

**EXPLANATORY NOTE FOR ELECTRONIC VERSION OF THIS REPORT:
All reproductions of the test advertisements are presented in Appendix A.**

I. Introduction

This report presents the primary findings of a large scale copy test project that was initiated to examine several issues relevant to the Commission's 1994 Enforcement Policy Statement on Food Advertising ("Statement").¹ That document provides guidance concerning the manner in which the Commission's approach to advertising relates to FDA's food labeling regulations that implemented the Nutrition Labeling and Education Act of 1990. These regulations define a series of standardized "nutrient content descriptors" that may be used in labeling to characterize the absolute and relative level of common nutrients in foods. The FDA rules also prescribe standards governing the use of health claims on labels.

The FTC Statement generally comports with the FDA standards, but provides for certain disclosures and qualifications that would allow advertisers to make health claims about products and diet-health relationships that could not appear in labeling. The copy test that is the subject of this memo was undertaken to help determine which types of disclosures and qualifications communicate most effectively truthful and nonmisleading information concerning the nutrient content and health attributes of food products and dietary supplements.

A. The FDA Regulations

As indicated, the FDA regulations define several nutrient content descriptors, such as "high" fiber, "low" fat, or "reduced" sodium. The regulations also specify a series of requirements that determine whether a food label may carry a health claim that links the nutrient profile of the food to a specific health benefit, such as a reduced risk of heart disease or cancer.

Specifically, the FDA regulations allow health claims only when (1) the labeled food does not contain more than 20 percent of the recommended daily value for total fat, saturated fat, sodium, or cholesterol, (2) the labeled food meets certain minimum or maximum nutrient level requirements specific to the health claim being made,² and (3) FDA has determined that there is "significant scientific agreement" supporting the claimed diet-disease relationship. To date, FDA

¹ The copy test was conducted jointly by the Bureaus of Consumer Protection and Economics. Principal participants from the Division of Advertising Practices of the Bureau of Consumer Protection were Theodore Hoppock, Michelle Rusk, and Marianne Watts. Representing the Bureau of Economics were Dennis Murphy of the Division of Consumer Protection and Pauline Ippolito of the Division of Economic Policy Analysis. Two outside consultants participated in the project. These were David W. Stewart, Professor of Marketing at the University of Southern California, and Kenneth L. Bernhardt, Professor of Marketing at Georgia State University.

² For example, a label cannot link a food's sodium level to a reduced risk of hypertension unless that food is low in sodium (140mg per serving or less). Labels cannot relate fat content to a heart disease benefit unless the food is low in fat (3g or less), saturated fat (1g or less), and cholesterol (20mg or less).

has approved health claims for eleven diet-health relationships. Claims concerning any other health relationship (including antioxidant vitamin supplements and cancer and trans fatty acids and heart disease) are currently prohibited.

B. The FTC Statement

The FTC Statement advises that the Commission will apply FDA's definitions for absolute and relative nutrient content descriptors when they are used in the same context in advertising. With respect to health claims, the Statement advises that advertised foods should meet the FDA regulations for the relevant health claim if the claim contains no qualifications concerning the extent of the benefit or the degree of scientific proof underlying the claim. The Statement does, however, provide for certain qualified health claims that could not appear in labeling.

First, the Statement does not automatically disqualify all health claims for products that contain more than 20 percent of the Daily Value for any of the four nutrients that can increase the risk of a diet-related disease. Instead, the Statement advises that it may be necessary under such circumstances for an advertisement to disclose the presence and possibly the significance of the risk-increasing ingredient if reasonable consumers might otherwise believe that the product did not present any health risk.³

Second, the Statement advises that, under certain limited instances, it may be possible to craft a truthful and nonmisleading comparative health claim in advertising for a product that does not meet the nutrient profile prescribed by FDA for the specific nutrient-disease area in question. If, for example, a product were substantially lower in total fat than other foods in its product category, but exceeded the 3-gram limit FDA stipulates for a heart health claim, an advertiser could still claim the product was heart healthy if the advertisement made clear that the benefit derived only from the substitution of the food for a significantly less healthful alternative.⁴

Third, the Statement does not necessarily limit health claims to the eleven diet-disease relationships for which FDA has determined there is significant scientific agreement. The

³ See Statement, pp. 22-23. Disclosure would almost certainly be necessary if the risk-increasing nutrient were closely related to the same diet-disease relationship that is the subject of the health claim. That is, a product low in fat and cholesterol, but high in sodium, could not make a heart health claim unless the advertisement also alerted consumers to the product's sodium content. It might also be necessary to explain the link between high-sodium diets and hypertension (and therefore heart disease).

