Neuringer, A., Jensen, G., & Piff, P. (2007). Stochastic matching and the voluntary nature of choice. *Journal of the Experimental Analysis of Behavior, 88*, 1-28.

ABSTRACT

Attempts to characterize voluntary behavior have been ongoing for thousands of years. We provide experimental evidence that judgments of volition are based upon distributions of responses in relation to obtained rewards. Participants watched as responses, said to be made by ''actors,'' appeared on a computer screen. The participant's task was to estimate how well each actor represented the voluntary choices emitted by a real person. In actuality, all actors' responses were generated by algorithms based on Baum's (1979) generalized matching function. We systematically varied the exponent values (sensitivity parameter) of these algorithms: some actors matched response proportions to received reinforcer proportions, others overmatched (predominantly chose the highest-valued alternative), and yet others undermatched (chose relatively equally among the alternatives). In each of five experiments, we found that the matching actor's responses were judged most closely to approximate voluntary choice. We found also that judgments of high volition depended upon stochastic (or probabilistic) generation. Thus, stochastic responses that match reinforcer proportions best represent voluntary human choice.

Is it human? Judgments of choice responding

PsyCrit, October 11, 2007

The target paper concluded that the degree to which people judged that an actor's (in reality, a computer's) performance on concurrent variable-interval (VI) schedules to be like voluntary human behavior or not depended on the degree to which the actor's performance approximated strict matching (a sensitivity of 1 in the generalized matching relation; Baum, 1974). The results are interesting, though some further considerations and further analysis would have been useful. The authors' interpretation of their results is questionable.

In Experiment 1, the actors responded according to the generalized matching relation (Baum, 1974):

$$\log \frac{B_1}{B_2} = a \log \frac{R_1}{R_2} + \log c \,,$$

where *B* refers to responses, *R* to reinforcers, and the subscripts denote the choices. The parameter *a* is called "sensitivity to reinforcement", and it measures the amount that log choice changes with a unit change in the log reinforcer ratio. Each of the 6 actors responded with a different value of sensitivity to reinforcement that ranged from 0 to 6—from no effect of changes in reinforcer ratio on choice to an extreme effect, respectively. There were 6 games that differed in the probability of reinforcement at each of 3 response alternatives. Judgement of human volition (but see below) was found to be maximal when the actor responded with a sensitivity of 1.0, and decreased with increasing deviations away from this sensitivity value.

I have two general problems with this research. The first concerns what was being measured, the second concerns the adequacy of the actor behavior as representative of human behavior.

What is being measured?

The instructions for the Experiments 1 and 2 were: "Your task is to rate the six actors in terms of how well each represents a *human* making *voluntary* choices in the same game that you've just played… Use a 0-100 point scale, with 0 indicating that the actor is absolutely *not representative of a real human being*, and 100 indicating maximum approximation to *real human behavior*. Of course, intermediate values will indicate intermediate levels of approximation to *voluntary hu-*

man choices…" (Neuringer et al., 2007, p. 6, italics added). Experiment 3 instructions were to rate "computer algorithm" or ''voluntary human." (Neuringer et al., p. 12). Experiment 4 instructions were: "(You should) attempt to decide whether the choices represented across the six games represent voluntary choices made by a real human player or not." (p. 13), and Experiment 5 "Not At All Human" … and "Certainly Human" (Experiment 5, p. 14). These instructions, at least for Experiments 1, 2 and 4, stress "human" [H] and "voluntary" [V] separately in Sentence 2, but Sentences 1 and 3 conflate "human" and "voluntary". So, what was the positive conditional stimulus (S+) that the participants were requested to report? – was it H with V irrelevant, or V with H irrelevant, or was it the conjunction of H and V together? One might presume that the last of these was intended, were it not that the paper and its title focus on "voluntary". Were the reports of each of the participants controlled by the same S+ or by different S+s? The same questions may be asked of the S- stimulus – was this Not H with V irrelevant, or Not V with H irrelevant, or Not (H and V). Given these various interpretations, was each participant's behavior controlled by the same S-? The extreme inter-participant variation in the patterns of responding over all 5 experiments (Appendix 2) suggests imperfect dimensional stimulus control.

