

Toward a Model of Impaired Reality Testing in Rats

Michael McDannald² and Geoffrey Schoenbaum^{1,2}

²Department of Anatomy and Neurobiology, University of Maryland School of Medicine, Baltimore, MD 21201

Schizophrenia is a chronic brain disorder that affects about 1.1% of the adult US population annually. Hallucinations, delusions, and impaired reality testing are prominent symptoms of the disorder. Modeling these symptoms is difficult because it is unclear how to assess impaired reality testing in animals. Animals cannot discuss their beliefs; however, a century of learning experiments has shown us that they, like us, construct complex internal representations of their world. Presumably, these representations can become confused with reality for animals in much the same way that they do for schizophrenic patients. Indeed, there is evidence from studies of Pavlovian conditioning that this happens even in normal animals. For example, early in training a cue that has been paired with reward elicits a highly realistic, sensory representation of that reward, which is to some extent indistinguishable from reality. With further training, this sensory hallucination of reward is replaced by a more abstract representation, termed a reward expectancy. Reward expectancies reflect the sensory and other qualities of the impending reward but are distinguishable from the actual reward. Notably, the hallucinatory representations depend on subcortical regions, such as amygdala, whereas reward expectancies require the progressive involvement of prefrontal areas, such as orbitofrontal cortex. Abnormal prefrontal function is associated with schizophrenia; impaired reality testing may result from a failure of the normal shift from highly realistic, sensory representations to more abstract, prefrontal expectancies. The Pavlovian procedures discussed here could be applied to animal models and schizophrenic patients to test this hypothesis.

Key words: schizophrenia/hallucination/delusion/devaluation/orbitofrontal/rat

Introduction

Consider a simple situation in which a hungry rat is placed in an experimental chamber. A tone is presented, and, following its termination, a sweet, sugar solution is delivered. The rat will learn after only a few trials that the tone predicts sugar. This is evident in the rat's behavior, which changes during the tone to reflect the impending sugar delivery. Yet, what is the rat "thinking" during the tone? It turns out that the tone evokes specific properties of the sugar, and the nature of these evoked properties changes in important ways from the earliest stages of training to later stages. As we will show, these changes reflect a shift or transition from a highly realistic, sensory representation of the reward, which the rat has trouble distinguishing from reality, to a more abstract representation of the expected reward, which the rat can readily distinguish from reality.

The evidence for this shift comes from studies done by Holland and colleagues using nausea to manipulate the value of the reward. It is well established that if one pairs a reward, such as a sugar solution, with nausea, rats will reduce consumption of the reward thereafter. The reward has been devalued. Amazingly, Holland found that if rats were presented with pairings of a tone and sugar and then made nauseous in the presence of the tone alone, the rats subsequently reduced their consumption of the sugar solution.¹ This effect is referred to as "mediated devaluation" because it is observed even though the sugar is not present during the nausea. Rather devaluation was mediated by the representation of sugar evoked by the tone. This tone-evoked representation was so realistic that the rat mistook it for actual sugar.

This inability to distinguish between a sensory representation and the presence of sugar is conceptually somewhat similar to the impaired reality testing that characterizes hallucinations and delusions in schizophrenic patients. Like schizophrenic patients, who misinterpret cues and other input from their environment to support simple and complex beliefs that are at odds with reality, the rats have misinterpreted the tone as reflecting the presence of reward, even though the reward is not present. Notably, in normal rats, this effect is only found very early in training; with further pairings of the tone and sugar solution, the tone rapidly loses its ability to mediate devaluation of the sugar solution (figure 1).

¹To whom correspondence should be addressed; University of Maryland School of Medicine, Departments of Anatomy and Neurobiology and Psychiatry, 20 Penn Street, HSF-2, Rm S251, Baltimore, MD 21201; tel: 410-706-3814, fax: 410-706-2512, e-mail: schoenbg@schoenbaumlab.org.