The Statement also advises that, depending on context, such disclosure might be necessary even when the nutrient that is the subject of the health claim and the risk-increasing nutrient relate to different diet-disease relationships, as would be the case for a food high in calcium (reduced risk of osteoporosis) but also high in saturated fat (increased risk of heart disease). The governing consideration is whether, absent qualifying disclosures, the osteoporosis health claim would convey to reasonable consumers that the food is healthful in all respects.

⁴ Statement, p. 25.

Statement allows qualified health claims for other relationships if the advertisement clearly discloses the extent of the scientific support for the claim, and the claim is not inconsistent with the larger body of evidence in that area.⁵ Thus, the Statement leaves the door open to nonmisleading advertising about diet-disease relationships where existing evidence is very promising, but has not yet attained the level of certainty that would constitute significant scientific agreement.

II. Overview of The Food Copy Test

A. Components of the Test

The consumer research project reported in this memo contained three components, each targeted at one of the three previously discussed areas where the FTC Statement departs from the FDA labeling regulations. In all cases, the study first probed whether additional disclosures and qualifications were actually needed to correct misimpressions that consumers might otherwise receive from unqualified health claims. The research also attempted to determine which of several types of disclosures and qualifications would remedy most successfully any misimpressions that consumers do receive from unqualified claims.

1. Halo Effect

The first component of the consumer research, which can conveniently be referenced as the "Halo Effect" segment, concerned food products that are high in both a clearly beneficial nutrient (such as fiber) and a risk-increasing nutrient (such as sodium), and therefore are disqualified from health claims on labels. One major goal of this phase of the project was to determine whether nutrient content and, in particular, health claims about the beneficial nutrient would cast a halo over the entire product and convey to consumers that there were no problematic elements in the food's overall nutrient profile. The other purpose of this segment was to determine the effectiveness of a series of increasingly detailed and explicit disclosures concerning the level and significance of the risk-increasing ingredient in the food.

2. Substitution Claim

The second component of the project, which will be referenced as the "Substitution Claim" segment, concerned health products that are relatively low in a risk-increasing nutrient, but not sufficiently low to meet the relevant FDA definition of low. Such products could not make a claim in labeling that explained the health benefits of reducing the intake of the problem nutrient.

Specifically, this phase of the research tested advertisements for a brand of packaged sliced turkey that claimed to have less than half of the sodium of other leading brands, but which still did not qualify as a low sodium food. All of the ads stated that switching to the lower sodium brand could help consumers cut back on sodium and help reduce the risk of high blood pressure.

⁵ *Id.*, p. 20.

All but the control ad contained disclosures specifying more precisely the levels of sodium in the advertised brand and in other leading brands.

The first object of the inquiry was to determine whether consumers would misinterpret the sodium comparison and health claim in the control ad to mean that the advertised product was in an absolute sense a low sodium food. The second goal was to determine the capacity of the tested sodium content disclosures to correct any such misimpressions.

3. Level of Scientific Certainty

The final portion of the research involved health claims that have not yet achieved the level of scientific support needed for FDA approval in labeling. As discussed, the Statement advises that such claims may be allowable in advertising if the ad accurately conveys the level of scientific support for the claim and if there is not a larger body of scientific evidence that contradicts the claim.

This component of the consumer research tested health claims for two diet-disease relationships that have not been approved by FDA--antioxidant vitamins and cancer and trans fatty acids and heart disease. Two sets of ads were tested--one for an antioxidant vitamin supplement and one for a margarine that was free of trans fatty acids. The research attempted to determine what level of certainty consumers would attach to unqualified health claims for these products, and then to gauge the impact of a series of disclosures that qualified the degree of scientific proof supporting the claims.