It is further unclear what reporting behavior was controlled by the anchors at the other end of these possible dimensions. The category "not human" could, I suppose, include non-human animal or mechanical, but it is hard to know how the demand characteristics of the situation controlled the behavior of each participant. It is even harder to determine the implication of the "not voluntary" end of the dimension – is this "involuntary", or maybe "constrained", in as much as a constraint on behavior might imply it was not freely emitted?

In summary, neither the S+ or the S- conditional stimuli controlling the behavior of the participants was clearly defined in the instructions. Moreover, no post-hoc attempts were made to help determine the controlling dimension (e.g., perhaps by asking the participants what controlled their choices). Admittedly, it is nearly impossible to carry out adequate psychophysics on subjective stimuli such as pain or diffidence, and it is equally difficult for the verbal community, or an experimenter, to train adequate reporting of the presence or absence of such stimuli. As always, the criterion is: How could we train a non-verbal animal to report (say) H and V versus Not H and V? To do this, we would have to define these categories before we commenced training, completely subverting any interpretation of the results in terms of what is voluntary human behavior. Is it better to leave the definition of the categories to an individual's history with her/his verbal community? If we do so, we are discovering more about the history of shaping by the verbal community than the behavior of the participant; but this, of course, convolves upon itself and becomes a consensual hallucination (Gibson, 1984).

Because I am unsure of the conditional stimuli, I will try to be generic about what I say henceforth.

The adequacy of the actor

The actor was designed to emit 3-alternative matching behavior with various generalized-matching sensitivities to reinforcement. Neuringer et al. (2007) *assume* that the experimental variation of sensitivity must be mapped in a one-to-one fashion to one or both of the human or voluntary dimensions, and that their results thus speak to control by one of these dimensions (the voluntary dimension is highlighted in the MS). It is not possible to say whether this is a valid assumption or not.

If we assume the assumption is valid, then a number of more technical points should be made about these results. They are: (1), humans rather seldom match with a sensitivity of 1.0 on concurrent VI VI schedules. Kollins, Newland, and Critchfield (1997), in their review of human matching, reported sensitivities that ranged from negative to positive greater than 1. Thus, it could be suggested that the ways in which participant reports changed with sensitivity in the target article were not accurately controlled by the "human" dimension, because, if so, the judgements *should* have been non-differential with respect to sensitivity; (2), both punishment for concurrent-schedule performance (Farley, 1980), and the imposition of travel time between alternatives (Baum, 1982; Davison, 1991), produce sensitivity values that reliably exceed 1 – if participants were judging "voluntary" rather then "human", and if "constrained" is the opposite of "voluntary", then the participants' judgements of overmatching as non-voluntary (= constrained) would be appropriate for performance under such constrained conditions—but a judgement of "non-human" would be inappropriate; (3), if the judgements are of "voluntary" rather than "human", the long tails of distributions of non-human animal sensitivity values (Baum, 1979; Taylor & Davison, 1983) would be judged as non-voluntary and constrained. This might be taken as appropriate, were it not for non-human sensitivity distributions having shorter tails (hence, more voluntary) than human sensitivity distributions (less voluntary; Kollins et al.); (4), but the nonhuman sensitivity distributions come from concurrent schedules using changeover delays (Herrnstein, 1961), which prevent reinforcers being obtained within a fixed period of time following changing to an alternative. As Herrnstein (1961) and Shull and Pliskoff (1967) showed, responding on concurrent schedules without changeover delays generally has sensitivity values considerably less than 1.0, and would be judged *non*-voluntary or constrained by Neuringer et al.'s (2007) participants; the introduction of a changeover delay brings sensitivity values closer to 1.0, which Neuringer et al.'s participants would judge as more voluntary. Thus, perversely, *adding a constraint* would increase judgements of volition; (4), Taylor and Davison showed that the means of sensitivity distributions are affected by how VI schedules are constructed, with arithmetic progressions giving lower mean sensitivity values than exponential progressions. If this were true for humans too, judgements of performances as representing voluntary human behavior will depend on the schedule progression used; (5), in non-human animals, sensitivity to reinforcement increases with overall reinforcer rate (Alsop & Elliffe, 1988; Elliffe & Alsop, 1996)—again, if this also occurs with humans, participants' judgements of the degree to which a performance is voluntary will change with overall reinforcer rate. In summary, while participants' performances may not be under the control of the "human" dimension, we should be very wary about accepting the implication that deviations from strict matching thus measure constraints to human volition, or that closeness to strict matching implies volition. If adding or increasing a constraint can ever increase judgements of human volition (as in the changeover-delay case) such verbal reports seem unlikely to be related to volition—unless the opposite of volition if not constraint.

