Available online at www.sciencedirect.com

SCIENCE DIRECT®

Behavioural Processes 73 (2006) 36–40

BEHAVIOURAL PROCESSES

www.elsevier.com/locate/behavproc

The internet as a research tool in the study of associative learning: An example from overshadowing

Miguel A. Vadillo, Raúl Bárcena, Helena Matute *

Departamento de Psicología, Universidad de Deusto, Apartado 1, 48080 Bilbao, Spain

Received 29 July 2005; received in revised form 31 January 2006; accepted 31 January 2006

Abstract

The present study aimed to replicate an associative learning effect, overshadowing, both in the traditional laboratory conditions and over the internet. The experimental task required participants to predict an outcome based on the presence of several cues. When a cue that was always trained together with a second cue was presented on isolation at test, the expectancy of the outcome was impaired, which revealed overshadowing. This experimental task was performed by undergraduate students ($N = 106$) in the laboratory and by a different set of anonymous participants over the internet ($N = 91$). Similar levels of overshadowing were obtained in both locations. These similarities show that web-delivered experiments can be used as a complement of traditional experiments.

Keywords: Associative learning; Predictive learning; Overshadowing; Internet

The expansion of the internet will surely be recalled as one of the most important events of the late twentieth century. Although its complete impact on our lives in the long run cannot be assessed yet, it is clear that it is already a fundamental part of our daily activities, either as a working tool or as a window to countless entertainment possibilities. Researchers from different areas of psychology are also becoming aware of the potential benefits that the use of the internet can provide. Not only is the internet being used as a means to obtain larger samples in surveys and personality questionnaires (Buchanan and Smith, 1999; Gosling et al., 2004; Schmidt, 1997), but it is also starting to be used in experimental studies on basic cognitive phenomena such as probability learning, decision making or causal induction (Birnbaum, 1999; Birnbaum and Wakcher, 2002; Steyvers et al., 2003; Vadillo and Matute, in press; Vadillo et al., 2005).

Of course, internet-based research clearly poses important methodological problems that cannot always be overcome. These problems are mainly related to the loss of control over the experimental conditions (e.g., the experiment is performed under circumstances that could diminish the attention participants pay to the task) and over participants (e.g., participants could decide to perform the experiment several times). However, many of these problems can usually be reduced by adopting appropriate solutions that have already been explored and published (e.g., Gosling et al., 2004; Reips, 2002; Schmidt, 1997). Additionally, the comparison of traditional studies and online studies has shown that similar results are usually obtained in both domains in a wide variety of research areas (e.g., Buchanan and Smith, 1999; McGraw et al., 2000; Vadillo and Matute, in press). This indicates that whatever potential variables are introduced by internet-based methodologies, their effect is not large. As a counterpart, web-based methodologies offer interesting opportunities to psychological science. For example, the internet allows the researcher: (1) to obtain incomparably larger, more diverse and more heterogeneous samples that guarantee the generality of the results, (2) to reduce the potential effect of the experimenter bias and (3) to reduce the costs of the experiment. Given these and other advantages, it is not surprising that many researchers are deciding to use this kind of research modality, in spite of the above mentioned methodological problems.

This situation is probably creating some discrepancies between those researchers willing to perform their studies over the internet and some members of the scientific community concerned with the validity of these investigations. One important source of this problem is that online experiments are usually presented as an alternative to traditional experiments. However, a much better approach would be to explicitly consider internet-based and traditional methodologies as complementary research

* Corresponding author. Tel.: +34 94 413 92 49; fax: +34 94 413 90 89.
E-mail address: matute@fice.deusto.es (H. Matute).

0376-6357/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.beproc.2006.01.014

strategies. Perhaps the most appealing feature of internet-based research is that it allows one to perform experiments with little cost. Research teams with few resources or insufficient availability of participants could run their pilot experiments over the internet, and perform similar experiments in the laboratory only when they have gathered sufficient parametric or preliminary data over the internet. This strategy would keep the advantages of traditional laboratory research (because the most critical data are obtained under well-controlled conditions), while reducing to a great extent the costs of the experimental series (see Vadillo and Matute, in press). Additionally, internet-based methodology provides access to larger and more heterogeneous samples than those of traditional studies. Replicating experiments over the internet can provide an easy and economical way to test the generality of the results in different populations. Instead of simply taking the generality of the studied effects for granted, researchers could repeat their experiments over the internet, which would provide stronger evidence to assess the generality.

