Original Article

The Influence of Anchoring on Pain Judgment

Paolo Riva, MS, Patrice Rusconi, MS, Lorenzo Montali, PhD,
and Paolo Cherubini, PhD
Department of Psychology, University of Milano-Bicocca, Milan, Italy

Abstract

Context. Research on decision making suggests that a wide range of spontaneous processes may influence medical judgment.

Objectives. We considered an easily accessible strategy, anchoring and insufficient adjustment, which might contribute to health care professionals' miscalibration of patients' pain.

Methods. A sample ($n = 423$) of physicians, nurses, medical students, and nursing students participated in a computerized task that showed 16 vignettes featuring fictitious patients reporting headache. In the experimental condition, participants were asked to evaluate the severity of the patient's pain before and after knowing the patient's rating. In the control condition, participants were shown all information about the patient at the same time and were required to make judgments in a single stage.

Results. When participants could express an initial impression before knowing the patient's rating, they fully anchored to their initial impressions in almost half of the responses. Moreover, even among those who revised their initial impression, the extent of the revision was often insufficient. Greater anchoring was associated with patients' ratings that were higher than the participants' initial impression. Finally, we provided evidence that anchoring increased pain miscalibration. We discuss our findings in terms of their contribution to the understanding of the cognitive processes involved in pain assessment.

Conclusion. When estimating patients' pain intensity, observers are driven by anchoring, a rule of thumb that might have pernicious consequences in terms of unwarranted overreliance on initial impressions and insufficient revision in light of relevant disconfirming evidence. Taking this heuristic into account might foster accurate pain assessment and treatment. J Pain Symptom Manage 2011;42:265–277. *© 2011 U.S. Cancer Pain Relief Committee. Published by Elsevier Inc. All rights reserved.*

Key Words
Decision making, anchoring effect, initial impressions, pain miscalibration, credibility attribution, patient's rating

Address correspondence to: Paolo Riva, MS, Department of Psychology, University of Milano-Bicocca, 1 Piazza Ateneo Nuovo, Milan 20126, Italy. E-mail: paolo.riva1@unimib.it

Accepted for publication: October 20, 2010.

0885-3924/$ - see front matter
doi:10.1016/j.jpainsymman.2010.10.264

Introduction

The notion that adequate pain management begins with an accurate assessment of others' pain is now widely recognized.[1] Yet, studies on pain judgment continue to show a systematic divergence of observers' estimates from the pain severity reported by patients.[2–4]

In the last four decades, several studies have identified the contribution of different variables to pain miscalibration. These include the presence of medical evidence,[5] the type of pain,[6] the observers' expertise,[4] the patients' demographic characteristics[4,7,8] and nonverbal expressions,[9] the availability of self-reports, and the presence of context cues.[10] Yet, researchers have called for a greater understanding of the cognitive processes that might intervene, along with the patients' and observers' characteristics and situational factors, in pain judgment.[11–13] Tait et al.[13] argue that observers often express their pain judgments under uncertainty about the symptoms that the patients report. This circumstance fosters the use of an intuitive/heuristic cognitive system (also known as "System 1"),[14–16] which is effortless, fast, and automatic and relies on prior expectations, and which ultimately might lead to a biased judgment of the patient's pain. Consistent with this emphasis on the cognitive underpinnings of pain judgment, we examined the influence of an easily accessible heuristic, labeled anchoring, that might contribute to miscalibration.

Previous studies assumed, either tacitly or explicitly, that the observers formulate their judgments in a unique stage when interacting with the patient.[4,17] We argue that it might prove useful to conceive of pain assessment as a sequential process, in which the information about the patient is acquired and integrated by the observers dynamically during their interaction with the patient.

In medical judgment—both for diagnostic and prognostic judgments—a tendency has been observed to generate preliminary hypotheses in the early stages of an interaction with a patient, and thereafter to revise them too slowly and insufficiently.[11,18–20] Similarly, it is plausible that in many instances (e.g., in physical examinations) the observers form an initial impression about the patient's pain (based on readily available clues) before asking the patient for an evaluation of his or her own pain level. Indeed, often the patient's rating is acquired after the perception of other characteristics that are readily accessible (e.g., the patient's gender, his or her facial expression). Because of its temporal primacy, an impression formed in the early stages of the assessment might outweigh an impression formed after learning other pieces of information.[21,22] Accordingly, observers might over-rely on initial impressions and subsequently discount relevant disconfirming evidence, a phenomenon known in psychology as "anchoring and insufficient adjustment"[23–26] or "focusing,"[27] which has previously been shown to affect medical judgment and decision making.[20,28–38]

The present study investigated the influence of anchoring on pain judgment by considering whether and to what extent an impression formed before knowing the patient's rating could serve as an "anchor" for the judgments made after learning the patient's rating. To address this issue, we devised two conditions (see Fig. 1): one in which participants were given the patient's rating immediately along with other clues (control condition), and one in which participants were required to make two judgments, before and after learning the patient's self-report (experimental condition). First, we hypothesized that the observers would exhibit a tendency to anchor to their initial impressions of pain judgment (Hypothesis 1). Secondly, based on previous studies that showed that both for chronic and acute pain, the higher the patient's rating, the greater the observers' discrepancy,[5,12,39,40] we tested the hypothesis of a different extent of anchoring depending on the patient's self-report. Specifically, we expected a greater extent of anchoring for the cases in which patients' ratings were higher than the participants' initial impression compared with those in which the patients' ratings were lower than the participants' initial impression (Hypothesis 2).

