Effect of Paper Color and Question Order on Exam Performance

Ilanit R. Tal, Katherine G. Akers, and Gordon K. Hodge
University of New Mexico

To deter cheating, teachers commonly use exams printed on differently colored paper or with varied question orders. Previous studies, however, reported that paper color and question order affect exam performance and suggested that teachers should adjust students’ scores accordingly and discontinue the use of alternate exam forms. We conducted 2 experiments testing the effects of paper color and question order on exam performance. Students performed worse on exams printed on primary blue paper, but pastel colors had no effect. We found no effect of exam question order (sequential vs. random). We conclude that colored paper or differently ordered questions do not affect exam performance.

Teachers often use alternate exam versions to curb cheating. However, differently colored paper (Sinclair, Soldat, & Mark, 1998; Skinner, 2004; Soldat, Sinclair, & Mark, 1997) and differently ordered questions (Balch, 1989) commonly distinguish alternate versions, but appear to affect performance.

Although Michael and Jones (1955) reported no effect of paper color on exam performance, several more recent studies have contradicted these results. Soldat et al. (1997) found students scored higher on difficult GRE-type questions printed on blue compared to red paper. Sinclair et al. (1998) reported students scored higher on a midterm exam printed on blue versus red paper. These researchers supposed that the color red induces positive affect, triggering a processing strategy that conserves cognitive resources. On the other hand, the authors suggested that the color blue induces negative affect, triggering detail-oriented processing that might be more effective when solving difficult problems. Thus, they argued that better performance on blue versus red exams may result from different affective states and cognitive processes induced by the different colors (cf. Elliot, Maier, Moller, Friedman, & Meinhardt, 2007). In contrast to these findings, Skinner (2004) reported that students scored higher on exams printed on red than on blue paper, with the highest scores occurring on exams printed on white paper. Skinner (2004) and Sinclair and colleagues (1998) urged teachers to discontinue use of differently colored exam versions or at least to adjust students’ exam scores to account for any disadvantages resulting from using different colors of paper.

Concerning question order, students performed better on exams containing questions paralleling the order of presentation in lectures and the textbook (sequential order) than on exams containing randomly ordered questions (Balch, 1989). To explain this effect, Balch suggested that items are best retrieved from memory in the same context in which they were encoded. Exams with sequentially ordered questions present the items in the same order in which they were presumably encoded, and therefore students would perform better under this condition. Other studies, however, have not replicated Balch’s effect (Goss Lucas & Bernstein, 2005; Neely, Springston, & McCann, 1994; Perlini, Lind, & Zumbo, 1996).

In our classes, we use pastel colors to distinguish different exam versions rather than the vivid primary colors used in previous research (N. E. Skinner, personal communication, November 14, 2005; Sinclair et al., 1998; Soldat et al., 1997), and we present multiple-choice questions and alternative answers in random rather than sequential order. Our concern was whether the effects found with primary colors also occur with pastel colors (i.e., red vs. pink). In this study, we attempted to replicate previously reported effects of primary-colored paper and also tested whether pastel colors of paper affect exam performance. To address whether question order (sequential vs. random order)
affects exam performance, we included question order in our study design. In addition, we included gender in our analyses because previous researchers (Elliot et al., 2007; Michael & Jones, 1955; Sinclair et al., 1998) considered student gender as a factor that might interact with paper color to affect exam performance.

Method and Results

Experiment 1: Exams on Primary-Colored Paper

Undergraduate introductory psychology students (364 men, 230 women) took a 40-question multiple-choice exam based on textbook (Davis & Palladino, 2004) and lecture content. Exams were printed in black ink on either white paper (Xerox Business Multipurpose 4200 Bright 92) or four primary colors of paper: RIV 02054 rojo red, RIV 02055 lemon yellow, RIV 02057 emerald green, and RIV 02059 marine blue (Riverside Paper Kaleidoscope Multipurpose). Within each color, questions appeared in either sequential or random order, resulting in a total of 10 exam versions. Students seated adjacent could not have the same color of exam; seating was unassigned.

