

BRIEF REPORT

In a randomized study of envelope and ink color, colored ink was found to increase the response rate to a postal questionnaire

Kate S.M. Taylor^a, Carl E. Counsell^{a,*}, Clare E. Harris^a, Joanna C. Gordon^a,
Sofia C.C. Fonseca^b, Amanda J. Lee^c

^aDepartment of Medicine and Therapeutics, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen, AB25 2ZD, United Kingdom

^bDepartment of Public Health, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen, AB25 2ZD, United Kingdom

^cDepartment of General Practice and Primary Care, Foresterhill Health Centre, Westburn Road, Aberdeen, AB25 2AY, United Kingdom

Accepted 31 March 2006

Abstract

Objective: To assess the effect of the colors of the envelope and ink on the response rate to a postal questionnaire in a study screening for undiagnosed parkinsonism in people aged 65 years and over in the community.

Study Design and Setting: A total of 2,524 people aged 65 years and over from five general practices in Aberdeen were randomized to receive a questionnaire about the symptoms of parkinsonism printed in either colored (green) or black ink, and sent out in either a brown or white envelope.

Results: The overall response rate was 63.5%. There was no significant interaction between envelope and ink color. The use of green ink compared to black significantly increased the response rate from 61.4% to 65.7% (OR 1.20, 95% confidence interval 1.02, 1.41). There was no overall effect of envelope color on response rate (62.3% brown and 64.8% white, OR 0.90, 95% confidence interval 0.76, 1.06) but there was significant heterogeneity between the general practices. When this general practice–envelope interaction was accounted for, brown envelopes had a significantly lower response rate than white ones (OR 0.49).

Conclusion: This study, along with existing evidence, has shown that the use of certain ink colors in postal questionnaires is likely to increase response rates relative to black ink. The effect of envelope color was inconsistent both within this study and between previous studies. © 2006 Elsevier Inc. All rights reserved.

Keywords: Data collection; Healthcare surveys; Postal questionnaire; Response rate; Envelope color; Ink color

1. Introduction

Many research studies rely on postal questionnaires but such studies can be compromised by poor response rates, which reduce statistical power and can also introduce non-response bias. For example, studies in the elderly often achieve response rates of 60% or less [1,2] and non-responders often have different characteristics or outcomes to responders [1]. Therefore, researchers need to identify and use methods that have been reliably shown to improve response rates. A Cochrane systematic review identified several techniques that have been proven to improve

response rates to postal questionnaires [3] and also highlighted several approaches worthy of further study. Two of the latter, which would be easily applicable, were whether the color of the envelope (brown or white) or the color of the ink on the questionnaire (colored or standard black/blue) affected the response rate.

Three studies with a total of 6,017 people from predominantly younger populations (a survey of headache [4], a survey of back pain in nurses [5], and an industry survey [3]) have assessed the effects of brown or white envelopes. Meta-analysis showed a significantly greater response rate with brown envelopes [3], but this result was entirely due to one study showing a large difference in favor of brown envelopes [5], whereas the other two studies showed no difference between the use of brown and white envelopes. A single study, of 3,540 tire dealers, from the field of industry rather than medicine assessed the use of colored (blue on yellow background) or black ink on the questionnaire [6], and showed a promising effect in favor of colored ink, but this had not been replicated in the medical field.

* Corresponding author. Tel.: +44-1224-551117; fax: +44-1224-554761.

E-mail addresses: ksmp@iahs.abdn.ac.uk (K.S.M. Taylor) or carl.counsell@abdn.ac.uk (C.E. Counsell) or c.harris@abdn.ac.uk (C.E. Harris) or j.c.gordon@abdn.ac.uk (J.C. Gordon) or a.j.lee@abdn.ac.uk (A.J. Lee).

