Holmes and Clemens suggest two alternative explanations to account for data presented in our recent paper. While a single study cannot rule out all possible explanations, the existing literature does not support their claims.

The authors propose that *"with extended training, persistence of the drug-seeking response after extinction of the drug-taking response may have been simply due to a direct association between the drug-seeking response and cocaine"*. They base their argument on work by Corbit and Balleine (2003) but fail to note four key points. First, Corbit and Balleine used a motivational shift procedure in which the food used for satiation differs from that used in the instrumental contingency. In devaluation procedures, outcome devaluation is only achieved when the food used to induce satiety is the same as that associated with the instrumental response. Prefeeding an alternative reinforcer is without effect. Second, even if one accepts the false premise that the motivational shift paradigm is analogous to standard devaluation procedures, satiation failed to affect the seeking response in the shift paradigm and only affected the seeking link when animals had opportunity to experience the food reinforcer in the satiated, non-food-deprived state (Balleine et al., 1995). Thus, the ability of the primary reinforcer to affect the seeking link was dependent on incentive learning. In our procedure, there is no shift in motivational estate or opportunity for incentive learning. Therefore, the devaluation effect we observed cannot simply be explained by a shift in motivational state. Third, analyses of chained schedules suggest that conditioned reinforcement provided by access to the terminal link maintains responding occurring during the early link. Although some research has suggested that responding during early links can be controlled by the primary reinforcer at the end of the chain across a temporal delay (Staddon, 1983), other work does not support this interpretation (Williams et al., 1995). Finally, the explanation of Holmes and Clemens requires not only that

drug seeking is controlled by association with the primary reinforcer but that this control gains relevance with extended training. To our knowledge, no such evidence exists in the literature. In its absence, the explanation remains conjecture.

The authors suggest that *"chunking of drug-seeking and drug-taking responses may have occurred."* In support of their assertion they refer to Ostlund et al. (2009) incorrectly stating that *"the same heterogeneous chain as Zapata et al. was used".* Our study employed a forced-choice, chained schedule. In contrast, Ostlund et al. used a free-choice two-lever sequential procedure, in which both levers were present during training and testing. Additionally, the sequences and reinforcer contingencies were concurrently active. Lever pressing in one sequence delivered a reinforcer; the reverse sequence resulted in delivery of an alternative reinforcer. In contrast to our study, each lever response was the proximal response for one reinforcer and the distal response for the alternative reinforcer. Thus, in this procedure, learning of specific sequences as single units is fostered, because differentiating the sequences determines obtained outcome and animals earn only a fixed number of the two reinforcers. Importantly, the sequence was no longer reinforced after obtaining the first reinforcer and alternative reinforcer presentation required completion of the alternative sequence. In our study, levers were not presented simultaneously. Upon completion of the seeking contingency, the seeking lever was withdrawn and the taking lever inserted. After completion of the taking contingency, the taking lever was withdrawn and a timeout, with no levers, enforced. Thus, in striking contrast to the paper cited, there is no opportunity to learn about sequence, schedule determines lever order and sequence is irrelevant. It is highly unlikely that chunking occurred under these conditions. Furthermore, Ostlund et al. (2009) showed that a sequence can become a single action unit associated with outcome and sensitive to outcome devaluation using a schedule which promotes goal-directed responding. Yet,

according to Holmes and Clemens, chunking occurs after extended training, when responding is habitual, and not after moderate training, when responding is still goal-directed. Thus, the study cited argues against the explanation proffered.

Lastly, Holmes and Clemens note that the chunking process depends upon the agranular cortex (Ostlund et al., 2009) and argue that, since this region innervates the dorsal striatum, the effects of dorsolateral striatum (DSL) inactivation may be explained by blocking of the chunking process, rather than habitual responding. Holmes and Clemens dismiss an extensive literature demonstrating the role of the DSL in habitual cocaine seeking and fail to note that the agranular cortex projects to different subregions of the dorsal striatum, each of which differentially modulates instrumental responding (Yin and Knowlton, 2006). We question ascribing the effects of agranular cortex inactivation to one of many projection sites in the absence of functional evidence. Finally, in the Ostlund study, agranular cortex lesions result in insensitivity of the sequence to devaluation indicating that this region is required for the sequence unit to function as a goal-directed action. Thus, given the established role of the dorsomedial striatum (DMS) in goal-directed behavior, this observation, in contrast to the assertion of Holmes and Clemens, suggests a functional connection of the agranular cortex with the DMS, rather than the DSL.

In summary, we used devaluation of the terminal link of a chained schedule to test the associative structure controlling responding on the distal link of a cocaine seeking/taking procedure. This devaluation procedure has been used previously in studies of both natural (Williams et al., 1995) and drug (Olmstead et al., 2001) reinforcers. The interpretations offered by Holmes and Clemens not only discount previous literature in the learning and memory field, but are based upon a seemingly incomplete understanding of the articles cited and a failure to recognize key procedural differences between those studies and ours. We hope that by clarifying

these inconsistencies the authors will recognize that although alternative explanations are indeed possible, the conclusions of our studies are the most parsimonious.

References

1. Balleine BW, Garner C, Gonzalez F, Dickinson A (1995) Motivational control of heterogeneous instrumental chains. Journal of Experimental Psychology 21: 203-217.
2. Corbit LH, Balleine BW (2003) Instrumental and Pavlovian incentive processes have dissociable effects on components of a heterogeneous instrumental chain. J Exp Psychol Anim Behav Process 29: 99-106.
3. Olmstead MC, Lafond MV, Everitt BJ, Dickinson A (2001) Cocaine seeking by rats is a goal-directed action. Behav Neurosci 115: 394-402.
4. Ostlund SB, Winterbauer NE, Balleine BW (2009) Evidence of action sequence chunking in goal-directed instrumental conditioning and its dependence on the dorsomedial prefrontal cortex. J Neurosci 29: 8280-8287.
5. Staddon JER (1983) Adaptive Learning and Behavior. Cambridge University Press.
6. Williams BA, Ploog BO, Bell MC (1995) Stimulus devaluation and extinction of chain schedule performance. Animal Learning & Behavior 23: 104-114.
7. Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. Nat Rev Neurosci 7: 464-476.