Methods

Participants

A total of 423 unpaid volunteers participated in the study during or after their shifts in several hospitals in three cities in Northern Italy. Two hundred twenty-three participants

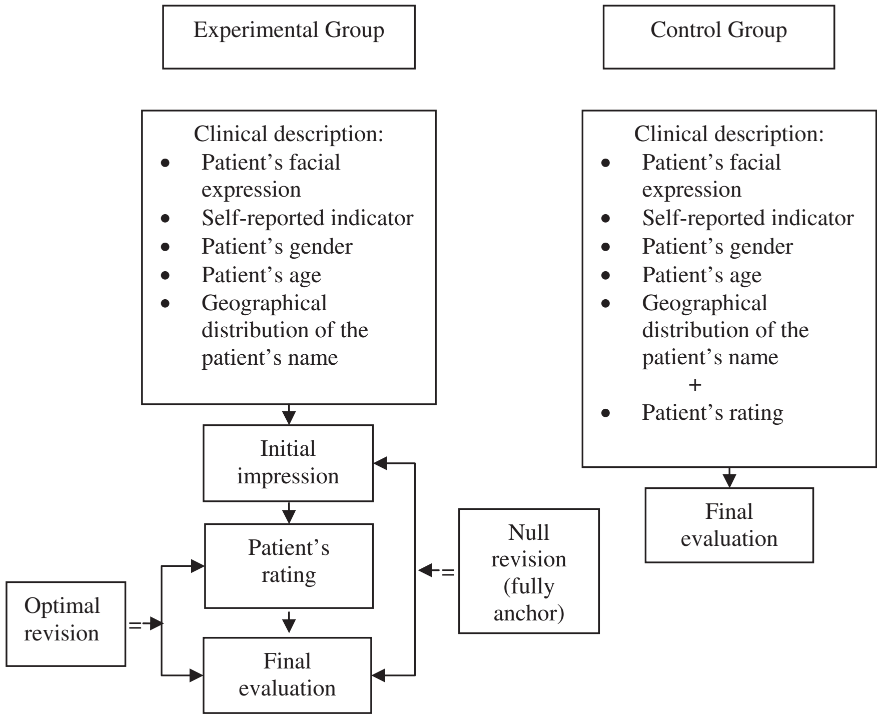

Fig. 1. Structure of the two conditions of the study.

were assigned to the experimental condition and the remaining 200 to the control condition. Because some of the information presented in the control condition was drawn from the experimental condition (see the Materials and Design section), we first ran participants in the experimental condition and then those in the control condition.

In the experimental group, the mean age of the 223 participants was 28 years (standard deviation [SD] = 9.0, range 18–67), and there were 128 females (57.4%). There were four groups of participants: 58 physicians, 40 nurses, 85 medical students, and 40 nursing students. The average length of employment was 9.4 years (SD = 9.7, range 2 months–35 years) for the physicians and 9.2 years (SD = 7.0, range 2 days–29 years) for the nurses. The medical students had been studying for 3.8 years (SD = 2.3, range 1–12 years) and the nursing students for 2.3 (SD = 0.8, range 1–4 years).

In the control group, the mean age of the 200 participants was 28 years (SD = 9.5, range 19–56), and there were 137 females (68.5%). As in the experimental condition, there were four groups of participants: 40 physicians, 40 nurses, 80 medical students, and 40 nursing students. The average length of employment was 8.4 years (SD = 9.0, range 8 days–28 years) for the physicians and 15.2 years (SD = 8.7, range 8 days–33 years) for the nurses. Medical students had been studying for 3.9 years (SD = 2.2, range 1–13 years) and nursing students for 1.4 years (SD = 0.5, range 1–3 years).

Materials and Design

Participants read text-based vignettes about hypothetical patients reporting headache, which were presented on computer screens (see Appendix).

To increase the face validity of the experiment, we included several clues that previous research had found to be relevant factors in pain judgment.[5,13,40,41] Specifically, we manipulated within subjects the patient's gender (female vs. male), facial expression (tense vs. relaxed), and a self-reported indicator that is consistent/inconsistent with headache pain (presence vs. absence of sensitivity to light in half of the vignettes, or to noise in the other

half), whereas between-group factors were the patient's age (old vs. young), and the geographical distribution of the patient's name (typically Southern Italian vs. typically Northern Italian names; this factor was taken into account in light of a previous exploratory study[42] that showed that health care professionals were influenced by the patients' places of origin) (Fig. 1). The inclusion of these variables in the vignettes was meant to provide participants with the minimal clues necessary to make a pain judgment. However, the effect of the manipulated variables was out of the scope of the current work (i.e., the influence of anchoring on pain judgment).

In the experimental condition, each participant read 16 vignettes about hypothetical patients who arrived at the emergency room for a wart and, during the visit, stated that they felt a headache (typically related to a primary headache, to prevent participants from focusing on possible causes instead of pain-related symptoms; see Appendix). We did not choose the pain to be evaluated (i.e., headache) as the main reason to go to the emergency department (ED), because the observers would have expected that a pain level so strong to induce a person to go to the ED would have been at least moderately high (e.g., greater than 6 on a 0–10 pain scale). Accordingly, to make equally plausible all the values greater than 0 in the pain scale, we chose a reason for going to the ED different from the actual object of the pain evaluation.

The task requirement was to rate the pain intensity (on a 0–10 numeric rating scale [NRS]) and to choose a therapy prescription (by selecting among increasing levels of intensity from 1 = no therapy, to 8 = an anti-inflammatory in association with opioids). The choice of treatment levels was based on the suggestions made by three independent physicians who advised us on the use of this range of treatment options. A close scrutiny of the responses revealed that although options 2 and 8 were very extreme, they were actually chosen 147 and 59 times, respectively.