As part of an incidence study of parkinsonism in north-east Scotland, we screened elderly patients living in the community for undiagnosed parkinsonism with a postal questionnaire [7]. We, therefore, took the opportunity to assess whether brown envelopes and colored ink increased the response rate in our cohort. To our knowledge, only one previous study had used a postal questionnaire to screen for undiagnosed parkinsonism [8], which showed an excellent response rate of 84% after three mailings. However, this study was small (200 people) and included a population younger than ours (over 60 as opposed to over 65 years).

2. Methods

2.1. Design and participants

All people aged 65 years and over on the lists of five general practices in Aberdeen were eligible provided that their general practitioner (GP) felt it was appropriate for them to receive a questionnaire. Samples of about 550 people per general practice were randomly selected using true random numbers (www.random.org). The GPs then excluded patients they felt were inappropriate for the study, for reasons such as terminal illness or severe anxiety, whereas some others died while the lists were being compiled. The remainder were randomly allocated a number from 1 to 4 (again using true random numbers) to determine

whether they would receive a questionnaire printed in green or black ink in a brown or white envelope. Randomization was not stratified by any variable. In line with best evidence to improve response rates [3], the questionnaires were short, were sent with an explanatory letter on university-headed paper personally signed by the principal investigator and with a stamped addressed reply envelope that matched the color of the outer envelope. We did not use monetary incentives or precontact. The four-page questionnaire included 13 screening questions for parkinsonism and a EuroQuol EQ-5D. The colored questionnaires were printed with dark green ink on a light green background to give clear readable text, which was particularly important for our elderly sample. Three mailings were sent (total response time 6 months), and each person received the same color envelope and ink for each mailing.

Limited demographic data were available from the general practice lists including gender (from all five practices), postcode and age (from four practices each). A social deprivation score (DEPCAT) was generated from the postcode in which greater deprivation is represented by a higher score [9].

2.2. Statistical methods

The sample size was based on power calculations for the screening project [7] not the response rate. However, post

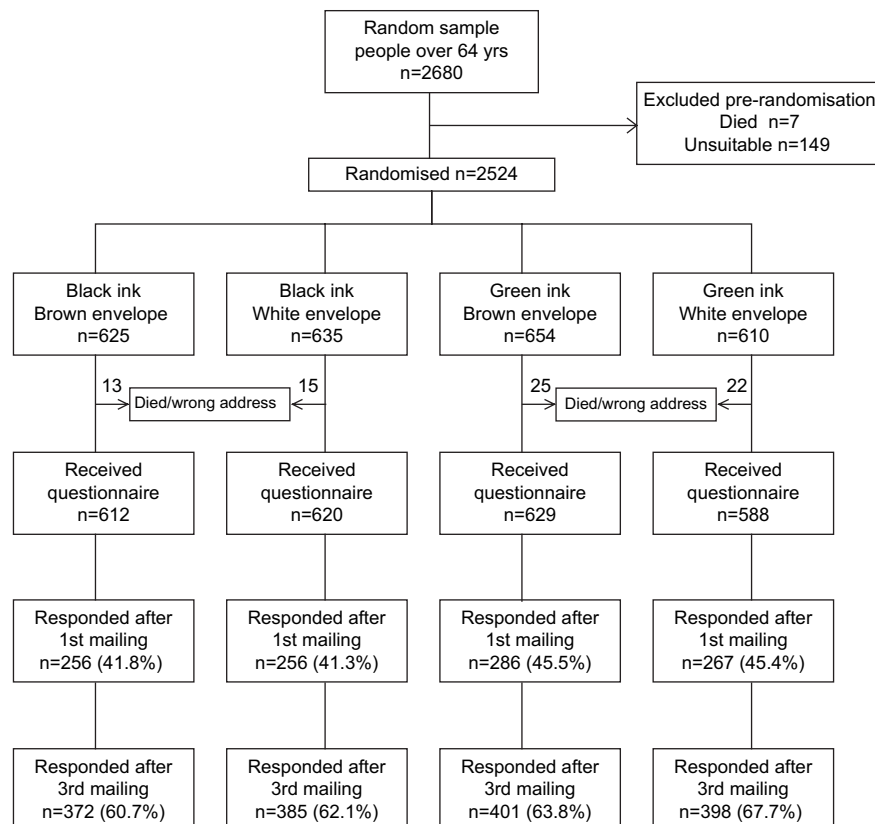


Fig. 1. Trial flow diagram.

