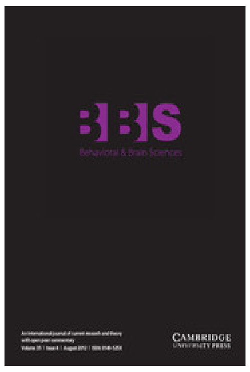
Behavioral and Brain Sciences

http://journals.cambridge.org/BBS



Implicit learning: What does it all mean?

David R. Shanks and Mark F. St. John

Behavioral and Brain Sciences / Volume 19 / Special Issue 03 / September 1996, pp 557 - 558
DOI: 10.1017/S0140525X00082194, Published online: 19 May 2011

Link to this article: http://journals.cambridge.org/abstract_S0140525X00082194

How to cite this article:
David R. Shanks and Mark F. St. John (1996). Implicit learning: What does it all mean?. Behavioral and Brain Sciences,19, pp 557-558 doi:10.1017/S0140525X00082194

Sometimes I seem to learn an infinity of things at once. When I am told that "multiplication is associative," I seem to learn that 1 × 2 = 2 × 1, 4 × 5 = 5 × 4, and so forth. I find it hard to believe that I could have been conscious of those infinitely many facts when I learned them, even though I "recognize" them as true when I hear them.

And sometimes I seem to learn things that I could not possibly be aware of. When I learn to swing a golf club properly, I learn to use certain muscles in certain ways. I am not aware of which muscles I use and how I use them differently. I find it hard to believe that I was aware of those things when I learned them. What I was aware of was what the golf pro said to me. But that was declarative, and what I learned (how to swing) was procedural. It took lots of practice to translate what was told (the declarative knowledge) into what I can do (the procedural knowledge). I was almost certainly aware of the former, but it is hard for me to see in what sense I might have been aware of the latter.

S&S seem to have a narrower view of learning than I do. They might say that I did not learn (say) that "France has a city that begins with a 'P'" from "Paris is the capital of France" when I heard "Paris is the capital of France." I inferred it later. I prefer to say that one *learns x at time t* if one did not know *x* before time *t* and knows it afterwards. Under my definition, all three of my examples would count as learning, and therefore as implicit learning.

It is almost impossible to prevent humans from learning far more than they could reasonably be aware of. Awareness would have to be very clever to latch onto everything one learns from "Paris is the capital of France" and very big to hold them all.

As a computer scientist, I look at implicit learning from a point of view somewhat different from S&S's. When I put things into a computer's memory, a great deal depends on where I put it and how it is therefore connected to the other information in that memory.

When I put a fact, like "Paris the capital of France," into a *database* it affects the future behavior – the way the system responds to queries – in all sorts of ways that I, who know both the fact and the organization of the database, cannot determine.

When I put a rule like "From X is the capital of France to infer X is Paris" into an *expert system*, I am aware of the rule. What I am not aware of is the way that the rule will influence the future responses of the system.

When I adjust the parameters in a *neural net* to represent the fact that Paris is the capital of France, it is hard for me to see how those parameter settings could possibly be the objects of the system's awareness.

It is not just that I am not aware of what such systems "learn" when a new piece of information is put into them. I cannot imagine how I could be aware. There is even a theorem, a consequence of Turing's (1936) proof of the unsolvability of the halting problem, that tells me that (if I have only the power of a computing machine) I cannot even figure out, in general, what all the consequences of making a change in a computer's data or program are. If I can't figure them out, how can I be aware of them?

Why, you might ask, should we care?

It seems to me that we should care because almost everything we learn that really matters is learned implicitly. If school teachers believed that all learning was explicit, they would miss most of what students learn in school. If I am right, what students learn to do, they learn implicitly. Nobody tells students that mathematics is difficult or that history is boring. But most of them learn it. And, because they learn it implicitly, the learning lasts and is hard to change.

There is a lot of merit in Shanks's & St. John's criticisms of attempts to demonstrate implicit learning. But the fact that these attempts have not wholly succeeded does not mean that what they are looking for does not exist. It may be hard to "see" implicit learning, but that does not necessarily mean that it is not there.

If you are piloting a ship, you can only see part of an iceberg. But it is useful to remember that it is the part you cannot see that you should worry about.

Authors' Response

Implicit learning: What does it all mean?

David R. Shanks[a] and Mark F. St. John[b]

[a]*Department of Psychology, University College London, London WC1E 6BT, United Kingdom;* [b]*Department of Cognitive Science, University of California at San Diego, La Jolla, CA 92093.* **david.shanks@psychol.ucl.ac.uk; mstjohn@cogsci.ucsd.edu**

Abstract: In the original target article (Shanks & St. John 1994), one of our principal conclusions was that there is almost no evidence that learning can occur outside awareness. The continuing commentaries raise some interesting questions, especially about the definition of learning, but do not lead us to abandon our conclusion.