After this task, participants were informed of the patient's self-report. Actually, the patient's estimate was a computer manipulation of the initial impression on pain judgment expressed by each participant. Specifically, the patient's rating was the participant's rating ±3 scale points on the 0–10 NRS: it was higher than the participants' ratings in eight of 16 vignettes, and it was lower than the participants' estimates in the remaining eight vignettes. The software randomly produced the sequence of higher and lower values by increasing the participants' initial impressions in half of the cases and decreasing them in the other half. For example, when participants initially rated a pain level of 5 and the software randomly assigned a decrease, the self-report of the fictitious patient was 2. By contrast, if the direction of the systematic variation was in terms of an increase, a patient's rating of 8 was shown. If participants provided a rating lower than 4 or greater than 7 and the direction of the systematic variation assigned by the computer was a decrease or an increase, respectively, the patient's rating was set at 1 (because 0 would have been interpreted as the absence of pain) or 10, respectively. The systematic variation of the patient's rating was required to address our main research goal; that is, the presence and the extent of anchoring to initial impressions in the face of contradicting evidence (e.g., the patient's rating). However, this is not an arbitrary, ecologically implausible manipulation because it has been shown that the correlation between patients' and observers' judgments is typically poor.[40,43] After learning of the patient's rating, participants were first asked whether they wanted to revise their initial impressions about the patient's pain intensity and then if they wanted to revise their therapy prescription. Participants who chose to revise were required to provide a new judgment of the pain level and of the therapy prescription.

The control condition was devised to determine the differential effect of knowing the patient's rating together with the other indicators (e.g., the facial expression) instead of acquiring it in a subsequent stage. Accordingly, participants were provided with the same vignettes used in the experimental condition; however, the patient's rating was presented on the computer screen along with the other features (i.e., not after having expressed an impression on pain level and therapy). For each scenario, we computed the patient's rating as the mean of the initial impressions of the pain intensity given by each of the four groups of participants in the experimental condition

to avoid providing the participants in the control group with patient ratings that were arbitrary. For instance, the experimental physicians' mean for the scenario in which a patient was an older female, with a typical Southern Italian name, displaying a relaxed face and reporting sensitivity to noise, was used for the respective scenario presented to the physicians in the control condition.

In both conditions, the instructions and a sample vignette were presented at the beginning of the session in a trial example so that participants could become familiar with the task. The order of presentation of the 16 vignettes was randomly set, and they were run on computers by means of E-Prime™ experimental software (Psychology Software Tools, Inc., Pittsburgh, PA).[44]

Procedure

Three laptops were used to administer the computerized vignettes and collect the responses. Participants were recruited by convenience sampling from several health care facilities and medical schools in three cities in Northern Italy. Telephonic and personal contacts were made with head physicians, charge nurses, and professors of medical schools to reach groups of potential participants. On consent to participate in the study, participants were informed about the general aim and the duration of the study. The participants were not paid to participate in the study. The study was conducted in accordance with the ethical standards of the American Psychological Association.

Each person was tested individually by the authors in studies that were available at the moment of testing. All participants in both the experimental and control groups completed the entire procedure. In the experimental condition, no participants reported having noticed that the patient's rating was a systematic manipulation of their initial judgment.

Statistical Analyses

In our first analysis (Analysis 1), we considered the frequency of null revisions of the initial judgments on both pain level and therapy prescription given by participants in the experimental condition. A null revision occurred when participants chose to fully anchor to their initial impression despite its discrepancy with the patient's rating computed by the software (Fig. 1). Thus, this analysis investigated the presence of the strongest tendency to anchor, namely the fully anchoring to the initial impression. To do this, we excluded from the analyses the cases in which participants' initial impressions were 1 or 10 and the patient's rating computed by the software was equal to the participant's rating; this occurred 95 times out of 3568 cases (2.66%).

In Analysis 2, we focused on the question of whether a tendency to anchor was present even for observers who chose to revise their initial impressions of pain level. Indeed, an optimal revision in light of the patient's rating would be a change of the initial pain estimate equal to ± 3, whereas any revision less than ± 3 scale points would reflect an influence of anchoring (Fig. 1). The exact patient's rating is not an arbitrary benchmark for assessing whether or not the revision of the pain level estimate is optimal. Indeed, in the literature on pain assessment, the Iafrati's criterion is widely used for determining whether the observer's estimate diverges or not from the patient's rating. According to this criterion, a rating falling outside the range of ± 1 is considered miscalibrated.[10,52] To investigate this, we focused on the extent of the participants' revisions of pain level. In this and in the following analyses, we did not consider the extent of revisions of the therapy prescriptions because there is not an optimal benchmark in treatment such as the patient's rating is for the pain level. Thus, we first selected the cases in which participants chose to revise their initial impressions of pain severity. This procedure eliminated the cases of nonrevision, which were 47.51% of the total responses (1650 of 3473 cases). Then, we excluded from this analysis the cases in which participants gave initial extreme ratings, and thus the randomized addition/subtraction of 3 could exceed the 0−10 scale. In those cases, the algorithm could not compute a patient's rating that was exactly 3 scale points distant from the initial participant's rating. Hence, the latter procedure assured that the difference between the initial impression and the patient's rating computed by the software was held constant. This operation eliminated an additional 5.44% of the total cases, leaving a total of 1634 of 3473 cases to analyze. The

dependent variable was the degree of revision of the initial pain rating.

