Certainties and mysteries in the within-trial contrast literature: A reply to Zentall (2008)

MARCO VASCONCELOS AND PETER J. URCUIOLI
Purdue University, West Lafayette, Indiana

Contrary to Zentall's (2008) hypothesis that extensive overtraining is required in order to obtain the work ethic effect in pigeons, we find no evidence of this in two separate paradigms designed to assess preference for stimuli encountered after high effort. In light of these findings, along with our previously published results and those of Arantes and Grace (2008), we maintain that the work ethic effect is not a reliable phenomenon in pigeons. On the other hand, we are uncertain about the reason(s) underlying the different preference profiles for high- versus low-deprivation stimuli as a function of successive versus simultaneous discrimination training (Vasconcelos & Urcuioli, 2008). Nevertheless, we disagree with Zentall that additional analyses of the sort he suggests support the conclusion that both training regimens yield within-trial contrast effects.

Zentall (2008) questions our conclusions regarding the reliability of the work ethic effect in pigeons (Vasconcelos, Urcuioli, & Lionello-DeNolf, 2007a, 2007b) as well as the absence of clear preferences for stimuli previously encountered under high deprivation over those previously encountered under low deprivation when each set of stimuli appears in the context of a simultaneous discrimination (Vasconcelos & Urcuioli, 2008). In the former case, he appeals to deficiencies in the amount of training prior to preference testing; in the latter case, he cites other statistical analyses that seem to support the generality of within-trial contrast across different training regimens.

Regarding our work ethic results (Vasconcelos et al., 2007a), Zentall mentions that our failure to replicate the work ethic effect is easily explained by insufficient overtraining. According to Zentall (see also Zentall & Singer, 2007b), reliable preferences for high-effort stimuli may require upward of twice the amount of overtraining that we gave our pigeons (viz., 20 sessions). However, given the findings of Arantes and Grace (2008) and other recently collected data mentioned in Vasconcelos and Urcuioli (2008, p. 17), our original failures to replicate are difficult to dismiss on this account.

For instance, we (Vasconcelos & Urcuioli, 2007) gave 6 experimentally naive pigeons 60 overtraining sessions after they had learned two S^+/S^- simultaneous discriminations, one of which was obtained after completing a fixed-ratio 1 (FR1) requirement to a vertical-line stimulus and the other after completing an FR30 requirement to a horizontal-line stimulus. (Signaling the different requirements should be especially conducive to within-trial con-

trast, according to Zentall & Singer, 2007a.) One pigeon could not be tested because it repeatedly stopped for 3 h or more on signaled FR30 trials, a testament to the aver-siveness of this requirement. For the other pigeons, five test sessions were run following their extensive overtraining in which nondifferentially reinforced trials involving a choice between the two S^+ stimuli (S^+_{FR1} vs. S^+_{FR30}) or between the two S^- stimuli (S^-_{FR1} vs. S^-_{FR30}) were interspersed among baseline training trials. Figure 1 (top panel) shows the average preferences for the FR30 S^+ and S^- on these trials, both for the first test session and over all five test sessions. None of the average preferences deviated significantly from chance [all $t_s(4) < 1.60$, two-tailed].

Likewise, we (Vasconcelos & Urcuioli, 2007) ran a pigeon version of Kacelnik and Marsh's (2002) effort-based experiment with starlings. In their experiment, starlings flew back and forth between two perches located 1 m apart and obtained food by pecking one hue after sixteen 1-m flights (high-effort trials) and by pecking another hue after four 1-m flights (low-effort trials). In our experiment, we used a chamber that had two intelligence panels 93 cm apart (a "long box"). Four experienced pigeons had to walk from one panel to the other either four times or sixteen times in order to obtain a discriminative stimulus that, when pecked, produced food. Different S^+ stimuli followed the four 93-cm runs and the sixteen 93-cm runs. Pigeons were probed for their S^+ preferences after 25, 50, and 60 training sessions. Afterward, we reversed the original contingencies and again probed pigeons' preferences after 25, 50, 60, and 70 training sessions. Figure 1 shows the average preference for the 16-runs (high-effort) stimulus before and after reversal (bottom left and right panels, respectively). Again, none of the preferences deviated significantly from chance [$-1.57 < \text{all } t_s(3) < 1.00$, two-tailed].