Table 1

Baseline characteristics by randomized group

| | Group 1 (n = 612) | Group 2 (n = 620) | Group 3 (n = 629) | Group 4 (n = 588) | Total (n = 2449) | P-value |
|---------------------|-------------------|-------------------|-------------------|-------------------|------------------|---------|
| Ink color | Black | Black | Green | Green | | |
| Envelope color | Brown | White | Brown | White | | |
| Male (%) | 236 (38.6%) | 248 (40.0%) | 265 (42.1%) | 256 (43.5%) | 1005 (41.0%) | 0.300* |
| DEPCAT 1–3 (%) | 156 (33.6%) | 151 (30.3%) | 171 (35.2%) | 158 (33.7%) | 636 (33.2%) | 0.421* |
| Mean age years (SD) | 75.1 (6.4) | 75.0 (7.1) | 75.6 (7.0) | 75.3 (7.0) | 75.2 (6.9) | 0.551** |

*Derived using chi-squared test; **derived using ANOVA test.

hoc analysis showed we had 65% power to detect a 4% difference in response rates with an alpha of 5% and a baseline response rate of 60%. Univariate odds ratios (OR) were calculated (SPSS version 12; SPSS, Inc, Chicago, IL) for ink and envelope color whereas multiple logistic regression also included gender, general practice, and interaction terms (envelope \times ink, ink \times practice, envelope \times practice). We could not correct for age and social deprivation, as these data were not available for all people. The results were added to the existing Cochrane review (Review Manager version 4.2), and heterogeneity was assessed using the I^2 statistic [10].

Ethical approval for the study was obtained from the Grampian Regional Ethics Committee.

3. Results

A total of 2,860 people aged 65 and over were randomly selected of whom 156 were deemed unsuitable by their GP or had died. Of 2,524 people randomized, 75 did not receive the questionnaire because they had died or it was returned undelivered because we had the wrong address (Fig. 1). These patients did not have the opportunity to respond and, therefore, the main analysis was restricted to 2,449 people. There were no significant differences in gender, social deprivation, or age between the four groups (Table 1) although DEPCAT data were available for only 1,917 people (78%) and age for 1,987 (81%).

The overall response rate was 63.5% (1,556/2,449). Univariate analysis showed that the response rate for brown envelopes (773/1,241, 62.3%) was not significantly different to the response rate for white envelopes (783/1,208, 64.8%) (OR 0.90, 95% CI 0.76, 1.06) whereas the response rate for colored ink (799/1,217, 65.7%) was significantly better than the response rate for black ink (757/1,232, 61.4%) (OR 1.20, 95% CI 1.02, 1.41). Similar differences were apparent even after the first mailing (brown 43.7% versus white 43.3%, green 45.4% versus black 41.6%). If those who were randomized but did not receive the questionnaires were included in the analysis as nonresponders (a true “intention to treat” analysis), the results for envelope color were unchanged (OR 0.90, 95% CI 0.77, 1.06), but the result for ink color became nonsignificant (OR 1.14, 95% CI 0.97, 1.34).

Multivariate analysis showed no significant interaction between envelope and ink color (OR 1.13, 95% CI 0.81, 1.57, $P = 0.478$), but there was a significant interaction between envelope color and general practice ($P = 0.001$). Thus, a model was chosen that considered the two main effects (ink and envelope color) adjusted for gender and general practice and the interaction term between general practice and envelope (Table 2).