Kugel's entertaining and provocative commentary presents a challenge to our arguments that did not arise in the earlier round of discussion. We claimed that most human learning is accompanied by awareness of what is being learned, but Kugel disputes this idea by proposing a new definition of learning that makes it implausible in the extreme that all learning is conscious. Since this definition has some superficial plausibility, it behooves us to look at Kugel's argument carefully.

To begin with, we must note that there is a deep confusion in the way learning is defined in **Kugel**'s commentary. He proposes that when a person learns that *multiplication is associative*, what is actually learned is just a set of facts (1 × 2 = 2 × 1, 4 × 5 = 5 × 4, etc.). But there is a major ambiguity here. The "definition" either means (1) that there is a set of facts (e.g., 1 × 2 = 2 × 1) to which people will now assent but to which they would not have assented previously, or (2) is circular. This circularity follows from the fact that learning is defined in terms of itself ("When I *learn* that *multiplication is associative*, I seem to *learn* that 1 × 2 = 2 × 1 . . . "). On the latter conception, how do we know what it is to learn that 1 × 2 = 2 × 1? If learning that 1 × 2 = 2 × 1 is defined in terms of another set of things that are learned (e.g., 1 × 2 × 3 = 2 × 1 × 3, 1 × 2 × 438 = 2 × 1 × 438, etc.), then we face an infinite regress with no true definition of learning in prospect. We will proceed on the assumption, therefore, that it is version (1) that Kugel is defending, because version (2) just does not seem to be a definition at all.

Version (1) defines learning in terms of a set of new behaviors or dispositions – it is, then, just a standard behaviorist conception of learning. Seen in this light, it should hardly be necessary for us to go into detail to show why it is inadequate, since this is a job that has been done for us by fifty years of philosophical and psychological debate (see Bechtel 1988). Briefly, the main difficulty with a behaviorist conception of learning is that it provides no basis for generalization. If I observe somebody assenting to the propositions 1 × 2 = 2 × 1, 4 × 5 = 5 × 4, 3 × 8 = 8 × 3, 127 × 458 = 458 × 127, I can make the reasonable prediction that they will also agree that 17 × 6 = 6 × 17. But this prediction is only licensed if I assume that the person has learned the proposition that *multiplication is associative*, which represents a generalization covering the observed behaviors. If all that has been acquired from the learning episode is a set of new behaviors and behavioral dispositions, then it must remain unexplained that the person, when tested, agrees that 17 × 6 = 6 × 17.

Kugel argues that it is often impossible to predict the way a system will behave when given a new fact; and, because learning is defined in terms of a set of new dispositions, Kugel concludes from this that learning cannot possibly be conscious. However, there are two difficulties here. First, it does not make sense to talk about a disposition being conscious or unconscious since it is a potential, not a real thing (and of course true behaviorists would never use terms like "conscious" or "unconscious" in the first place). But a disposition can no more be conscious than unconscious, so Kugel must be equally committed to denying that any learning is conscious: there can be no such thing as explicit learning. Yet Kugel seems unaware of this commitment, because he readily agrees that he can be aware of what the golf pro says to him. How can this be the case if what he learns from the golf pro is just a set of dispositions? Second, we dispute the claim that it is often impossible to predict the way a system will behave when given a new fact (at least with respect to standard symbolic databases). It may be *difficult* to predict behavior, but it is still the case that in such systems behavior is determined by inferences made over stored facts.

On our view, learning should be construed in terms of the acquisition of mental states whose content fixes what is learned. The content of the acquired knowledge is not picked out by specifying a large set of new behaviors and behavioral dispositions. Rather, these behaviors and dispositions are caused by mental states that are implicit in what is originally learned (we are using "implicit" here in the informational sense of Kirsh 1990) rather than being explicitly represented.

It is undoubtedly true that the term "learning" is extremely hard to define (indeed **Kugel**'s confusion attests to this). But we doubt that the behaviorist definition that Kugel seems to be advocating – which would admittedly challenge the thesis of our target article – is sustainable.

Furedy & Kristjansson confirm what we concluded in the target article, namely, that CS-US awareness is necessary for human Pavlovian conditioning (i.e., conditioning does not occur without awareness). Contrary to what they suggest, we made no claims about the possible causal relationship of the one in producing the other. The indisputable fact that awareness is not sufficient for conditioning seems to have no bearing on the central issues of our target article. Furedy & Kristjansson have a curious conception of conditioning, which we have previously commented upon (Shanks 1990).