B. Experimental Design

The copy test project employed the same experimental design for each of the three subject areas that were explored. Respondents were recruited in shopping malls in 12 geographically dispersed cities across the country.⁶ All respondents were screened to meet age and sex quotas and to determine whether they used products in the food category under examination.⁷

Consumers who passed the screening criteria were then escorted to an on-site testing facility, where they initially were shown three print ads. Two of these were "clutter ads" and one was the ad to be tested.⁸ The ads were then removed from sight and respondents were questioned to determine whether they remembered seeing the test ad. If so, the respondents were asked the very general open-ended question, "What do you think were the main points of the ad?" After responding, the respondents were allowed to view the test ad again, and the full questionnaire was administered (with the ad once more out of sight.) After the main interview concluded, all

⁶ The testing facilities were in Birmingham, Boston, Colorado Springs, Los Angeles, Orlando, Philadelphia, Seattle, Springfield, St. Louis, Tampa, Toledo, and Wichita.

⁷ The age limits for the study were 21-60, distributed approximately equally over 4 categories (21-34, 35-44, 45-54, and 55-60). Approximately 75 percent of the respondents were female.

⁸ The two clutter ads were control ads from other segments of the study.

respondents completed a questionnaire that measured their attitudes toward nutrition issues and their nutrition knowledge.

In general, one group of respondents in each segment of the research was shown a "tombstone" control ad for a fictitious product brand.⁹ This ad contained very general claims about such attributes as taste and/or price and made no explicit references to any health benefits from consuming the product. Results from the control ads presumably would reflect prior beliefs and knowledge about nutrition that consumers brought to the interviews, and therefore could provide a point of reference in gauging the impact of the test ads.

A second group of consumers viewed a very similar ad that contained some form of unqualified health claim.¹⁰ Remaining groups of respondents were exposed to one of several "remedy" ads that contained successively more detailed disclosures qualifying the overall health benefits of the products or the extent of scientific support for the health claim. In all, 27 ads were tested, with approximately 60 respondents exposed to each ad. A total of 1,623 respondents participated in the project.

As will be discussed in greater detail below, the results from the Halo Effect segment of the study revealed that the original version of the most detailed remedy disclosure failed to provide any additional clarification concerning the high level of risk-increasing nutrients in the two advertised products. Since this poor performance appeared to be due to a flaw in the experimental design, a small retest involving 120 respondents and two ads was conducted in February of 1997 using revised wording for this remedy disclosure. Thus over 1,700 respondents participated in either the initial or follow-up portions of the consumer research.

III. Detailed Discussion of Halo Effect Component

A. Hypotheses Tested

⁹ All but one of the ads used in the project were for a fictitious product brand. The decision to use fictitious rather than real-world brands was driven both by logistical and research design considerations. From a practical standpoint, we did not have sufficient resources or expertise to produce professional ads for real products or alter existing ads convincingly. From a theoretical standpoint, we were also concerned that consumers' prior beliefs concerning specific brands might color the results, notwithstanding the use of control ads. Our final strategy was to use a computer graphics program to produce ads that consumers might construe as preliminary versions of promotional material for a new product introduction. We recognized, however, that there were also risks to this approach, in that consumers might discount the credibility of the fictitious ads and participate less seriously in the interview process. As a check on this possibility, we included in the Strength of Science component one actual print ad for a real brand of vitamin supplements to help determine whether there were systematic differences in consumer reactions to the real and fictitious ads.

¹⁰ In two cases, consumers also saw an intermediate ad with a "nutrient content" claim that referenced high levels of fiber or calcium, but did not link the nutrient to a specific health benefit.

There were two broad objectives in the Halo Effect segment. The first was to determine whether a health claim (or, to a lesser extent, a simple nutrient content claim) for a food that contains high levels of a beneficial and a risk-increasing nutrient will imply to consumers that there are no problematic elements in the advertised product's nutrient profile. More formally, the first portion of the Halo Effect research was designed to test one primary and one secondary hypothesis:

Hypothesis I: A health claim relating a nutrient level to a reduction in the risk of a diet-related disease will imply to a substantial number of consumers that the product contains no other nutrient at levels that will increase the risk of a diet-related disease.

Secondary Hypothesis Ia: The implication that a food is free of any risk-increasing nutrients will be stronger for a health claim than for a nutrient content claim.

