Misinformation increases symptom reporting: a test – retest study

Harald Merckelbach • Marko Jelicic • Maarten Pieters

Faculty of Psychology and Neuroscience, Maastricht University, PO Box 616, 6200 MD, Maastricht, The Netherlands
Correspondence to: Harald Merckelbach. Email: H.Merckelbach@maastrichtuniversity.nl

Summary

Objectives   We examined whether misleading information (i.e. misinformation) may promote symptom reporting in non-clinical participants.

Design   A test–retest study in which we collected baseline data about participants' psychological symptoms and then misinformed them that they had rated two target symptoms relatively highly. During an interview, we determined whether participants would notice this misinformation and at direct and one-week follow-up, we evaluated whether the misinformation would exacerbate retest measures of the same symptoms.

Setting   A psychological laboratory.

Participants   A total of 78 undergraduate students.

Main outcome measures   Participants' scores on a widely used self-report measure of psychological symptoms.

Results   We found that most participants (63%) were blind to the discrepancies between their original symptom ratings and the upgraded scores they were misinformed with. Furthermore, at the one-week follow-up retest, blind participants revised their symptom ratings in the direction of the misinformation (i.e. they increased their ratings of these symptoms).

Conclusion   Introspective monitoring of common psychological symptoms is poor and this creates an opportunity for misinformation and symptom escalation. Our finding bears relevance to theories about the iatrogenic amplification of medically unexplained symptoms.

Introduction

It has long been recognized that symptom labelling and the medical (e.g. history-taking) and social (e.g. attending patient support groups) actions that accompany such labelling may have iatrogenic effects.[1,2] The mechanisms underlying such effects are poorly understood. Misleading information (i.e. misinformation) with which medical experts accidentally provide their patients might be a candidate mechanism. An extensive body of research shows that misinformation may profoundly bias human memory.[3] However, research on misinformation and symptom

reporting is scarce. So far, no study tested whether misinformation about psychological symptoms leads non-clinical participants to adopt these symptoms in the long run. It would *a priori* seem unlikely that participants would do so, as long as they recognize that the misinformation is at odds with their perception of symptom intensity.

How accurate are people in monitoring the intensity of their psychological symptoms? Research on what has been dubbed 'choice blindness' demonstrates that people may not be very good in monitoring their choices, intentions and preferences. In typical studies illustrating this phenomenon, participants were shown pairs of photographs of faces. They were instructed to select the face they found most attractive. Next, they were given the selected face and asked to describe the reasons for their choice. On some trials, however, participants were given the wrong photograph (i.e. they were misinformed about their choice). Most participants did not notice the discrepancies and these 'blind' participants even tended to confabulate reasons for choices they had not made.[4,5] Other studies found that when people are misled into believing that they selected a certain option (e.g. a vacation destination or a symptom), their endorsement of that option becomes stronger.[6,7]

We examined what happens when participants are misinformed about their symptom scores. More specifically, we explored whether something akin to choice blindness ('symptom blindness') would occur under such circumstances and if so, whether being blind to symptom misinformation would exacerbate one-week follow-up ratings of these symptoms. Our study bears relevance to situations in which healthy individuals with benign symptoms seek medical help and are provided with labels that have strong medical connotations (e.g. fibromyalgia).

Methods

Undergraduate students (*n* = 78; mean age = 20.7 yrs; 60 women) volunteered to participate in the study in return for a small financial compensation. Participants were tested individually. Figure 1 provides an overview of the different stages in the experiment. At baseline (T1), all participants first completed the Symptom Checklist-90 (SCL-90; Derogatis, Lipman, & Covi, 1973).[8] This widely used 90-item scale addresses a broad range of psychological symptoms (e.g. depression, anxiety, fatigue). Participants indicate on a 5-point scale (anchors: 0 = not at all; 4 = all the time) to what extent they experienced each symptom in the past week. We were not interested in total SCL-90 scores, but rather focused on two selected target items embedded in a series of eight control items (see below).

After they had completed the SCL-90, participants received a filler task consisting of two Sudoku puzzles. Each puzzle took 5–7 min. While the participant was busy solving the Sudoko puzzles, two target items of the already completed SCL-90 were manipulated. More specifically, we increased participants' scores for these target items by two full scale points. For example, when the participant rated the target item about concentration difficulties as 0 (i.e. not at all), this score was erased and replaced by a 2 (i.e. occasionally). In those rare instances in which participants had scored the target items with a 3 (a lot) or 4 (all the time), the manipulation consisted of decreasing the items by two full scale points (i.e. they were recoded as 1 or 2, respectively). During subsequent statistical analyses, these scores were reversed coded. In total, 78 × 2 = 156 target items were manipulated. Of these, only four (3%) concerned downgrading.