In the present paper we replicate a well-known phenomenon under traditional laboratory conditions and over the internet as a means to test the appropriateness of internet-based research for the study of associative learning. Among the many associative learning effects discussed in recent literature, few are as important as the so-called cue interaction effects. Cue interaction shows that the strength of the response to a cue does not only depend on the number of cue–outcome pairings, but also on other predictors of the outcome that are present while the training of the target cue takes place. For example, Pavlov (1927) showed that the conditioned response elicited by a conditioned stimulus paired with an unconditioned stimulus was less intense if another stimulus was present during the target conditioned stimulus-unconditioned stimulus pairings, an effect known as overshadowing. Overshadowing and other cue interaction effects (e.g., blocking; see Kamin, 1968) have been traditionally studied in the area of animal conditioning (e.g., Mackintosh, 1975; Miller and Matzel, 1988; Pearce and Hall, 1980; Rescorla and Wagner, 1972; Wagner, 1981), but have also had a major impact in several areas of psychology such as causal learning (e.g., Price and Yates, 1993; Van Hamme and Wasserman, 1994), categorization (e.g., Gluck and Bower, 1988; Kruschke and Johansen, 1999) and social psychology (e.g., Van Rooy et al., 2003). Given that overshadowing is the simplest and probably the most robust cue interaction effect, we tried to replicate it in the area of human predictive behavior, as a means to check whether cue interaction effects can be observed in internet studies.

1. Method

1.1. Participants and apparatus

The laboratory sample was composed of 106 undergraduate students from Deusto University who volunteered to take part in the experiment. All these participants performed the experiment in a large computer room, with each participant at least at 1.5 m apart from the adjacent participant. The internet sample was composed of 91 anonymous participants who visited our virtual laboratory's web page (www.labpsico.com). In order to comply with ethical regulations for human research over the internet (Frankel and Siang, 1999) we decided not to register any variables that could affect participants' privacy. Thus, the conditions in which these participants performed the experiment are completely unknown to us.[1] The experimental program was embedded in an HTML document, using JavaScript code to manage the presentation of the stimuli on the screen and to collect participants' responses (see Bárcena et al., 2003). The program preloads all the stimuli in the computer's memory before participants can start the experiment, so that differences in the connection speed cannot influence the pace of the experiment.

1.2. Procedure and design

The experimental preparation used in this study has been previously used in several laboratory experiments on human associative learning (Pineño and Matute, 2000; see also Dieussaert et al., 2002). Participants were asked to imagine that they were soldiers of the United Nations and that they had to rescue people that were hidden in a ramshackle building. On each trial they were allowed to place a number of refugees in a truck and to take them to a safe place, by pressing the space bar. They could control the number of refugees they put into the truck by pressing the space bar more or less times. However, they could not be sure that the refugees they put into the truck would survive because the road could sometimes contain mines. Each truck was equipped with a spy-radio that signalled the state of the road by presenting some color lights on its panel: some lights indicated that the road would be safe and that the refugees in the truck would survive (and, therefore, that the participant should place as many refugees as possible in the truck) and some other lights indicated that the road would be mined and that the refugees placed in the truck would die (and, therefore, that the participant should avoid placing refugees in the truck). However, participants were not told which lights predicted which outcome. They had to learn this by paying attention to what happens when each color light was presented. Participants earned one point for each refugee that survived and lost one point for each refugee that died. The number of refugees placed in the truck on each trial was thus taken as a measure of the degree to which participants expected that the cue (i.e., the color light) presented in that trial predicted

[1] An unfortunate consequence of this is that it cannot be guaranteed that some participants have not performed the experiment several times. There are, however, some means by which one can control whether or not this is happening. For example, one can register the IP of the computer from which experimental data arrive. If different data files arrive from the same computer and especially if these data files arrive close in time, one can suspect the same participant is taking part in the experiment several times. Deleting these data files can reduce the noise introduced by these participants. However, it would still be possible that some participants performed the experiment several times from different computers, though this behavior does not seem to make much sense. The experimenter can try to compensate the noise that this might introduce in the data by using large samples where the potential effect of this factor is minimized. Alternatively, the experimenter can try to reproduce the results in the laboratory, in order to check that the potentially negative effects of this factor have not altered the outcome of the experiment. In our experiment we have chosen this later control procedure (see Reips, 2002, for alternative means to reduce the likelihood or impact of multiple data submission).

that the road would be safe. (For a demonstration of this task, as well as for complete instructions and source code, see Bárcena et al., 2003).