References

Baer, P. E. & Fuhrer, M. L. (1982) Cognitive factors in the concurrent differential conditioning of eyelid and skin conductance responses. *Memory and Cognition* 10:135–140. [JJF]

Bechtel, W. (1988) *Philosophy of mind: An overview for cognitive science.* Erlbaum. [DES]

Brewer, W. F. (1974) There is no convincing evidence for operant or classical conditioning in adult humans. In: *Cognition and the symbolic processes*, ed W. B. Weimer & D. S. Palermo. Erlbaum.[JJF]

Dawson, M. E. (1973) Can classical conditioning occur without contingency learning? A review and evaluation of the evidence. *Psychophysiology* 10:82-86. [JJF]

Dawson, M. E. & Furedy, J. J. (1976) The role of awareness in human differential autonomic classical conditioning: The necessary-gate hypothesis. *Psychophysiology* 13:50–53. [JJF]

Furedy, J. J. (1973) Some limits on the cognitive control of conditioned autonomic behavior. *Psychophysiology* 10:108–111. [JJF]

Furedy, J. J. (1988) Arguments for and proposed tests of a revised S-R contiguity-reinforcement theory of human Pavlovian autonomic conditioning: Some contra-cognitive claims. *Biological Psychology* 27:77–78. [JJF]

Furedy, J. J. (1990) Sharing a common language about conditioning requires accurate characterizations of each others' positions: Reply to Shanks. *Biological Psychology* 30:181–187. [JJF]

Furedy, J. J. (1991) Alice-in-Wonderland terminological usage in, and communicational concerns about, that peculiarly American flight of technological fancy: The CQT polygraph. *Integrative Physiological and Behavioral Science* 26:241–247. [JJF]

Furedy, J. J. (1992) Reflections on human Pavlovian decelerative heart-rate conditioning with negative tilt as US: Alternative approaches. *Integrative Physiological and Behavioral Science* 27:347–355. [JJF]

Furedy, J. J., Arabian, J. M., Thiels, E., & George, L. (1982) Direct and continuous measurement of relational learning in human Pavlovian conditioning. *Pavlovian Journal of Biological Science* 17:69–79. [JJF]

Furedy, J. J. & Poulos, C. X. (1976) Heart-rate decelerative Pavlovian conditioning with tilt as UCS: Towards behavioral control of cardiac dysfunction. *Biological Psychology* 4:93–106. [JJF]

Furedy, J. J. & Riley, D. M. (1987) Human Pavlovian autonomic conditioning and the cognitive paradigm. In: *Conditioning in humans*, G. Davey. Wiley & Sons. [JJF]

Furedy, J. J. & Schiffmann, K. (1973) Concurrent measurement of autonomic and cognitive processes in a test of the traditional discriminative control procedure for Pavlovian electrodermal conditioning. *Journal of Experimental Psychology* 100:21–217. [JJF]

Jones, J. E. (1962) Contiguity and reinforcement in relation to CS-UCS intervals in classical aversive conditioning. *Psychological Review* 69:176–186. [JJF]

Kirsh, D. (1990) When is information explicitly represented? In: P. P. Hanson (ed.), *Information, language, and cognition*, University of British Columbia Press. [DES]

Lovibond, P. F. (1992) Tonic and phasic electrodermal measures of human aversive conditioning with long duration stimuli. *Psychophysiology* 29:621–632. [JJF]

Marinkovic, K., Schell, A. M., & Dawson, M. E. (1989) Awareness of the CS-UCS contingency and classical conditioning of skin conductance responses with olfactory CSs. *Biological Psychology* 29:39–60. [JJF]

Rescorla, R. A. (1967) Pavlovian conditioning and its proper control procedures. *Psychological Review* 74:71–80. [JJF]

Rescorla, R. A. (1988) Pavlovian conditioning: It's not what you think it is. *American Psychologist* 43:151–160. [JJF]

Rescorla, R. A. & Wagner, A. R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In: *Classical conditioning*, (vol. 2), eds., A. H. Black & W. F. Prokasy. Appleton-Century-Crofts. [JJF]

Schiffmann, K. & Furedy, J. J. (1977) The effect of CS-US contingency variation on GSR and on subjective CS/US relational awareness. *Memory and Cognition* 5:273–277. [JJF]

Shanks, D. R. (1990) On the cognitive theory of conditioning. *Biological Psychology* 30:171–179. [DRS]

Shanks, D. R. & St. John, M. F. (1994) Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences* 17:367–395. [JJF, PK]

Stewart, M., Stern, J. A., Winokur, G., & Fredman, S. (1961) An analysis of GSR conditioning. *Psychological Review* 687:60–67. [JJF]

Turing, A. M. (1936). On computable numbers with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* 42:230–265. [PK]