Analysis 3 investigated the mean absolute value of the difference between the final evaluation and the initial impression provided by participants (Fig. 1): the greater the difference, the less the extent of the anchoring. The largest magnitude of anchoring occurred if the participants decided not to revise their initial impression and thus the difference between the final and the initial judgment was null. This analysis was performed through a paired *t*-test comparing the difference in estimates when the patient's rating computed by the computer was higher vs. when it was lower than the participant's initial impression. We excluded from the analysis the 95 cases in which the participants could not revise their initial pain estimates because their initial ratings were either "1" or "10" and the algorithm produced the same value. The remaining cases were then analyzed across participants (i.e., 223 cases).

Finally, after investigating for the presence of the anchoring effect in pain judgment, we analyzed the consequences of anchoring on observers' miscalibration (Analysis 4). Accordingly, we ran an independent-samples *t*-test comparing the experimental vs. control conditions. The dependent variable was the absolute value of the mean miscalibration. Specifically, in the experimental condition, it was the absolute value of the mean difference between the participants' initial estimate and the patient's rating when participants chose to confirm their initial impression, or the mean difference between the final evaluation and the patient's rating when participants chose to revise their initial impression. In the control group, the miscalibration was computed as the absolute value of the mean difference between the unique pain judgment and the patient's rating. To avoid any confound because of the use of extreme scale points, we excluded all the cases (in the experimental condition) in which the algorithm did not compute a patient's rating of ± 3 with respect to the initial impressions provided by the observers. This left 2868 cases in the experimental condition, which, along with the 3200 cases of the control condition, were then aggregated across participants, giving rise to a total of 423 cases to analyze (note, however, that the analysis including all trials yielded highly consistent results).

Results

Null Revisions: Was There a Tendency to Fully Anchor to the Initial Impressions?

Table 1 summarizes, at a descriptive level, participants' responses in the experimental condition, specifically the mean initial and final estimates, and the mean proportions of their revisions of both pain and therapy levels.

First, we examined whether there was a tendency to fully anchor to the impression formed before knowing the patient's rating (Analysis 1). We hypothesized the presence of a tendency to anchor to the initial impressions (Hypothesis 1). Thus, in this analysis, we predicted the presence of a tendency not to revise the initial impressions in light of the patient's rating.

We compared the frequency of revisions/nonrevisions of the initial judgments through nonparametric tests, given that the dependent variable was, originally, categorical. A series of Wilcoxon signed-rank tests on the count of revisions vs. nonrevisions (range 0−16) showed that the number of revisions of the initial pain judgment ($Mdn = 8$) was not significantly different from the number of nonrevisions ($Mdn = 7$) ($Z = -1.30$, asymptotic $P = 0.19$, two-tailed). As to the therapy, it turned out that the number of nonrevisions of the initial prescription ($Mdn = 10$) was significantly higher than the number of revisions ($Mdn = 5$) ($Z = -7.85$, asymptotic $P = 0.001$, two-tailed, $r = -0.37$). The number of nonrevisions of the initial therapy prescriptions ($Mdn = 10$) was significantly higher than the number of

Table 1
Mean (SD) Initial and Final Estimates of Pain and Therapy Levels Provided by Participants in the Experimental Condition and Mean (SD) Proportions of Revisions of the Initial Judgments

	Pain Intensity Level	Therapy Level
Initial estimate	4.64 (1.09)	3.63 (1.24)
Final estimate[a]	4.61 (1.1)	3.71 (1.18)
Proportion of revisions	0.52 (0.29)	0.34 (0.25)

SD = standard deviation.
[a]We considered as "final" estimate either the revised judgment in light of the patient's rating or, when participants decided for a nonrevision, the first judgment.

nonrevisions of the initial pain judgments ($Mdn = 7$) ($Z = -9.42$, asymptotic $P = 0.001$, two-tailed, $r = -0.45$).

Further, there was a significant positive correlation between the number of nonrevisions of the initial pain level and the number of nonrevisions of therapy prescriptions, Spearman's $r_s = 0.58$, $n = 223$, $P = 0.001$, two-tailed, showing the association between the tendency to fully anchor to the initial impression of pain intensity and the decision not to revise the hypothesized treatment.

Therefore, supporting Hypothesis 1, we found a considerable tendency to fully anchor to the initial impression. This tendency was not driven by a small group of observers who chose not to revise their initial impressions many times, but was widespread across the observers. Indeed, 40.36% of participants did not revise their initial impressions in more than half of their responses (i.e., more than eight responses out of 16). The anchoring effect was even greater for therapy judgments. Indeed, 68.61% of observers did not revise their first therapy prescriptions in more than half of their responses.

Two Kruskal-Wallis tests on the number of nonrevisions revealed that there were no significant differences among the four groups of participants (i.e., physicians, nurses, medical students, nursing students) either when considering pain judgments, χ^2 (3, $n = 223$) = 3.09, $P = 0.38$ or when examining therapy prescriptions, χ^2 (3, $n = 223$) = 0.26, $P = 0.97$. Accordingly, and for the sake of concision, in the further analyses we no longer included the participant's group as a factor.

In our vignettes, we manipulated several variables: the patient's facial expression, the self-reported indicator, patient's gender, and age, and the geographical distribution of his or her name (Fig. 1). This was aimed to increase the plausibility of our scenarios, but their effects on participants' judgments were not the main focus of the present study. However, we present for completeness the results of a preliminary analysis to check for any possible influence of those variables on anchoring.