Overshadowing was assessed using a within-subjects design in which all participants were presented with three types of training trials. In some trials, two colors, cues A and X, appeared simultaneously in the spy-radio signaling the possibility of gaining points by pressing the space bar (i.e., the road was safe in those trials). In other trials, a different color, cue Y, was presented alone also signaling the possibility of gaining points. This was used as the control cue. Finally, to avoid participants' responding in all trials and to make them discriminate between colors, there was a third type of trial in which a different color, cue B, appeared alone indicating that participants would lose points if they pressed the space bar during those trials (i.e., the road was mined in those trials). Four trials of each type were presented in random order during training. The colors that served as cues A, B, X and Y were blue, yellow, red and green, all counterbalanced in a latin square. After the training sequence of trials, cues X and Y were presented in two test trials (counterbalanced for order) in which participants were given no feedback. Less responding to cue X (that had always been presented together with cue A during training) than to cue Y (that was trained alone) was indicative of overshadowing.

2. Results

Before performing the statistical analyses, a data selection criterion was applied to eliminate participants that might have been paying little attention to the task. Specifically, data from participants who responded more in the last training trial with cue B (for which responding was penalized) than in the last training trials with AX or Y (for which responding was rewarded) were eliminated from the analyses. The data from 10 participants in the laboratory condition and from 12 participants in the internet condition were eliminated following this criterion.

The mean responses to cues X and Y in the test trials are depicted in Fig. 1. As this figure shows, the number of space bar presses in the X test trial was lower than the number of responses in the Y test trial, suggesting overshadowing both in the laboratory and in the internet conditions. Additionally, participants in the laboratory condition responded slightly more to both cues than participants in the internet condition. These impressions were confirmed by a 2 (location: laboratory versus internet) × 2 (cue: X versus Y) mixed ANOVA performed on the number of responses at test. This ANOVA showed a marginally significant main effect of location, $F(1, 173) = 3.54$, MSE = 688.69, $p = 0.061$, and a significant effect of cue, $F(1, 173) = 16.2$, MSE = 5675.43, $p < 0.001$. No interaction between these two factors was found, $F < 1$. Planned comparisons showed that the main effect of cue was due, as expected, to participants' responding to X being lower than their responding to Y both in the laboratory, $t(95) = 3.03$, $p < 0.01$, and in the internet, $t(78) = 2.78$, $p < 0.01$. Thus, overshadowing was obtained in both conditions.

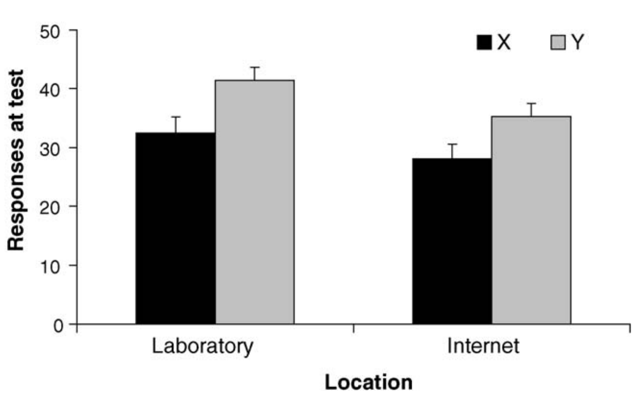

Fig. 1. Mean responses to cues X and Y at test for participants in the laboratory condition and in the internet condition. Error bars denote standard errors of the means.

3. Discussion

The pattern of results confirms the presence of a significant overshadowing effect regardless of the location in which the experiment is conducted. Overshadowing is defined as a low level of responding to a stimulus due to its being trained as a predictor of an outcome always in the presence of an additional stimulus. In other words, participants are more reluctant to predict an outcome after a cue if this cue has been paired with the outcome always in compound with a second cue. There are several reasons why this effect might take place. Some learning theories (e.g., Rescorla and Wagner, 1972) propose that all the cues that are trained together as predictors of an outcome compete among themselves to become associated with the outcome. Other models suggest that all the cues that predict an outcome become associated with it and that cue interaction arises because there can be post-training interferences between different predictors of the same outcome if these predictors have been trained in compound (e.g., Miller and Matzel, 1988). A third type of models (e.g., Pearce, 1987, 1994) explains cue interaction effects, such as overshadowing, as a generalization decrement. From this point of view, participants would have learned during training that the compound AX predicts the outcome. In spite of this, participants do not respond to X at test because the configural stimulus X alone is different from the compound AX that has been trained as a predictor. Thus, not all the knowledge about the AX compound will be generalized to X alone, which results in low responding to X.