Target items pertained to repeated unpleasant thoughts, trouble remembering things, blaming yourself for things, worrying too much about things, trouble concentrating, and uneasy when others watch you (SCL-90 items 3, 9, 26, 31, 55, and 61, respectively). These items load on the general emotional distress factor of the SCL-90.[9] We employed three sets of two target items counterbalanced across the sample. Set 1 consisted of SCL-90 items 3 and 9, set 2 of items 26 and 31, and set 3 of items 55 and 61. The experimenter showed participants their SCL-90 answer sheets and interviewed them about why they had rated two targets and eight control items the way they did. Thus, for target items, the interview conveyed misinformation. For example, in the case of the target item about concentration difficulties, the experimenter might ask: 'Could you please tell me why you said that you occasionally have troubles concentrating?' when in fact the participant had responded with not at all to this symptom. Target items were positioned half way in the

Figure 1

Timeline of the experiment. T1 = baseline; T2 = after misinformation about target items; T3 = one-week follow-up; SCL-90 = Symptom Checklist-90 (at T2 and T3 short versions); MCSDS = Marlowe-Crowne Social Desirability Scale

T1 SCL-90 (*N*=78)	Filler tasks	Misinformation	Interview	**T2** MCSDS (*N*=78) SCL-90 (*n*=39)	**T3** SCL-90 (*N*=78)
15 min	14 min		5 min	10 min	1 week later
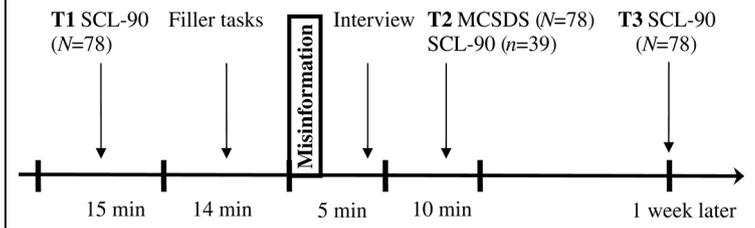
series, with one control item in between the target items. The interview took about 5 min.

At the end of the session (T2), all participants filled out a social desirability scale, the 33-item Marlowe-Crowne Social Desirability Scale (MCSDS).[10] Furthermore, to explore the immediate effects of misinformation, 39 randomly selected participants were given a short 30-item version of the SCL-90. After one week (T3), all participants were contacted by email and asked to complete an electronic version of the 30-item version and to return it by email. Items of the short version were also rated on a 5-point scale. The short version contained the two target and eight control items Participants were fully debriefed after they had returned the electronic version. The study was approved by the standing ethical committee of the Faculty of Psychology and Neuroscience, Maastricht University (ECP-9008-03-2010, April 2010).

Results

Proportion of 'blind' participants

During the interview, 58 (74%) participants accepted the misinformation about the first target item, i.e. explained why they had scored the item with X when in fact they had scored it with X ± 2. Participants were assigned to the 'blind' group if and only if they had failed to notice the discrepancies for both targets items. In total, 49 participants (63%) were blind to both target manipulations. Percentages of blind participants for the three sets of target items ranged from 56–72%, but did not differ significantly as a function of the sets, $\chi^2 (2) = 1.53$, $P = 0.47$.

Social desirability (i.e. MCSDS) scores for blind and non-blind participants were 16.71 (SD = 4.90) and 15.97 (SD = 6.05), respectively, $t(76) < 1.0$, n.s. Thus, blind participants did not have higher social desirability scores than non-blind participants.

Follow-up measures

Figure 2 depicts mean scores of blind and non-blind participants during baseline (T1), and at immediate (T2) and one-week follow-up retests (T3). At baseline, blind participants did not score higher on target than on control symptoms, $t(48) < 1$, n.s. Yet, at the immediate retest (T2) and at the one-week follow-up retest (T3), blind participants scored higher on target symptoms than on control symptoms, $t(24) = 3.95$, $P < 0.01$ and $t(48) = 3.41$, $P < 0.01$, respectively. For the non-blind group, all comparisons between target and control symptoms at T1, T2, and T3 fell short of significance, all $ts < 1.21$, all $Ps > 0.11$.

We also performed a 2 (group: blind vs. non-blind) × 2 (symptoms: target vs. control) repeated measures analysis of variance (ANOVA) on baseline (T1) data. This revealed that before the misinformation manipulation, blind participants already had higher symptom scores than non-blind participants, $F(1, 76) = 8.16$, $P < 0.01$, $\eta_P^2 = 0.10$, although there were no main or interaction effects of symptom (both $Fs (1, 76) < 1$, n.s.).