A series of Wilcoxon signed-rank tests and of Mann-Whitney tests on the number of nonrevisions of pain judgments showed that none of the variables we considered significantly affected participants' decisions not to revise their initial estimates, $Z \leq -1.09$, asymptotic $P_s \geq 0.27$, two-tailed, except for the patient's facial expression, $Z = -4.76$, asymptotic $P = 0.001$, two-tailed, $r = -0.23$ (anchoring was greater when the face was relaxed, $Mdn = 4$, than when it was tense, $Mdn = 3$), and the geographical distribution of the patient's name, $Z = -2.04$, asymptotic $P = 0.041$, two-tailed, $r = -0.14$ (greater anchoring when the patient's name was typically Southern Italian, $Mdn = 8$, than when it was Northern Italian, $Mdn = 6$).

Similarly, as to the nonrevisions of the therapy, it turned out that none of the variables affected participants' judgments, $Z \leq -0.86$, asymptotic $P_s \geq 0.39$, two-tailed, except for the patient's facial expression, $Z = -4.01$, asymptotic $P = 0.001$, two-tailed, $r = -0.19$ (greater anchoring when the face was relaxed, $Mdn = 6$, than when it was tense, $Mdn = 5$), and the geographical distribution of the patient's name, $Z = -2.46$, asymptotic $P = 0.014$, two-tailed, $r = -0.16$ (again anchoring was more pronounced when the patient's name was typically Southern Italian, $Mdn = 11$, than when it was typically Northern Italian, $Mdn = 9$).

Suboptimal Revisions and Optimal Revisions: To What Extent Did Participants Adjust Their Initial Impressions of the Patient's Pain Level?

In Analysis 2, we considered to what extent participants adjusted their initial impressions of the patient's pain level when they decided for a revision.

Supporting Hypothesis 1, the analysis showed that only 13% of the participants who chose to revise the initial impression of pain intensity fully complied in a systematic way with the patient's self-report, optimally revising their impression by ±3 scale points. All of the remaining participants (87%) were influenced by their initial impressions, suboptimally revising them (e.g., their revisions were inferior to ±3 scale points).

Did the Direction of the Systematic Variation of the Patient's Rating Influence the Extent of Anchoring?

We then tested our hypothesis of a moderation effect of the patient's rating on the extent of anchoring (Hypothesis 2; Analysis 3). Supporting our prediction, we found that the patient's rating affected the extent of participants' anchoring

to the initial impressions of pain intensity. Specifically, participants were more prone to large revisions of their initial estimate when the patient's rating was lower than their initial impression ($M = 1.23$, $SD = 0.87$) than when it was higher ($M = 1.12$, $SD = 0.91$), $t(222) = 2.08$, $P = 0.039$, $d = 0.12$ (95% confidence interval [CI] 0.01, 0.23). In other words, as predicted, the magnitude of anchoring was greater when the algorithm produced a patient's rating that was higher than their initial impression of pain severity.

The Consequences of Anchoring: How Did Anchoring to the Initial Impression Influence the Participants' Miscalibration?

To investigate the consequences of anchoring on pain judgment, we analyzed whether participants' miscalibration varied across the experimental conditions. Miscalibration occurs when the observer—knowing the sufferers' pain rating—expresses a divergent pain estimate. A difference in the miscalibration between the experimental vs. control conditions could be interpreted as the consequence of anchoring. Indeed, the only difference between the two conditions is the possibility to form an impression of pain intensity in the absence of the patient's rating (experimental condition) or in the presence of the patient's rating along with the other indicators (control condition).

Analysis 4 yielded a significant difference between the two conditions, $t(410) = 12.30$, $P = 0.001$, $d = 1.21$ (95% CI 0.76, 1.05). Specifically, further supporting Hypothesis 1, the analysis showed that the discrepancy between the observers' final evaluations and the patients' rating was lower in the control condition ($M = 0.82$, $SD = 0.65$) than in the experimental condition ($M = 1.73$, $SD = 0.86$).

Discussion

Since the early 1970s, research has identified that patients and their health care providers diverge in their estimates of pain severity and that this discrepancy might affect pain treatment.[3] Yet, according to many clinical practice guidelines, the pain level recorded by the observers should conform to the patients' reports,[45–47] although absolute agreement remains an ideal because sufferers and observers do not share the same bases for estimation.[48] A wealth of studies has found several variables that consistently affect this divergence. In this regard, there remains a lack of empirical research on the cognitive underpinnings of pain miscalibration.[11–13,49]

The present study is an attempt to unveil a cognitive process that might intervene in pain assessment. We focused on the use of a readily accessible heuristic, namely anchoring and adjustment, whereby people rely excessively on an initial impression and interpret subsequent incoming evidence with reference to that impression. Specifically, we investigated whether and to what extent observers anchor to an initial impression (formed before knowing the patient's rating) after learning the patient's actual estimate of his or her pain level.

Tendency to Anchor to Initial Impressions

Our results support the hypothesized tendency to anchor to initial impressions in light of the patient's rating (Hypothesis 1). Indeed, participants tended to maintain the experimentally set discrepancy with the patient's self-report (in the cases in which the systematic variation of the patient's rating occurred). In particular, we found that in almost half of the responses (48%), the observers fully anchored to their initial impression of pain intensity. This percentage is relatively high, especially if we consider the recommendations of the aforementioned clinical practice guidelines.[45–47] Moreover, when considering the therapy prescription, this proportion was even greater (66%).

Further corroborating the hypothesis that observers would confirm their initial impressions after learning of the patient's rating, the extent of revisions was rarely optimal. Indeed, only 13% of the participants who chose to modify their initial pain evaluation in light of the patient's self-report systematically revised it by ± 3 scale points. The remaining 87% of our sample was influenced by the hypothesized bias toward the initial impression, insufficiently adjusting the pain judgment.