Our experiment was not designed to test these explanations against each other. Thus, we cannot favor any of them over the others. Instead, the purpose of the present study was to demonstrate that overshadowing could be obtained both under the traditional laboratory conditions and over the internet. Moreover, we have demonstrated that the location of the experiment (laboratory versus internet) does not affect the size of the overshadowing effect (i.e., there is no interaction between location and overshadowing). This result is encouraging for researchers lacking resources to perform pilot studies to test their procedures or to adjust experimental parameters. In light of the evidence presented here and elsewhere (Birnbaum, 1999; Birnbaum and Wakcher, 2002; Buchanan and Smith, 1999; McGraw et al., 2000; Steyvers et al., 2003; Vadillo and Matute, in press; Vadillo et al., 2005), there is no reason to expect differences in the nature of the studied process arising from the use of an internet-based methodology, as compared to the more traditional laboratory methods. A preparation that works well for the study of one phenomenon in the internet is likely to work similarly well for the study of the same phenomenon in the laboratory.

We found, however, a marginally significant main effect of the location that should be considered carefully. This effect indicates that responses tended to be generally lower in the internet condition than in the laboratory. The lower level of responding in the internet might be due to several reasons such as, for example, participants in the internet condition being less motivated or more reluctant to respond. Another possibility is that this might be a product of the different hardware used in both locations. In this task, participants can place people in the truck rapidly by keeping the space bar pressed down. However, the precise amount of responses registered per second depends on the keyboard's setup.[2] Thus, these results might be due to the laboratory computers' having a different configuration from that of the computers used by the internet participants. The potentially negative effect of the participants' computer having diverse and uncontrolled configurations should always be taken into account by researchers performing their experiments over the internet. However, if the sample is large enough, the effect of these uncontrolled variables should be equally distributed across all experimental conditions. Thus, even if they have some effect, this would be evident in a larger variance or in a main effect of the location, rather than in an interaction between the studied effect and the location.

The results of the experiment presented here also suggest that the learning phenomena under study are general enough to appear even under less controlled situations than those of the traditional laboratory. Therefore, they provide more support to the claim that cue interaction phenomena in humans are not mere artifacts of the experimental procedures or of the kind of samples used, but they reflect the natural way human beings in general process information. In the present experiment we did not collect socio-demographic data about our sample in order to keep participants' privacy. We cannot, therefore, guess how different our internet sample was from the laboratory sample. It would not be surprising if our internet sample was composed mostly of psychology students and people curious about psychology. However, one might still expect more heterogeneity (e.g., regarding age, country, etc.) in this sample than in the laboratory one. Moreover, the heterogeneity can be encouraged, if necessary, by means such as posting links to the experimental programs in web sites with diverse contents.

Regardless of the problems that internet studies might sometimes pose to draw firm conclusions from the results, some of their peculiar characteristics make them especially useful to achieve certain goals that usually remain out of the scope of traditional studies. As we have argued, the access to larger and more heterogeneous samples allows to test the generality of many basic cognitive processes that are commonly studied using only college students as participants. Internet-based research clearly complements these traditional studies by providing more support to researchers' attempts to generalize their results to a wide population. Additionally, the low costs (not only regarding money, but also regarding time and the availability of participants) of internet studies also makes them an appealing means to run preliminary and parametric studies that can then inspire better designs for subsequent traditional laboratory experiments. Our personal opinion is that this strategy keeps the best of both research modalities and results more constructive than either ignoring the potential of online experiments or trusting excessively on them to draw strong conclusions. Instead of discussing *whether* internet experiments should be performed, we should better discuss *how* they should be used.

[2] This is similar to what happens in a word processor when one key is kept pressed down: The speed at which instances of the same letter appear on the screen depends on the keyboard's configuration.

Acknowledgements

Support for this research was provided by Grant PI-2000-12 from Departamento de Educación, Universidades, e Investigación of the Basque Government to Helena Matute. Miguel A. Vadillo was supported by F.P.I. fellowship from the Basque Government (Ref.: BFI01.31). The authors would like to thank Fernando Blanco and Cristina Orgaz for their helpful comments on an earlier draft of this paper.

References

Bárcena, R., Vadillo, M.A., Matute, H., 2003. Spy-Radio Task for Windows [Computer Program], http://www.labpsico.com (retrieved May 6, 2005).

Birnbaum, M.H., 1999. Testing critical properties of decision making on the internet. Psychol. Sci. 10, 399–407.

Birnbaum, M.H., Wakcher, S.V., 2002. Web-based experiments controlled by JavaScript: an example from probability learning. Behav. Res. Methods Instrum. Comput. 34, 189–199.