This model confirmed a 20% increase in the odds of response to questionnaires printed in green ink compared to black ink. The model also showed a doubling of the odds of response to questionnaires posted and returned in white compared to brown envelopes. Although there were significant differences between the five general practices in terms of socioeconomic status (DEPCAT) and mean age (but not gender), the significant interaction between general practice and envelope color was not readily explained by these differences (Table 3). When we analyzed the results by age and socioeconomic status in those from whom we had these data, the only suggestion of an interaction was between colored ink and socioeconomic status: green ink had lower response rates than black ink (OR 0.70, 95% CI 0.47, 1.03) in those with least deprivation (DEPCAT 1–2) but higher response rates (OR 1.23, 95% CI 0.99, 1.53) in those with greater deprivation (DEPCAT 3–6).

Adding our unadjusted results to the Cochrane review [3] showed that there was clear evidence that questionnaires printed in colored ink compared to those printed in black ink increased response rates (fixed effect OR 1.28, 95% CI 1.13, 1.45, $P < 0.0001$, I^2 28.6% based on two trials with 5,989 participants in total). There was a nonsignificant trend for greater response rates with brown envelopes compared to white envelopes (random effects OR 1.19, 95% CI

Table 2
Results of multiple logistic regression

| | OR (95% CI) | P-value | Direction of effect |
|----------------------------------|-------------------|---------|----------------------------|
| Envelope color | 0.49 (0.33, 0.73) | <0.001 | Favors white envelope |
| Ink color | 1.20 (1.01, 1.41) | 0.035 | Favors colored (green) ink |
| General practice | 1.28 (1.07, 1.54) | 0.007 | |
| Gender | 0.81 (0.68, 0.96) | 0.013 | Favors male |
| Envelope color \times practice | 0.82 (0.73, 0.92) | 0.001 | |

Table 3

Baseline characteristics and response rates to envelope by general practice

| | Practice 1 | Practice 2 | Practice 3 | Practice 4 | Practice 5 | Total | P-value |
|------------------------|-------------|-------------|-------------|-------------|-------------|--------------|---------|
| Characteristic | | | | | | | |
| Number of people | 477 | 467 | 517 | 462 | 526 | 2449 | |
| Number male (%) | 201 (42.1%) | 197 (42.2%) | 217 (42.0%) | 173 (37.4%) | 217 (41.3%) | 1005 (41.0%) | 0.533* |
| Number DEPCAT 1–3 (%) | 125 (26.2%) | 193 (41.3%) | 123 (23.8%) | 195 (42.2%) | Unavailable | 636 (33.1%) | <0.001* |
| Mean age in years (SD) | 74.4 (6.3) | 76.0 (7.0) | 74.9 (6.9) | Unavailable | 75.7 (7.0) | 75.2 (6.9) | 0.002** |
| Response rate to | | | | | | | |
| Brown envelope | 56.1% | 63.1% | 65.4% | 65.0% | 61.5% | 62.3% | |
| White envelope | 66.0% | 71.4% | 67.5% | 66.2% | 52.9% | 64.8% | |

*Derived using chi-squared test; **derived using ANOVA test.

0.76, 1.88, four trials with 8,466 participants in total), but there was extreme heterogeneity (I^2 94.9%).

4. Discussion

This study has certain limitations. Firstly, it was performed as a substudy of a screening project, and so the power calculations were derived for the screening project, not the response rates. Secondly, in the main analysis we excluded those who were randomized but did not receive the questionnaire because they had no opportunity to respond. If we included these people, the results for ink color changed slightly but the meta-analysis remained unchanged. Thirdly, we did not have data on age and social deprivation for all people, and so we could not control for these variables in the logistic regression analyses for the whole sample.

However, we do not believe these limitations detract from the main findings. Taken together with previous similar research [6], we have shown that a questionnaire printed with colored as opposed to black ink increases the response rate in postal surveys. The absolute difference was small (4.3%), but any improvement in response rate is worthwhile providing it can be achieved with little effort and cost. Although we did not perform a formal cost-effectiveness analysis, the additional cost of using green ink in our study was very small (total ink costs were approximately £10 for green and £2 for black) and was at least partly offset by lower postage costs (green ink resulted in higher response rates after the first mailing and so fewer people required second and third mailings).