Many of the points raised above would be interesting to test in Neuringer et al.'s (2007) procedure. In many cases above, it could be that an actor sensitivity that deviated substantially from 1 *would* engender a report of a high approximation to real voluntary human behavior. For instance, if an actor travel time is introduced and made obvious to the participant, an actor with a high sensitivity might be judged human-like, while one with unit sensitivity would be not be judged human-like. Similar points can be made for punishment, for schedule progressions, and for overall reinforcer rates. It is possible, therefore, that sensitivity to reinforcement, or behavior approximating a sensitivity of 1.0, may not generally be judged as voluntary human behavior, and that the control of participants' verbal behavior is much more subtle. Neuringer et al.'s results would be considerably strengthened if participant judgements depended only on sensitivity and were independent of procedure—in which case, the addition of punishment or travel time would seem, quite appropriately, to be making behavior less voluntary. But if participant judgements depended on both sensitivity and experimental procedure, some participant judgements would be based on details of human-like behavior generally, rather than on voluntary human behavior.

A related caveat is that the computer program that was the actor was surely not an adequate and detailed model of human matching—at our current state of knowledge, this is simply not available. While we know nothing about the *details* of human matching to three alternatives as used by Neuringer et al., animal data as mentioned above surely have some bearing. In particular, animals show "preference pulses" after reinforcers and effects of sequences of up to 3 previous reinforcers (Davison & Baum, 2003). It appears that the stochastic matcher (Appendix 2 of Neuringer et al., 2007) did not have such properties. We also have as-yet unpublished data showing that pigeons in 3- and 4-alternative choice make transitions from each alternative to the other alternatives according to a power function of the reinforcer ratio on the other available alternatives. It is not evident that the stochastic matcher had these properties either.

The general point is the same general point that makes the use of simulation to test theory difficult: Unless you have a simulator (actor, here) that behaves exactly (so far as we know at any one time) like a real animal or human, the results of the simulation are unconvincing. If a participant judges such an impaired actor (in the sense of one that does not show all the detailed local behavioral features of a real human), but one that does show an overall sensitivity of 1.0, as representing human voluntary performance, what can we conclude? Perhaps that the details of performance are irrelevant to such a judgement, and only the extended performance is the controlling stimulus; but there is a myriad of other possible interpretations that remain untested.

A second technical point is that the matching relation is but one of a number of descriptions of performance on concurrent VI VI schedules. Which extended matching model one espouses is, to some extent at least, a matter of choice. Personally, I like the contingency-discriminability model (Davison & Jenkins, 1985; Davison & Nevin, 1999), mostly because the discriminability parameter in this model, d_r, directly implies the operations that change its value. The model is:

$$\frac{B_1}{B_2} = c\frac{d_r R_1 + R_2}{d_r R_2 + R_1},$$

in which *B* refers to responses and *R* to obtained reinforcers. The contingency-discriminability parameter is constrained to fall between 1 (no discriminability) and infinity (perfect discriminability); *c* is inherent bias.