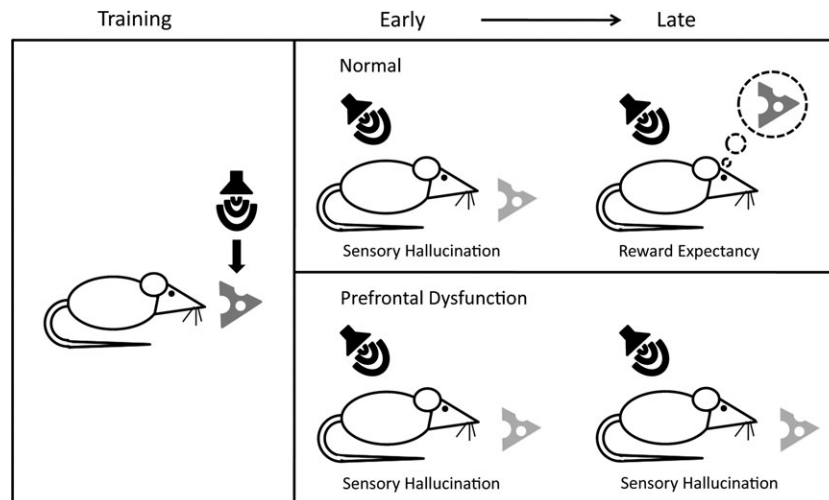


Fig. 1. A Cartoon Depicts Sensory Hallucinations and Reward Expectancies Over the Course of Pavlovian Conditioning in Normal Rats and Our Hypothesis for Rats With Prefrontal Dysfunction. Here, training consists of a tone that predicts cheese (left panel). In normal rats (top right panel) early in training presentation of the tone alone will produce a sensory hallucination of cheese that, to some extent, is not distinguishable from cheese itself. Late in training, the representation evoked by the tone becomes distinguishable from that evoked by cheese, corresponding to an expectancy of cheese. That is, the tone does not mean cheese is present but will soon be delivered. If this transition from sensory hallucination to reward expectation depends on prefrontal control then rats with prefrontal dysfunction might show normal behavior early in learning (bottom right panel). However, as training continues there will be no transition to expectancy, resulting in a persistent hallucination of cheese after normal rats have begun to generate an expectancy of cheese.

Thus, the representation of the sugar reward evoked by the tone cue becomes less real or somehow more distinguishable from that evoked by actual sugar reward.

Interestingly, the loss of the tone's ability to support mediated devaluation is paralleled by the development of the tone's ability to evoke what have been called reward expectancies—predictions that sugar is about to be received.² The operation of these expectancies can be demonstrated with reward devaluation. In reward devaluation, many pairings of tone and sugar are given. Following extensive training, an aversion is formed by pairing the sugar directly with nausea. Finally responding to the tone alone is tested. Normal rats exhibit a spontaneous decrease in responding to the tone following devaluation of the reward.^{3,4} This result would only occur if the rats had formed an expectancy of sugar in the presence of the tone. Thus, the early sensory hallucination of sugar is replaced by a more abstract representation of the expected sugar delivery; this representation is no longer confused with actual sugar and can still be used to guide behavior appropriately (figure 1).

Neural Mechanisms Underlying Hallucinations and Expectancies

The underlying associative processes that support sensory hallucinations and reward expectancies appear to depend, in part, on different neural substrates. Dwyer and Killcross⁵ have shown that the basolateral amygdala (BLA) is critical to the ability of cues to support mediated devaluation. They found that when rats were exposed to

a context predictive of a specific reward and then made ill, consumption of that specific reward was later reduced. This was not true for rats with bilateral, neurotoxic BLA lesions. The deficit of BLA-lesioned rats cannot be attributed to deficits in forming taste aversions; the BLA is not involved in forming taste aversions to familiar reward.^{6,7} Rather, the deficit seems to stem from an inability of the BLA-lesioned rats to experience the sensory hallucination of the reward in its predictive context. It is likely that the BLA is producing this sensory hallucination through connections with gustatory cortex^{8,9} and possibly gustatory thalamus.¹⁰ Unfortunately, we know of no other studies that have examined the neural mechanisms underlying mediated devaluation. Clearly much work remains to be done in this area.

By contrast, many studies have been performed on reward expectancy and its underlying neural mechanisms. It has been demonstrated many times that the BLA and orbitofrontal cortex (OFC) are critical to acquiring and using reward expectancies to guide behavior. Neurotoxic lesions of the BLA or OFC prevent changes in conditioned responding following reward devaluation in rats^{6,11–13} and monkeys,^{14–16} as does disconnection of these structures, via contralateral or crossed lesions.¹⁷ When reward devaluation procedures are employed in humans, functional magnetic resonance imaging reveals that cue-evoked blood oxygen level-dependent response in BLA and OFC changes to reflect the current value of predicted reward.¹⁸ Thus, across rats, monkeys, and humans reward devaluation studies reveal the BLA and OFC to be critical for signaling reward expectancies.

Relevance of Pavlovian Models to Schizophrenia

Early in training, it is normal for a cue to elicit a highly realistic, sensory representation of the reward it predicts. We have suggested that this process is similar in some ways to a hallucination: perceiving the presence of a reward that is not really there. Later in training, the rat's behavior reflects the reality that the predicted reward is not actually present but rather that it is expected in the near future. Sensory hallucinations are replaced with reward expectancies.

Although the underlying data are certainly incomplete, the pattern of results described above suggests a model in which prefrontal areas, such as the OFC, are essential for modifying sensory representations acquired by other brain areas. This would be consistent with the general proposal that prefrontal areas are essential for working memory, executive function, abstract representation, and manipulation of information that is entailed in these processes. Of course, prefrontal and even orbitofrontal dysfunction has been well documented in patients with schizophrenia.^{19–24} If our proposal is correct, then dysfunction of these prefrontal areas would disrupt the normal transition from sensory hallucinations to reward expectancies, thereby resulting in impaired reality testing (figure 1).