Our long-box findings coincide with the other failure-to-replicate results obtained in standard operant chambers (Arantes & Grace, 2008; Vasconcelos & Urcuioli, 2007; Vasconcelos et al., 2007a). In addition, they suggest that Clement, Feltus, Kaiser, and Zentall's (2000) failure to obtain a work ethic effect after successive discrimination training may have less to do with impulsivity on test trials (Zentall, 2008, p. 21) and more to do with the unreliability of this effect in pigeons. After all, like Kacelnik and Marsh's (2002) starlings, the pigeons that walked back and forth multiple times between panels were relatively far away from the S^+ when it appeared on each trial.

Despite the ineffectiveness of our effort-based manipulations, we (Vasconcelos & Urcuioli, 2008) did obtain results consistent with within-trial contrast using a deprivation manipulation. However, pigeons' preference for an S^+ experienced under high deprivation over an S^+ experienced under low deprivation (i.e., following prefeeding) occurred only after successive discrimination training

Figure 1. Results from unpublished data by Vasconcelos and Urcuioli (2007). Top panel: Average percentage choices ($\pm$ SEM) of the FR30 S⁺ stimulus (filled bars) and the FR30 S⁻ stimulus (unfilled bars), on the first test session and over all five test sessions, after 60 overtraining sessions. Bottom panels: Average percentage choices ($\pm$ SEM) of the 16-runs stimulus from training across different levels of overtraining, both before (left panel) and after (right panel) reversal.

(Experiment 1), not after simultaneous discrimination training (Experiment 2).

Zentall argues that other analyses lend support for claiming within-trial contrast in both experiments. For instance, he claims that when the preference results are averaged over both test sessions in our Experiment 2, 5 of the 6 pigeons (or 83%) showed a within-trial contrast effect. But this seems to be overstating the case. The high-deprivation S⁺ preferences for the 6 pigeons averaged over the two test sessions were 57.5%, 25%, 82.5%, 67.5%, 52.5%, and 62.5%. Significance ($p < .05$) in a one-tailed binomial test with 40 trials requires 65% accuracy, so by this statistical decision-making rule, only 2 of the 6 pigeons (33.3%) showed a within-trial contrast effect. Moreover, 1 of the remaining 4 pigeons showed the opposite effect! Even using the number of birds with any preference score above 50% (cf. Kacelnik & Marsh, 2002) as the dependent variable (which we regard as inappropriate), a result based on 5 out of 6 subjects is not significant according to a binomial test ($p > .05$).

Zentall also argues that the results from the two experiments should differ significantly in order to conclude that a preference effect was obtained in one experiment but not the other. In fact, a t test performed on S⁺ preferences averaged across both test sessions in each experiment reveals no significant between-experiment difference [$t(8) = 1.71, p = .13$, two-tailed]. The question is what to make of this result. Zentall seems to advocate that the reliable within-trial contrast effect observed in Experiment 1 is in fact evident in Experiment 2. However, another pos-

sible interpretation is that no within-trial contrast occurs in *either* experiment and that our assertion of such an effect in Experiment 1 is a Type I error (i.e., no true effect was there). Because no logical solution is available solely on the basis of this type of analysis, an analysis of individual preferences may be informative: All pigeons in Experiment 1 showed a significant preference for the high-deprivation stimulus, but in Experiment 2, only 2 of 6 did, and 1 of the remaining 4 preferred the low-deprivation stimulus. To us, this supports our conclusion that the effect was present in Experiment 1 but not in Experiment 2. The conclusion is further strengthened, in our opinion, by a significant between-experiment difference using the data from only the first test session [$t(8) = 2.40, p < .05$, two-tailed], which was less likely to be impacted by the nondifferential reinforcement for any preference choice.