We did not attempt to identify why colored ink resulted in better response rates. It may just be because it was more appealing to the eye. It remains unclear whether the increased response to colored ink is widely generalizable because this was only the second such trial. However, it is encouraging that it appears to work in medical and non-medical settings [6] and in people of working age (although the mean age of people in the industry-based study [6] was not given) and elderly populations. Whether such an effect would be seen in postal surveys in much younger people remains unknown. Although we found that colored ink was less effective in more affluent people, this conclusion

should be treated cautiously as it was based on a subgroup analysis of an incomplete sample. It is also unclear whether there is one specific ink color that is best. The previous study used blue on yellow, whereas we used dark on light green. The improved response rate may, therefore, be a general effect of color rather than being specific to certain colors. However, whatever color is used it must obviously be easily readable.

We did not confirm previous suggestions that brown envelopes produce better response rates than white [5]. In fact, we found that white envelopes were better overall although there was significant heterogeneity between general practices: four had better responses with white envelopes and one with brown. We were unable to explain this heterogeneity on the basis of age or socioeconomic status. Other factors may explain this heterogeneity such as people's previous experience with questionnaires or how much "junk" mail people receive in brown or white envelopes. Similarly, there was enormous heterogeneity in the meta-analysis, largely due to the one trial that showed a much higher response rate with brown envelopes [5]. It may be that this was a chance or biased finding or alternatively that the effect of envelope color is genuinely different across different populations. More trials of brown versus white envelopes are required to clarify this.

5. Conclusions

This study, along with existing evidence, has shown that questionnaires printed with certain ink colors are likely to improve response rates in postal surveys compared to those printed with black ink. We also found that white envelopes improved response over that found with brown envelopes but that major heterogeneity in this effect, both within this study and between this study and previous studies, limited the generalizability of this finding.

Acknowledgments

This study was supported by a grant from the BUPA Foundation, and was part of a larger study funded by the Aberdeen branch of the Parkinson's Disease Society and

the Doris Hillier Award from the British Medical Association.

References

- [1] Paganini-Hill A, Hsu G, Chao A, Ross RK. Comparison of early and late respondents to a postal health survey questionnaire. *Epidemiology* 1993;4:375–9.
- [2] Hardie JA, Bakke PS, Mørkve O. Non-response bias in a postal questionnaire survey on respiratory health in the old and very old. *Scand J Public Health* 2003;31:411–7.
- [3] Edwards P, Roberts I, Clarke M, DiGiuseppi C, Prata S, Wentz R, et al. Methods to increase response rates to postal questionnaires. *The Cochrane Database of Methodology Reviews* 2003, Issue 4. Art. No.: MR000008. doi:10.1002/14651858.MR000008.pub2.
- [4] Newland CA, Waters WE, Standford AP, Batchelor BG. A study of mail survey method. *Int J Epidemiol* 1977;6:65–7.
- [5] Roberts I, Coggan C, Fanslow J. Epidemiological methods: the effect of envelope type on response rates in an epidemiological study of back pain. *J Occup Health Saf Aust New Zealand* 1994;10:55–7.
- [6] La Garce R, Kuhn LD. The effect of visual stimuli on mail survey response rates. *Ind Market Manage* 1995;24:11–8.
- [7] Taylor KSM, Counsell CE, Gordon JC, Harris CE. Screening for undiagnosed parkinsonism in people aged 65 years and over in the community. *Parkinsonism Relat Disord* 2006;12:79–85.
- [8] Meara RJ, Bisarya S, Hobson JP. Screening in primary health care for undiagnosed tremor in an elderly population in Wales. *J Epidemiol Commun Health* 1997;51:574–5.
- [9] McLoone, Philip. Carstairs scores for Scottish postcode sectors from the 2001 census. MRC Social and Public Health Sciences Unit, University of Glasgow. http://www.msoc-mrc.gla.ac.uk/Publications/pub/Carstairs_MAIN.html (accessed March 29, 2006).
- [10] Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.