This model, as Davison and Jenkins showed, makes predictions that are not discriminably different from the generalized-matching relation under two conditions: (1), when d_r is relatively high, say > 10); and (2), when the reinforcer ratio is within the range 1:10 through 10:1. If the first constraint is not met, sensitivity will be less than 1, falling to zero when $d_r = 1$. If the second constraint is not met, and d_r is not too high, predictions of the contingency-discriminability model will become less extreme than the generalized-matching relation as reinforcer ratios become more extreme. Jones and Davison (1995) showed that this deviation does occur when d_r was arranged to be moderately low. Additionally, this model predicts that response ratios are *not* infinite on concurrent VI Extinction schedules, as indeed empirically they are not (Davison & Hunter, 1976; Davison & Jones, 1998).

Neuringer et al. (2007) arranged actor reinforcer ratios that were 6:1 at their most extreme when reinforcers were delivered, but in 3 of the 6 games (Games 3, 4, and 6), they arranged no reinforcers on only 1 or 2 of the 3 alternatives (producing infinite reinforcer ratios). An actor responding according to generalized matching with *any* sensitivity value would not have responded on the Extinction alternatives in Games 3, 4, and 6. Of the remaining 3 games, Game 1, which arranged equal reinforcer rates on all alternatives, provided no differential signal to the participant with changes in matching sensitivity because any generalized-matching sensitivity (or contingency-discriminability) value would produce the same response distribution.

The important questions are: Did the participants judge all 6 games equally across actors with differing sensitivity? Game 1 provided no differential across-actor sensitivity; Games 2 and 5 provided differential response allocations across actors, but the reinforcer ratios were varied only over a very limited range; and Games 3, 4, and 6 provided response distributions consistent with generalized matching with various sensitivities, but also some response distributions that could have been incompatible (if d_r was less than infinite) with contingency-discriminability predictions. From the generalized matching perspective, Game 2 provided the greatest discriminative signal with reinforcer ratios of 2:1, 3:1 and 6:1, and, presuming contingency discriminability is high, Game 2 provided little signal for the contingency-discriminability model. The important question that needs answering is: Which game did the participants choose to view in the final phase of Experiment 1? These data were not reported, but the observing-response part of the procedure could provide highly salient data on the features of the assumed stimuli that were controlling the verbal behavior of the participants.

Another way of interpreting Neuringer et al.'s (2007) data is to assume that peoplerespond according to the contingency-discriminability model, and that the participants were reporting deviations from unconstrained discriminability ($d_r = \infty$, and the absence of punishment and similar constraints) in this model. Under this possibility, zero responding on the Extinction alternative(s) in Games 3, 4 and 6 would look like a $d_r = \infty$ (aka a sensitivity of 1.0) for all actor sensitivity values—but the other pairwise choices in Games 3 and 6 would look like a $d_r < \infty$, making the overall performance look unnatural. Games 2 and 5 would not look $d_r = \infty$ except when actor sensitivity was 1.0. Thus, it would be of major interest to look at the final-phase data according to Game and actor sensitivity, if such were available. This might give some useful pointers in the choice between the generalized matching and the contingency discriminability models. The choice between the two actor models discussed here, however, requires the use of a contingency-discriminability actor, and a considerably wider range of obtained reinforcer ratios.