Buchanan, T., Smith, J.L., 1999. Using the internet for psychological research: personality testing on the world wide web. Br. J. Psychol. 90, 125–144.

Dieussaert, K., Schaeken, W., d'Ydewalle, G., 2002. The quality of test context and contraevidence as a moderating factor in the belief revision process. In: Proceedings of the 24th Annual Meeting of the Cognitive Science Society, pp. 280–285.

Frankel, M.S., Siang, S., 1999. Ethical and Legal Aspects of Human Subjects Research on the Internet. Report of a Workshop convened by the American Association for the Advancement of Science, Program on Scientific Freedom, Responsibility, and Law, Washington, DC, http://www.aaas.org/spp/dspp/sfrl/projects/intres/main.htm (retrieved May 6, 2005).

Gosling, S.D., Vazire, S., Srivastava, S., John, O.P., 2004. Should we trust web-based studies? A comparative analysis of six preconceptions. Am. Psychol. 59, 93–104.

Gluck, M.A., Bower, G.H., 1988. From conditioning to category learning: an adaptive network model. J. Exp. Psychol.: Gen. 117, 227–247.

Kamin, L.J., 1968. Attention-like processes in classical conditioning. In: Jones, M.R. (Ed.), Miami Symposium on the Prediction of Behavior: Aversive Stimulation. University of Miami Press, Miami, FL, pp. 9–31.

Kruschke, J.K., Johansen, M.K., 1999. A model of probabilistic category learning. J. Exp. Psychol. Learn. Mem. Cogn. 25, 1083–1119.

Mackintosh, N.J., 1975. A theory of attention: variations in the associability of stimuli with reinforcement. Psychol. Rev. 82, 276–298.

McGraw, K.O., Tew, M.D., Williams, J.E., 2000. The integrity of web-delivered experiments: can you trust the data? Psychol. Sci. 11, 502–506.

Miller, R.R., Matzel, L.D., 1988. The comparator hypothesis: a response rule for the expression of associations. In: Bower, G.H. (Ed.), The Psychology of Learning and Motivation, vol. 22. Academic Press, San Diego, CA, pp. 51–92.

Pavlov, I.P., 1927. Conditioned Reflexes. Oxford University Press, London.

Pearce, J.M., 1987. A model for stimulus generalization in Pavlovian conditioning. Psychol. Rev. 94, 61–73.

Pearce, J.M., 1994. Similarity and discrimination: a selective review and a connectionist model. Psychol. Rev. 101, 587–607.

Pearce, J.M., Hall, G., 1980. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol. Rev. 87, 532–552.

Pineño, O., Matute, H., 2000. Interference in human predictive learning when associations share a common element. Int. J. Comp. Psychol. 13, 16–33.

Price, P.C., Yates, J.F., 1993. Judgmental overshadowing: further evidence of cue interaction in contingency judgment. Mem. Cognit. 21, 561–572.

Reips, U.-D., 2002. Standards for internet-based experimenting. Exp. Psychol. 49, 243–256.

Rescorla, R.A., Wagner, A.R., 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black, A.H., Prokasy, W.F. (Eds.), Classical Conditioning II: Current Research and Theory. Appelton-Century-Crofts, New York, pp. 64–99.

Schmidt, W.C., 1997. World-wide web survey research: benefits, potential problems, and solutions. Behav. Res. Methods Instrum. Comput. 29, 274–279.

Steyvers, M., Tenenbaum, J.B., Wagenmakers, E.-J., Blum, B., 2003. Inferring causal networks from observations and interventions. Cogn. Sci. 27, 453–489.

Vadillo, M.A., Matute, H. Predictions and causal estimations are not supported by the same associative structure. Q. J. Exp. Psychol., in press.

Vadillo, M.A., Miller, R.R., Matute, H., 2005. Causal and predictive-value judgments, but not predictions, are based on cue–outcome contingency. Learn. Behav. 33, 172–183.

Van Hamme, L.J., Wasserman, E.A., 1994. Cue competition in causality judgments: the role of nonpresentation of compound stimulus elements. Learn. Motiv. 25, 127–151.

Van Rooy, D., Van Overwalle, F., Vanhoomissen, T., Labiuse, C., French, R., 2003. A recurrent connectionist model of group biases. Psychol. Rev. 110, 536–563.

Wagner, A.R., 1981. SOP: A model of automatic memory processing in animal behavior. In: Spear, N.E., Miller, R.R. (Eds.), Information Processing in Animals: Memory Mechanisms. Erlbaum, Hillsdale, NJ, pp. 5–47.