Although it was out of the main scope of the present contribution, we found that the tendency to over-rely on the initial impression could be moderated by some characteristics of the target. Specifically, anchoring and insufficient adjustment was greater (on both pain

evaluation and treatment decision) when the patient's face was relaxed (compared with when it was described as tense) and when s/he had a typical Southern Italian name (compared with a typical Northern one). Although the former effect might reflect the observers' inclination to disregard what the patient reports when s/he has a relaxed face (perhaps because this is somehow inconsistent with the patient's complaint of headache pain), the latter could represent a similar tendency to invalidate what the patient says about his or her pain if s/he belongs to a group stereotypically viewed as unreliable.[42] Future studies should work out in depth the issue of the target's and/or the observer's characteristics that can moderate the influence of anchoring on pain judgment.

The Influence of the Level of the Patient's Rating on the Extent of Anchoring

We then tested the hypothesis of a greater extent of anchoring when the patient's rating was higher than the initial impression compared with when it was lower (Hypothesis 2). Accordingly, we found that the tendency to anchor to the initial impressions was not homogeneous: When the algorithm produced a patient's rating higher than the initial impression provided by the participants, the magnitude of anchoring was greater than when the patient's rating was lower than the initial participants' impressions. These results cohere with previous studies that showed that, both for chronic and acute pain, the higher the patient's rating, the greater the observers' disbelief.[5,12,39,40] However, it should be noted that this effect was really very small, being the Cohen's d of 0.12, which is not even a small effect according to Cohen's benchmark[50] of 0.2. The statistical significance of this effect using a weak manipulation (e.g., text-based vignettes) suggests that it is worth examining how strong the effect might be with more naturalistic paradigms.[51]

The Role of Anchoring to Initial Impressions in Pain Miscalibration

We argue that the relevant consequence of anchoring and insufficient adjustment in pain judgment is an increase in pain miscalibration. Our data supported this view, showing that the discrepancy between the patient's self-report and the observers' judgments was greater when participants were given the opportunity to form an impression before knowing the patient's rating than when they knew the latter from the beginning of the evaluation process. We noted that this comparison between the miscalibration in the experimental condition and that in the control condition, which supported the presence of anchoring in the experimental condition, yielded a significant and large effect (Cohen's d of 1.21). Furthermore, our findings are not limited to the pain intensity judgment given that we found that anchoring to the initial impression of pain level not only affected pain assessment (in terms of a greater miscalibration), but also pain treatment, as indicated by the positive correlation between nonrevision of pain levels and therapy prescriptions.

Overall, we showed the usefulness of adopting a cognitive perspective to shed light on the discrepancy between the observer's and the sufferer's pain estimates. The cognitive psychology literature indicates that people can resort to both spontaneous and analytic processes when making judgments. Under conditions of symptom uncertainty or under high cognitive load, health care professionals are likely to rely on spontaneous processes in generating diagnoses and treatments.[13] Specifically, we showed that people are prone to use first judgments on patient's pain severity and treatment, made in the absence of the patient's rating, as a "natural starting point"[26] for their subsequent evaluation of the patient's estimate and pain treatment. This can reflect the observers' tendency to strive for internal consistency more than for the compatibility of their judgments with other held beliefs,[26] such as the knowledge of the subjective nature of the pain experience, which implies taking into account the patient's self-report. Hence, when assessing pain, health care professionals should consider the influence of anchoring, which is an economical shortcut but might lead to overlooking important clues.

Limitations of the Study

There are some critical study limitations to consider. First, we hypothesized that the features different from the patient's self-report could be acquired before inquiring about the patient's rating. We acknowledge that this order might represent one of the possible versions of

reality and that different combinations might occur. However, it seems reasonable that—in a face-to-face interaction—when health care providers elicit the pain estimate from a patient, they should have previously detected (with different degrees of consciousness) the patient's features that are readily accessible, such as the patient's age, gender, or facial expression.

Second, to investigate the pain judgment process by means of an experimental design, we chose to require a judgment in terms of a numerical scale. Hence, our findings can be reasonably applied to the cases in which the observers are able to form an explicit initial impression of the patient's pain. However, further studies should investigate whether the same rationale might apply to other circumstances, for example, when observers form more implicit and less conscious impressions.

Finally, we are aware of the limited external validity of this study because of the use of the computerized text-based vignettes, which, nonetheless, allowed the investigation of the impact of each manipulated variable while avoiding uncontrolled confounds.

Disclosures and Acknowledgments

This research was partly funded by a Programma di ricerca di Rilevante Interesse Nazionale 2008 grant to PC by the Italian government. The authors declare no conflicts of interest.

The authors wish to thank Giorgio Costantino, Cristina Dragone, Giovanni Di Pietro, Giorgio Fumagalli, Maura Lusignani, Stefania Negri, Alessia Riva, Dario Ronchi, Maurizio Sala, Loredana Ubino, and Gianluca Vago for their help in recruiting participants. They also thank Luciano Giromini, Marco Marelli, and Simona Sacchi for their comments on an early version of the article, Carlo Toneatto for his helpful technical support, and Sally Couchman for proofreading the article.

References

1. Craig KD. The social communication model of pain. Can Psychol 2009;50:22–32.

2. McCaffery M, Ferrell BR. Nurses' knowledge of pain assessment and management: how much progress have we made? J Pain Symptom Manage 1997;14:175–188.

3. Marks RM, Sachar EJ. Undertreatment of medical inpatients with narcotic analgesics. Ann Intern Med 1973;78:173–181.

4. Marquié L, Raufaste E, Lauque D, et al. Pain rating by patients and physicians: evidence of systematic pain miscalibration. Pain 2003;102:289–296.