Finally, Zentall (p. 21) reaffirms that within-trial contrast usually requires a considerable amount of training, typically more than the twelve 20-trial sessions used in Vasconcelos and Urcuioli (2008). We have already argued against this position on the basis of data from our effort-based manipulations (Vasconcelos & Urcuioli, 2007) and from those of Arantes and Grace (2008). The position also seems unconvincing for deprivation-based manipulations. To date, all within-trial contrast effects arising from deprivation manipulations have been obtained with relatively small amounts of training: to be precise, eight or ten 10-trial sessions (Marsh, Schuck-Paim, & Kacelnik, 2004), ten 20-trial sessions (Pompilio & Kacelnik, 2005), six 8-trial sessions (Pompilio, Kacelnik, & Behmer, 2006),

and twelve 20-trial sessions (Vasconcelos & Urcuioli, 2008, Experiment 1). To us, this pattern says that something else is likely to be responsible for the null results from our simultaneous discrimination experiment (although we don't yet know what that "something else" may be).

In summary, we feel certain that the work ethic effect in pigeons is not reliable and that amount of overtraining does not matter. Nevertheless, other manipulations (e.g., deprivation) do appear to generate within-trial contrast in pigeons. However, at this time, the conditions under which the latter manipulations induce preferences indicative of within-trial contrast remain poorly understood.

AUTHOR NOTE

Correspondence concerning this article should be addressed to M. Vasconcelos, Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907-2081 (e-mail: marcov@psych.purdue.edu).

REFERENCES

- ARANTES, J., & GRACE, R. C. (2008). Failure to obtain value enhancement by within-trial contrast in simultaneous and successive discriminations. *Learning & Behavior*, **36**, 1-11.
- CLEMENT, T. S., FELTUS, J. R., KAISER, D. H., & ZENTALL, T. R. (2000). "Work ethic" in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review*, **7**, 100-106.
- KACELNIK, A., & MARSH, B. (2002). Cost can increase preference in starlings. *Animal Behaviour*, **63**, 245-250.
- MARSH, B., SCHUCK-PAIM, C., & KACELNIK, A. (2004). Energetic state during learning affects foraging choices in starlings. *Behavioral Ecology*, **15**, 396-399.
- POMPILIO, L., & KACELNIK, A. (2005). State-dependent learning and suboptimal choice: When starlings prefer long over short delays to food. *Animal Behaviour*, **70**, 571-578.
- POMPILIO, L., KACELNIK, A., & BEHMER, S. T. (2006). State-dependent learned valuation drives choice in an invertebrate. *Science*, **311**, 1613-1615.
- VASCONCELOS, M., & URCUIOLI, P. J. (2007). *Extensive training is insufficient to produce the work ethic effect in pigeons*. Manuscript submitted for publication.
- VASCONCELOS, M., & URCUIOLI, P. J. (2008). Deprivation level and choice in pigeons: A test of within-trial contrast. *Learning & Behavior*, **36**, 12-18.
- VASCONCELOS, M., URCUIOLI, P. J., & LIONELLO-DENOLF, K. M. (2007a). Failure to replicate the "work ethic" effect in pigeons. *Journal of the Experimental Analysis of Behavior*, **87**, 383-399.
- VASCONCELOS, M., URCUIOLI, P. J., & LIONELLO-DENOLF, K. M. (2007b). When is a failure to replicate not a Type II error? *Journal of the Experimental Analysis of Behavior*, **87**, 405-407.
- ZENTALL, T. R. (2008). Within-trial contrast: When you see it and when you don't. *Learning & Behavior*, **36**, 19-22.
- ZENTALL, T. R., & SINGER, R. A. (2007a). Within-trial contrast: Pigeons prefer conditioned reinforcers that follow a relatively more rather than a less aversive event. *Journal of the Experimental Analysis of Behavior*, **88**, 131-149.
- ZENTALL, T. R., & SINGER, R. A. (2007b). Within-trial contrast: When is a failure to replicate not a Type I error? *Journal of the Experimental Analysis of Behavior*, **87**, 401-404.

(Manuscript received October 30, 2007;
accepted for publication October 30, 2007.)