Importantly, the simple Pavlovian procedures described here provide a way to test this hypothesis. Of particular interest would be the performance of schizophrenic patients in reward devaluation and mediated devaluation tasks. Schizophrenic patients should fail to show reduced behavioral responding to cues predicting devalued rewards. Further, we would expect to find persistent mediated devaluation even after extended training in schizophrenic patients and also in animal models, such as neonatal ventral hippocampus lesion-lesioned rats.²⁵ This pattern of results would mirror the persistent hallucinations that characterize schizophrenia and provide a potential animal model in which to assess the efficacy of treatment.

References

- Holland PC. Acquisition of representation-mediated conditioned food aversions. *Learn Motiv.* 1981;12:1–18.
- Holland PC. Amount of training effects in representation-mediated food aversion learning: no evidence of a role for associability changes. *Learn Behav.* 2005;33:464–478.
- Holland PC, Rescorla RA. Effect of 2 ways of devaluing unconditioned stimulus after 1st-order and 2nd-order appetitive conditioning. *J Exp Psychol Anim Behav Process.* 1975;104:355–363.
- Holland PC, Straub JJ. Differential effects of two ways of devaluing the unconditioned stimulus after Pavlovian appetitive conditioning. *J Exp Psychol Anim Behav Process.* 1979;5:65–78.
- Dwyer DM, Killcross S. Lesions of the basolateral amygdala disrupt conditioning based on the retrieved representations of motivationally significant events. *J Neurosci.* 2006;26:8305–8309.
- Hatfield T, Han JS, Conley M, Gallagher M, Holland P. Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *J Neurosci.* 1996;16:5256–5265.
- Schafe GE, Thiele TE, Bernstein IL. Conditioning method dramatically alters the role of amygdala in taste aversion learning. *Learn Mem.* 1998;5:481–492.
- Krettek JE, Price JL. Projections from the amygdaloid complex to the cerebral cortex and thalamus in the rat and cat. *J Comp Neurol.* 1977;172:687–722.
- Allen GV, Saper CB, Hurley KM, Cechetto DF. Organization of visceral and limbic connections in the insular cortex of the rat. *J Comp Neurol.* 1991;311:1–16.
- Nakashima M, Uemura M, Yasui K, Ozaki HS, Tabata S, Taen A. An anterograde and retrograde tract-tracing study on the projections from the thalamic gustatory area in the rat: distribution of neurons projecting to the insular cortex and amygdaloid complex. *Neurosci Res.* 2000;36:297–309.
- Johnson AW, Gallagher M, Holland PC. The basolateral amygdala is critical to the expression of Pavlovian and instrumental outcome-specific reinforcer devaluation effects. *J Neurosci.* 2009;29:696–704.
- Pickens CL, Sadoris MP, Setlow B, Gallagher M, Holland PC, Schoenbaum G. Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *J Neurosci.* 2003;23:11078–11084.
- Gallagher M, McMahan RW, Schoenbaum G. Orbitofrontal cortex and representation of incentive value in associative learning. *J Neurosci.* 1999;19:6610–6614.
- Izquierdo A, Suda RK, Murray EA. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J Neurosci.* 2004;24:7540–7548.
- Izquierdo A, Murray EA. Combined unilateral lesions of the amygdala and orbital prefrontal cortex impair affective processing in rhesus monkeys. *J Neurophysiol.* 2004;91:2023–2039.
- Malkova L, Gaffan D, Murray EA. Excitotoxic lesions of the amygdala fail to produce impairment in visual learning for auditory secondary reinforcement but interfere with reinforcer devaluation effects in rhesus monkeys. *J Neurosci.* 1997;17:6011–6020.
- Baxter MG, Parker A, Lindner CC, Izquierdo AD, Murray EA. Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J Neurosci.* 2000;20:4311–4319.
- Gottfried JA, O'Doherty J, Dolan RJ. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science.* 2003;301:1104–1107.
- Nakamura M, Nestor PG, Levitt JJ, et al. Orbitofrontal volume deficit in schizophrenia and thought disorder. *Brain.* 2008;131(pt 1):180–195.
- Gur RE, Cowell PE, Latshaw A, et al. Reduced dorsal and orbital prefrontal gray matter volumes in schizophrenia. *Arch Gen Psychiatry.* 2000;57:761–768.
- Hoptman MJ, D'Angelo D, Catalano D, Mauro CJ, Shehzad ZE, Kelly AM. Amygdalofrontal functional connectivity and aggression in schizophrenia. *Schizophr Bull.* March 30, 2009; doi:10.1093/schbul/sbp012.
- Radulescu AR, Mujica-Parodi LR. A systems approach to prefrontal-limbic dysregulation in schizophrenia. *Neuropsychobiology.* 2008;57:206–216.

23. Barch DM, Carter CS, Braver TS, et al. Selective deficits in prefrontal cortex function in medication-naive patients with schizophrenia. *Arch Gen Psychiatry*. 2001;58:280–288.
24. O'Donnell P. Increased cortical excitability as a critical element in schizophrenia pathophysiology. In: O'Donnell P, ed. *Cortical Deficits In Schizophrenia: From Genes to Function*. Springer US; 2008:219–236.
25. Tseng K, Chambers RA, Lipska BK. The neonatal ventral hippocampal lesion as a heuristic neurodevelopmental model of schizophrenia. *Behav Brain Res*. In press.