5. Chibnall JT, Tait RC, Ross LR. The effects of medical evidence and pain intensity on medical student judgments of chronic pain patients. J Behav Med 1997;20:257–271.

6. Choiniere M, Melzack R, Girard N, Rondeau J, Paquin MJ. Comparison between patients' and nurses' assessment of pain and medication efficacy in severe burn injuries. Pain 1990;40:143–152.

7. Hirsh AT, George SZ, Robinson ME. Pain assessment and treatment disparities: a virtual human technology investigation. Pain 2009;143:106–113.

8. Weisse CS, Sorum PC, Dominguez RE. The influence of gender and race on physicians' pain management decisions. J Pain 2003;4:505–510.

9. Craig KD, Prkachin KM, Grunau RE. The facial expression of pain. In: Turk DC, Melzack R, eds. Handbook of pain assessment. New York: Guilford Press, 2001: 153–169.

10. Kappesser J, Williams AC, Prkachin KM. Testing two accounts of pain underestimation. Pain 2006;124:109–116.

11. Prkachin KM, Solomon PE, Ross J. Underestimation of pain by health-care providers: towards a model of the process of inferring pain in others. Can J Nurs Res 2007;39:88–106.

12. Solomon PE. Congruence between health professionals' and patients' pain ratings: a review of the literature. Scand J Caring Sci 2001;15:174–180.

13. Tait RC, Chibnall JT, Kalauokalani D. Provider judgments of patients in pain: seeking symptom certainty. Pain Med 2009;10:11–34.

14. Stanovich KE, West RF. Individual differences in reasoning: implications for the rationality debate? Behav Brain Sci 2000;23:645–665.

15. Evans JS. In two minds: dual-process accounts of reasoning. Trends Cogn Sci 2003;7:454–459.

16. Gilovich T, Griffin D, Kahneman D, eds. Heuristics and biases: The psychology of intuitive judgment. Cambridge, UK: Cambridge University Press, 2002.

17. Igier V, Mullet E, Sorum PC. How nursing personnel judge patients' pain. Eur J Pain 2007;11:542–550.

18. Elstein AS, Shulman LS, Sprafka SA. Medical problem solving: An analysis of clinical reasoning. Cambridge, MA: Harvard University Press, 1978.

19. Nickerson RS. Confirmation bias: a ubiquitous phenomenon in many guises. Rev Gen Psychol 1998;2:175–220.

20. Poses RM, Bekes C, Copare FJ, Scott WE. What difference do two days make? The inertia of

physicians' sequential prognostic judgments for critically iii patients. Med Decis Making 1990;1:6–14.

21. Lingle JH, Ostrom TM. Principles of memory and cognition in attitude formation. In: Petty RE, Ostrom TM, Brock TC, eds. Cognitive responses in persuasion. Hillsdale, NJ: Lawrence Erlbaum, 1981: 399–420.

22. Sherman SJ, Zehner KS, Johnson J, Hirt ER. Social explanation: the role of timing, set, and recall on subjective likelihood estimates. J Pers Soc Psychol 1983;44:1127–1143.

23. Arkes HR. Costs and benefits of judgment errors: implications for debiasing. Psychol Bull 1991;110:486–498.

24. Chapman GB, Johnson EJ. Incorporating the irrelevant: anchors in judgments of belief and value. In: Gilovich T, Griffin D, Kahneman D, eds. Heuristics and biases: The psychology of intuitive judgment. Cambridge, UK: Cambridge University Press, 2002: 120–138.

25. Einhorn HJ, Hogarth RM. Ambiguity and uncertainty in probabilistic inference. Psychol Rev 1985;92:433–461.

26. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. Science 1974;185: 1124–1131.

27. Del Missier F, Ferrante D, Costantini E. Focusing effects in predecisional information acquisition. Acta Psychol (Amst) 2007;125:155–174.

28. Bravata DM. Making medical decisions under uncertainty. Semin Med Pract 2000;3:6–14.

29. Brewer NT, Chapman GB, Schwartz JA, Bergus GR. The influence of irrelevant anchors on the judgments and choices of doctors and patients. Med Decis Making 2007;27:203–211.

30. Cioffi J. Decision making by emergency nurses in triage assessments. Accid Emerg Nurs 1998;6: 184–191.

31. Cioffi J. A study of the use of past experiences in clinical decision making in emergency situations. Int J Nurs Stud 2001;38:591–599.

32. Cioffi J, Markham R. Clinical decision-making by midwives: managing case complexity. J Adv Nurs 1997;25:265–272.

33. Estrada CA, Isen AM, Young MJ. Positive affect facilitates integration of information and decreases anchoring in reasoning among physicians. Organ Behav Hum Decis Proc 1997;72:117–135.

34. Fisher A, Fonteyn M. An exploration of an innovative methodological approach for examining nurses' heuristic use in clinical practice. Sch Inq Nurs Pract 1995;9:263–276.

35. Glascoe FP, Dworkin PH. Obstacles to effective developmental surveillance: errors in clinical reasoning. J Dev Behav Pediatr 1993;14:344–349.

36. Judd J. Strategies used by nurses for decision-making in the paediatric orthopaedic setting. J Orthop Nurs 2005;9:166–171.

37. Redelmeier DA. The cognitive psychology of missed diagnoses. Ann Intern Med 2005;142:115–120.

38. Trowbridge RL. Twelve tips for teaching avoidance of diagnostic errors. Med Teach 2008;30:496–500.

39. Marquié L, Raufaste E, Lauque D, et al. Further results about pain rating by patients and physicians: reply to Chibnall and Tait. Pain 2004;107:194–195.