Neuringer et al. (2007, p. 18) state: "We propose … that the term "voluntary" refers to a demonstrable behavioral phenomenon and that identifying functional relationships between instances and judgments will help us to understand that phenomenon. Stated differently, ''voluntary'' is a discriminable behavioral characteristic (rather than an abstract metaphysical attribute), one that can best be identified in terms of behavior-environment relationships. The reality of volition is supported by the facts just described, namely that animals and people match responses to reinforcements in a stochastic fashion, and, as indicated by the present research, so doing is judged by an external observer to involve voluntary choice." Their interpretation hinges on the functional relation between instances and judgements, but it is arguable whether they achieved this, given that they have clearly controlled neither the presence nor the absence of instances of human volition, and their instructions muddied the ends of the presence/absence dimension. In the paragraph prior to that reproduced above, they contrast their position on the "reality" of volition with that of volition being a shared consensual agreement about a non-existing entity (which is to say that the verbal behavior about volition is controlled by the verbal community). But the agreement across participants concerning non-existing entities (Santa Claus and unicorns, at least in some cultures) would show much more agreement across subjects than the presence/absence of human volition, simply because of better control by the verbal community. One simply cannot argue for the "reality" of volition when its presence or absence is much less well discriminated than are the presence and absence of concepts that themselves are surely not real.

Michael Davison
The University of Auckland, New Zealand
m.davison@auckland.ac.nz

REFERENCES

lsop, B., & Elliffe, D. (1988). Concurrent-schedule performance: Effects of relative and overall reinforcer rate. *Journal of the Experimental Analysis of Behavior, 49*, 21-36.

Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 22*, 231-242.

Baum, W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior,* 32, 269-281.

Baum, W. M. (1982). Choice, changeover, and travel. *Journal of the Experimental Analysis of Behavior, 38*, 35-49.

Davison, M. (1991). Choice, changeover, and travel: A quantitative model. *Journal of the Experimental Analysis of Behavior, 55*, 47-61.

Davison, M., & Baum, W. M. (2003). Every reinforcer counts: Reinforcer magnitude and local preference. *Journal of the Experimental Analysis of Behavior, 80*, 95-129.

Davison, M. C., & Hunter, I. W. (1976). Performance on variable-interval schedules arranged single and concurrently. *Journal of the Experimental Analysis of Behavior, 25,* 335-345.

Davison, M., & Jenkins, P. E. (1985). Stimulus discriminability, contingency discriminability, and schedule performance. *Animal Learning & Behavior, 77*, 77-84.

Davison, M., & Jones, B. M. (1998). Performance on concurrent variable-interval extinction schedules. *Journal of the Experimental Analysis of Behavior, 69*, 49-57.

Davison, M., & Nevin, J. A. (1999). Stimuli, reinforcers, and behavior: An integration. *Journal of the Experimental Analysis of Behavior, 71*, 439–482.

Elliffe, D., & Alsop, B. (1996). Concurrent choice: Effects of overall reinforcer rate and the temporal distribution of reinforcers. *Journal of the Experimental Analysis of Behavior, 65*, 445-463.

Farley, J. (1980). Reinforcement and punishment effects in concurrent schedules: A test of two models. *Journal of the Experimental Analysis of Behavior, 33*, 311-326.

Gibson, W. (1984). *Neuromancer*. London, UK: Gollancz.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4*, 267-272.

Jones, B. M. & Davison, M. (1995). A quantitative analysis of extreme choice. *Journal of the Experimental Analysis of Behavior, 64,* 147-162.

Kollins, S. H., Newland, C. M., & Critchfield, T. S. (1997). Human sensitivity to reinforcement in operant choice: How much do consequences matter? *Psychonomic Bulletin & Review, 4*, 208-220.

Neuringer, A., Jensen, G., & Piff, P. (2007). Stochastic matching and the voluntary nature of choice. *Journal of the Experimental Analysis of Behavior, 88*, 1-28.

Shull, R. L., & Pliskoff, S. S. (1967). Changeover delay and concurrent schedules: Some effects on relative performance measures. *Journal of the Experimental Analysis of Behavior, 10*, 517-527.

Taylor, R., & Davison, M. (1983). Sensitivity to reinforcement in concurrent arithmetic and exponential schedules. *Journal of the Experimental Analysis of Behavior, 39*, 191-198.

Acknowledgements

I thank Douglas Elliffe for his very helpful comments on this manuscript, and also (alphabetically) Annelle, Kristie, Siope, Susan, and Lydia who discussed this paper in Psych 711, Advanced Operant Behaviour.