A 2 (group) × 2 (time: T1 vs. T2) × 2 (symptoms) ANOVA performed on the data gathered in the immediate retest subgroup ($n = 39$) again revealed that overall, blind participants ($n = 25$) had higher symptom scores than non-blind participants, $F(1, 37) = 4.13$, $P < 0.01$, $\eta_P^2 = 0.10$. Also, both target and control symptoms received higher scores over time, $F(1, 37) = 12.73$, $P < 0.01$, $\eta_P^2 = 0.26$. At the immediate retest (T2), target scores were higher than control scores, but only so for the blind group, as evidenced by a group × time × symptom interaction, $F(1, 37) = 16.11$, $P < 0.01$, $\eta_P^2 = 0.30$.

A 2 (group) × 2 (time: T1 vs. T3) × 2 (symptom) ANOVA on baseline and one-week follow-up data indicated that blind participants had higher symptoms scores than non-blind participants, $F(1, 76) = 16.38$, $P < 0.01$, $\eta_P^2 = 0.18$. Furthermore, targets

Figure 2
Mean scores of blind ($n = 49$) and non-blind ($n = 29$) participants on target and control symptoms (range 0–4) during baseline (T1), and at immediate (T2) and one-week follow-up retests (T3). Error bars represent standard errors of the mean
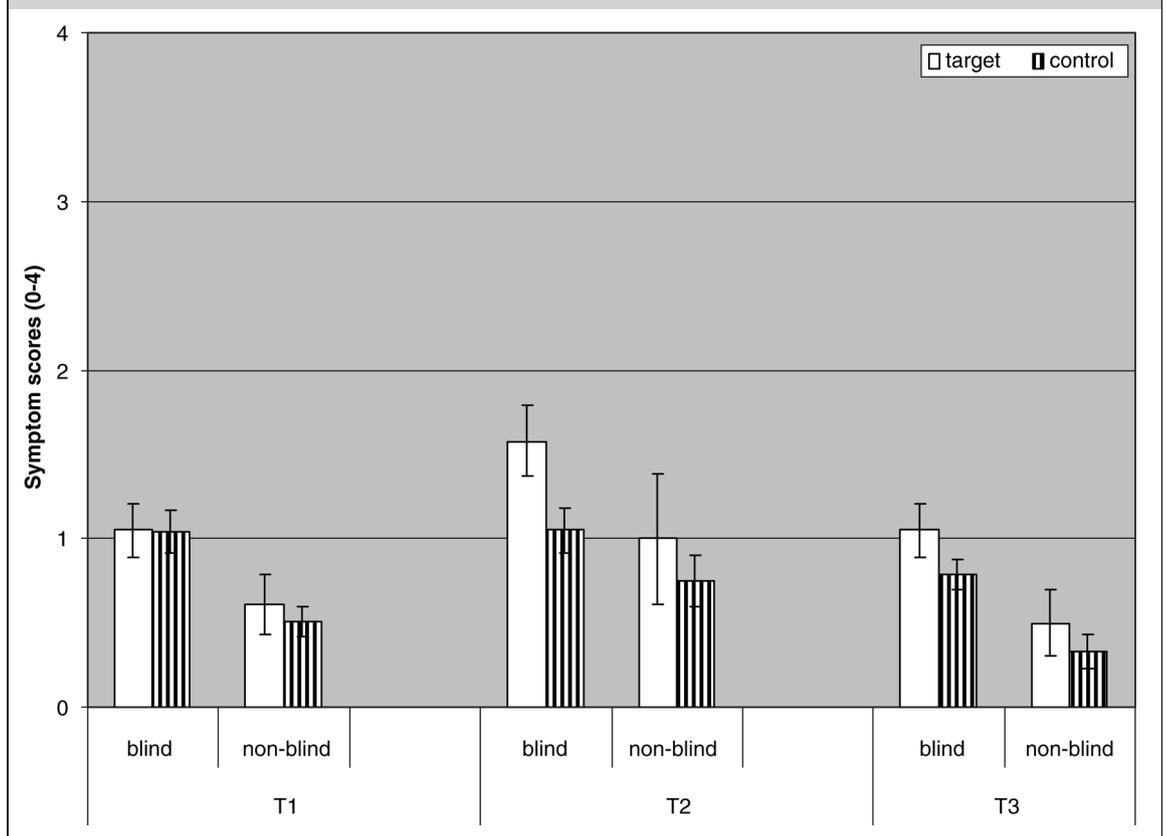
received higher scores than control symptoms at follow-up, $F(1, 76) = 4.37$, $P < 0.05$, $\eta_P^2 = 0.05$, but this was qualified by the critical group $\times$ time $\times$ symptoms interaction that approached significance, $F(1, 76) = 3.62$, $P = 0.06$, $\eta_P^2 = 0.05$.