40. Tait RC, Chibnall JT. Physician judgments of chronic pain patients. Soc Sci Med 1997;45: 1199–1205.

41. Chibnall JT, Tait RC. Observer perceptions of low back pain: effects of pain report and other contextual factors. J Appl Soc Psychol 1995;25:418–439.

42. Montali L, Colombo M, Riva P. Theories and practices in pain management: a research on doctors' representations. Psicologia della Salute 2009;1: 33–56.

43. Drayer RA, Henderson J, Reidenberg M. Barriers to better pain control in hospitalized patients. J Pain Symptom Manage 1999;17:434–440.

44. Schneider W, Eschman A, Zuccolotto A. E-Prime reference guide. Pittsburgh, PA: Psychology Software Tools, Inc., 2002.

45. Bell F. A review of the literature on the attitudes of nurses to acute pain management. J Orthop Nurs 2000;4:64–70.

46. McCaffery M, Ferrell BR, Pasero C. Nurses' personal opinions about patients' pain and their effect on recorded assessments and titration of opioid doses. Pain Manag Nurs 2000;1:79–87.

47. Thorn M. A survey of nurses' attitudes towards the assessment and control of postoperative pain. J Orthop Nurs 1997;1:30–38.

48. Kappesser J, Williams AC. Pain estimation: asking the right questions. Pain 2010;148:184–187.

49. Rusconi P, Riva P, Cherubini P, Montali L. Taking into account the observers' uncertainty: a graduated approach to the credibility of the patient's pain evaluation. J Behav Med 2010;33:60–71.

50. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum, 1988.

51. Prentice DA, Miller DT. When small effects are impressive. Psychol Bull 1992;112:160–164.

52. Iafrati NS. Pain on the burn unit: patient vs. nurse perceptions. J Burn Care Rehabil 1986;7: 413–416.

Appendix

Sample Vignettes From the Experimental and the Control Conditions

Experimental Group

Slide 1. You are going to be presented with some vignettes about patients who come into the ED for a problem on the back of the hand. Your diagnosis is that it is a wart. During the visit, the patients also report having a headache (with this expression we do not mean a headache deriving from injuries, cancer, ictus, etc., but a primary headache). You will be asked to express some judgments about the pain caused by the headaches. Please note that we will not ask you to make any diagnosis. Remember that there are no right or wrong responses, we are just interested in your point of view. To continue please press the space bar.

[The first case was an example].

Slide 2. [patient's name] is a [young 25—40/old 65—80]-year-old [man/woman]. When [s/he]says that [s/he] has a headache, you notice that [her/his] face appears [tense/relaxed]. The patient [does not report/reports] sensitivity to [light/noise]. Overall, which seems to you the most plausible level of [patient's name]'s pain? *Press a key from 0 to 10 (where 0 means "no pain" and 10 "maximum pain").*

Slide 3. Which therapy would you prescribe for [patient's name]'s pain?

1. No therapy
2. 100 mL physiological saline solution
3. 125 mg of paracetamol (in tablets)
4. 250 mg of paracetamol (in tablets)
5. 500 mg of paracetamol (in tablets)
6. 1 g of paracetamol (in tablets)
7. 1 g of paracetamol (into vein)
8. 1 g of paracetamol (in tablets) in association with 50 mg (in 100 mL physiological saline solution) of tramadol—opioid (into vein).

Slide 4. When you ask [her/him] to express [her/his] pain level on a scale from 0 to 10, [patient's name] says [the participants initial impression ±3]. After considering this information, do you want to revise your previous judgment about [patient's name]'s pain level? *Yes (press the key "Y")/No (press the key "N").*

Slide 5. (only if the participant's answer in slide 4 was "Yes"). Hence, which seems to you the most plausible level of [patient's name]'s pain? *Press a key from 0 to 10 (where 0 means "no pain" and 10 "maximum pain").*

Slide 6. Do you want to revise your previous decision about [patient's name]'s pain therapy? *Yes (press the key "Y")/No (press the key "N").*

Slide 7. (only if the participant's previous answer in slide 6 was "Yes"). Hence, for [patient's name]'s pain, which therapy would you prescribe? *[list of therapies].*

Slide 8. What was the [sex/age/geographical distribution of the name] of the patient that you have just evaluated? *Type the answer that you consider more appropriate. If you do not remember type "I do not remember."*

Control Group

Slide 1 was the same shown to the experimental group.

[The first case was an example].

Slide 2. [patient's name] is a [young 25—40/old 65—80]-year-old [man/woman]. When [s/he]says that [s/he] has a headache you notice that [her/his] face appears [tense/relaxed]. The patient [does not report/reports] sensitivity to [light/noise]. When you ask [her/him] to express [her/his] pain level on a scale from 0 to 10, [patient's name] says [mean of initial impressions given in the experimental group]. Overall, which seems to you the most plausible level of [patient's name]'s pain? *Press a key from 0 to 10 (where 0 means "no pain" and 10 "maximum pain").*

Slide 3. Which therapy would you prescribe for [patient's name]'s pain?

1. No therapy
2. 100 mL physiological saline solution
3. 125 mg of paracetamol (in tablets)
4. 250 mg of paracetamol (in tablets)
5. 500 mg of paracetamol (in tablets)
6. 1 g of paracetamol (in tablets)
7. 1 g of paracetamol (into vein)
8. 1 g of paracetamol (in tablets) in association with 50 mg (in 100 mL physiological saline solution) of tramadol—opioid (into vein)

Slide 4. What was the [sex/age/geographical distribution of the name] of the patient that you have just evaluated? *Type the answer that you consider more appropriate. If you do not remember type "I do not remember."*