We performed an ANOVA (SPSS Version 12.0) with number of questions correct as the dependent variable and paper color, question order, and gender as fixed factors. We used Type II sums of squares to correct for unequal cell sizes and tested a priori contrasts based on previous research. We found only a significant main effect of color, \( F(4, 594) = 5.53, \ p < .001, \eta^2 = .04 \), but this model explained only 3.3% of the variance in exam performance (adjusted \( R^2 \)). Contrasts revealed that students performed better on white exams than all other colored exams (M difference = 2.15), \( t(183.26) = 2.85, \ p < .01 \), and worse on blue than all other colored exams (M difference = -3.27), \( t(197.6) = -4.34, \ p < .001 \). We corrected degrees of freedom for unequal variances. We found no effects for question order or gender.

Experiment 2: Exams on Pastel-Colored Paper

A different group of undergraduate introductory psychology students (434 men, 247 women) took a 40-question multiple-choice exam under similar conditions as those described in Experiment 1, except that exams were printed on either white paper (Xerox Business Multipurpose 4200 Bright 92) or four pastel colors of paper: 32571 pink, 32541 light yellow, 32561 light green, 32521 light blue (Wausau Paper Exact Multipurpose).

We performed an ANOVA with number of questions correct as the dependent variable and paper color, question order, and gender as fixed factors. We tested a priori contrasts as in Experiment 1. We used Type II sums of squares to correct for unequal sample sizes and found no effects of color, question order, or gender.

Discussion

We found that printing exams on primary but not pastel colors of paper affected exam performance. Consistent with Skinner (2004), students performed better on exams printed on white paper than on vividly colored paper. Also consistent with Skinner, students performed worse on blue exams compared to other colored exams; this finding, however, contradicted the findings of Sinclair et al. (1998) and Soldat et al. (1997), who found worse performance on red exams. Although the conflicting evidence regarding the directionality of the red and blue paper effects on exam performance may be a result of differences in the hues of paper colors used among studies, the size of our primary color effect was relatively small and toward the lower end of the normal range of .01 to .14 (Cohen, 1988). We could retroactively calculate effect size (Laake, 1983) only for the Sinclair et al. (1998) study, and it was small (\( \eta^2 = .02 \)) as well. Given the small effect sizes and contradictory findings across studies with respect to which colors improve or decrease performance, we conclude there are no meaningful or consistent effects of primary color on exam performance.

We repeated the experiment using pastel colors of paper, which is how we distinguish alternate exam versions in our classes, and found no effect of paper color on exam performance. This lack of effect may occur because lighter colors do not induce the same affective and cognitive states as more vivid colors. Regardless, concerns regarding unfair advantages resulting from different colors of exam paper are alleviated when teachers use pastel rather than primary colors of paper.

In addition to paper color, we also tested whether question order affected exam performance. Although Balch (1989) found that students performed better on exams containing questions in sequential compared to random order, the results of our study as well as those by Goss Lucas and Bernstein (2005), Neely et al. (1994), and Perlini et al. (1998), did not support a question
order effect. Also, consistent with previous studies (Elliot et al., 2007; Michael & Jones, 1955; Sinclair et al., 1998), we found no effects for gender and no interaction with gender.

In summary, we found that paper color and question order do not meaningfully predict differences in exam performance. Faculty can use differently colored paper or differently ordered questions to deter cheating and not worry about affecting grades.

References


Notes

1. Our thanks to Kathryn Wiggins, Harold Delaney, Randolph Smith, and three anonymous reviewers.
2. Send correspondence to Gordon K. Hodge, Department of Psychology, Logan Hall, MSC 03 2220, University of New Mexico, Albuquerque, NM 87131-0001; e-mail: ghodge@unm.edu.