Discussion

Statement of principal findings

The main findings of our study can be catalogued as follows. First, when confronted with misleading information about certain symptoms, many participants will fail to detect the misinformation. Second, participants who are blind to symptom misinformation tend to adopt this misinformation, as manifested in their raised symptom intensity ratings at follow-up tests. Third, blind participants do not differ from non-blind participants in social desirability, making it unlikely that the crucial difference between both groups is sensitivity to social demands. Rather blindness for symptom misinformation seems to reflect poor monitoring of symptom intensity, just as choice blindness reflects poor introspective insight into choices, intentions and preferences.[4–7]

This interpretation is supported by the finding that already at baseline (T1), blind participants had higher symptom ratings than non-blind participants. Apparently, a non-zero symptom intensity level introduces ambiguity; thereby raising the probability that misinformation is accepted. This pattern is compatible with literature indicating that neurotic people are more suggestible than non-neurotic people.[11] It is also compatible with research that found certain personality

traits to be involved in somatization, i.e. the tendency to experience benign and trivial sensations as noxious and pathological.[12]

Strengths and weaknesses of the study

A limitation of our study is that it relied on undergraduate students and their intensity ratings of psychological symptoms. We do not know to what extent our results can be generalized to patient groups and to intensity ratings of bodily symptoms. Clearly, the boundary conditions of blindness for symptom misinformation deserve systematic study.

Our study is also silent as to the issue of whether misinformation can be used to reduce symptom intensity ratings. Germane to this issue are studies showing that misinformation about a negative experience with fattening food may lead people not only to believe that they had gotten sick after eating this food, but also to avoid this food.[13] It would be interesting to explore whether such positive effects can be paralleled in studies on misinformation and symptom intensity.

Strengths and weaknesses in relation to other studies

Previous studies determined that in certain patient groups, false feedback information about physiological parameters lead to immediate symptom escalation. For example, false feedback of asthmatic wheezing sounds elicits breathlessness in asthmatic children.[14] Likewise, false heart rate feedback increases anxiety in panic patients.[15] One recent experiment in our lab found that when healthy people are misinformed about their psychological symptoms, this will immediately increase their ratings of these symptoms, but only so if they fail to detect the misinformation.[7]

The current study extends previous research in that it provides the existence proof for long-term effects of misinformation on symptom reporting. Our finding is reminiscent of the extensive research on how misinformation may create robust pseudo-memories. Thus, confronting individuals repeatedly with false information about a fictitious event (e.g. as a child having had a skin sample removed as part of a medical procedure), leads many of them to develop detailed recollections of this event.[3,16,17] Such recollections are often maintained even when participants are confronted with contradictory evidence. We suggest that symptom intensity escalation due to misinformation behaves like pseudo-memories. For example, a recent case study illustrated how diagnostic misinformation convinced a 58-year-old woman that she suffered from Alzheimer's disease and how she clung to this diagnosis even when it was proven to be wrong.[18] Our findings were collected in a sample of healthy undergraduates and therefore, they need to be supplemented by clinical case studies on, for instance, diagnostic errors and their long-term effects on symptom reporting.

Meaning of the study: possible mechanisms and implications for clinicians

Medically unexplained or – as some authors prefer to call them – functional symptoms are not uncommon in patients attending general medical facilities, with prevalence estimates being as high as 12%. These symptoms often involve vague psychological symptoms (e.g. feelings of fatigue, depression, anxiety, tension) and the diagnostic label that is given to this constellation may depend on the medical experts that patients consult. Some authors have noted that there is a substantial overlap between labels such as chronic fatigue, fibromyalgia, irritable bowel syndrome and chronic pain.[1,2,19] There are good reasons to assume that iatrogenic factors may exacerbate symptom intensity in these patient groups. More specifically, expressing concern about the possibility of an underlying disease and, related to this, excessive investigation and attending patient support groups may all contribute to symptom escalation. What these interventions have in common is that they convey the message to the patient that his or her symptoms might be more intense and severe than he/she thinks they are. Our study suggests that blindness to unintended misinformation about the severity of the symptoms may underlie escalation of symptoms. Some experts have pointed out that appropriate interview techniques might reduce the risk of iatrogenic damage in these patients groups.[1] We agree and

would add that an appropriate interview style is one that refrains from exploring the whole spectrum of possible symptoms.

Unanswered questions and future research

Our study demonstrates that symptom reporting is susceptible to misinformation when this misinformation is not recognized as such. However, our study is silent as to whether this effect reflects altered symptom perception or reporting bias. Blind and non-blind participants did not differ in their social desirability scores, indicating that an account in social psychological terms (e.g. conformity) is not very likely. The same is true for our finding that at one-week follow-up, blind participants tended to give higher intensity ratings for target than for control symptoms, even though these follow-up data were obtained by email and so the demands that might be created by face-to-face contact with the researchers were absent. Again, this argues against a purely social psychological explanation. Taken together, our findings are best explained by assuming that misinformation biased blind participants' perception of ambiguous symptoms. Nevertheless, the issue of whether misinformation effects on symptom intensity scores reflect reporting or perception warrants further study. But even if symptom blindness would only reflect a reporting bias phenomenon, it would be relevant to clinical practice and research, because therapy outcome, particularly in the domain of psychiatry and clinical psychology, is often based on self-reports of symptoms.[20]

References

1 Page LA, Wessely S. Medically unexplained symptoms: Exacerbating factors in the doctor-patient encounter. *J R Soc Med* 2003;**96**:223–7

2 Hamilton WT, Gallagher AM, Thomas JM, White PD. The prognosis of different fatigue diagnostic labels. A longitudinal survey. *Fam Prac* 2005;**22**:383–8

3 Loftus EF. Our changeable memories: Legal and practical implications. *Nat Rev Neurosci* 2003;**4**:231–4

4 Johansson P, Hall L, Sikström S, Olsson A. Failure to detect mismatches between intention and outcome in a simple decision task. *Science* 2005;**310**:116–19

5 Hall L, Johansson P, Tärning B, Sikström S, Deutgen T. Magic at the marketplace: Choice blindness for the taste of jam and the smell of tea. *Cognition* 2010;**117**:54–61

6 Sharot T, Velsaques CM, Dolan RJ. Do decisions shape preference? Evidence from blind choice. *Psychol Sci* 2010;**21**:1231–5

7 Merckelbach H, Jelicic M, Pieters M. The residual effect of feigning: How intentional faking may evolve into a less conscious form of symptom reporting. *J Clin Exp Neuropsyc* 2011;**33**:131–9

8 Derogatis LR, Lipman RS, Covi L. SCL-90: An outpatient psychiatric rating scale: preliminary report. *Psychopharmacol Bull* 1973;**9**:13–28

9 Zack M, Toneatto T, Streiner DL. The SCL-90 factor structure in comorbid substance abusers. *J Subst Abuse* 1998;**10**:85–101

10 Crowne DP, Marlowe DA. A new scale of social desirability independent of psychopathology. *J Consul Psychol* 1960;**4**:349–54

11 Gudjonsson GH. Suggestibility, intelligence, memory recall, and personality: An experimental study. *Brit J Psychiat* 1983;**142**:35–7

12 Barksy AJ. The validity of bodily symptoms in medical outpatients. In: Stone AA, Turkkan JS, Bachrach CA, Jobe JB, eds. *The science of self-report: Implications for research and practice*. Mahwah, NJ: Erlbaum, 2000:339–61

13 Bernstein DM, Laney C, Morris EK, Loftus EF. False beliefs about fattening foods can have healthyconsequences. *P Natl Acad Sci USA* 2005;**102**:13724–31

14 Rietveld S, Kolk AM, Prins PJM, Colland VT. The influence of respiratory sounds on breathlessness in children with asthma: A symptom-perception approach. *Health Psychol* 1997;**6**:547–53

15 Ehlers A, Margraf J, Roth W, Taylor CB, Birbaumer N. Anxiety induced by false heart rate feedback in patients with panic disorder. *Behav Res Ther* 1988;**26**:1–11

16 Frenda SJ, Nichols RM, Loftus EF. Current issues and advances in misinformation research. *Curr Dir Psychol Sci* 2011;**20**:20–3

17 Mazzoni G, Memon A. Imagination can create false memories. *Psychol Sci* 2003;**14**:186–8

18 Merckelbach H, Jelicic M, Jonker C. Planting a misdiagnosis of Alzheimer's disease in a person's mind. *Acta Neuropsychiatr* (in press)

19 Mayou R, Farmer A. Functional somaticsymptoms syndromes. *BMJ* 2002;**325**:265–8

20 Rogler LH, Mroczek DK, Fellows MMA, Loftus STMA. The neglect of response bias in mental health research. *J Nerv Ment Dis* 2001;**189**:182–7