The second major objective of this phase of the research was to determine what types of disclosures would succeed in alerting consumers to the amount and health significance of the risk-increasing nutrient, assuming any such disclosures were needed.¹¹

As discussed, the FDA labeling regulations prohibit health claims for the types of food products we would be testing. Nutrient content claims are, however, allowed in labeling if the claim is accompanied by a disclosure referring consumers to the nutrition facts panel for information on the risk-increasing nutrient and other nutrients. For example, a complying nutrient content claim on the label of a frozen dinner that was low in fat but high in sodium might appear as follows: "Low in Fat. See back panel for information about sodium and other nutrients."

We did not consider the FDA disclosure appropriate for advertising, since, among other reasons, consumers viewing an ad could not conveniently consult the nutrition facts panel to determine the nutrient profile of the product. We therefore based most of our test disclosures on the information that consumers would find if they could in fact view the nutrition facts panel in conjunction with the advertisement. Our disclosures generally conformed to metrics and abbreviations specified by FDA. The first disclosure tested was simply the absolute amount of the problem nutrient stated in grams or milligrams ("absolute disclosure"). The second test disclosure added a statement of the absolute amount of the problem nutrient expressed as a percentage of the Daily Value that FDA specifies for that nutrient ("relative disclosure").¹²

¹¹ To economize on resources, the remedy disclosures were only applied to ads carrying a full health claim rather than a simple nutrient content claim. Since it was hypothesized that any halo effect from a health claim would be stronger than that associated with a simple nutrient content claim, a disclosure that proved effective in a health claim ad presumably would also have remedied misleading inferences from a nutrient content claim.

¹² Our test disclosures referenced FDA's "Daily Value" as the "Maximum Daily Value," since FTC orders in food advertising cases have specified this variation on the FDA term in an attempt
(continued...)

The final disclosure (the "strong disclosure") initially was designed to test whether consumers would be better able to interpret the absolute and relative nutrient content information if additional information (not found on the Nutrition Facts Panel) were provided on the health significance of high levels of the problem ingredient. The Statement advises that such a disclosure may frequently be needed to prevent deceptive inferences when health claims are made for products with high levels of a risk-increasing nutrient.¹³ We therefore added a health advisory to the relative disclosure of the form: "Diets high in (the problem nutrient) may increase the risk of (the associated disease).

As indicated, this disclosure performed no better than the absolute or relative disclosures in alerting consumers to the high levels of risk-increasing nutrients in the test products and, in some instances, increased the number of consumers responding that the advertised products were low in these nutrients. After further analysis of the results and a reassessment of the experimental design, we realized that consumers were likely to understand the significance of the tested health advisory only if they also interpreted the absolute and relative nutrient content disclosures correctly in the first instance and already knew that the tested foods were high in a problem nutrient. If consumers did not realize from the nutrient content metrics that the test food was high in the relevant ingredient, a statement that warned about diets high in that ingredient could be--and was--interpreted to mean that the advertised food was particularly healthy because it did not contribute to an unhealthy diet.

To correct this design flaw, the absolute and relative nutrient content disclosures were removed from the strong disclosure and replaced with a simple statement warning that the advertised food was high in the risk-increasing ingredient. This direct alert was followed by the original advisory concerning the health risks of diets high in the problem nutrient.¹⁴ The revamped strong disclosure was then retested in three of the original twelve shopping malls under conditions that replicated the initial research.¹⁵

¹²(...continued)

to communicate more clearly that the Daily Value for risk-increasing nutrients is the upper limit on the desired daily intake.

¹³ Statement, p. 22.

¹⁴ Had sufficient resources been available, we also would have tested a truncated version of the strong disclosure that deleted the dietary warning concerning heart disease. This experimental design would have allowed us to isolate how much of the superior performance of the revised strong disclosure was due to a clearer communication of the level of sodium and saturated fat in the advertised products, and how much was attributable to the explanation of the health significance of the problem nutrients. The decision to devote the limited funds to a test of the complete revised disclosure was driven by our negative experience with the original remedy disclosures. Given that all of the initial attempts to find an effective remedy had failed, we wanted to concentrate the follow-up research on the remedy that appeared most likely to succeed.

¹⁵ The three malls were located in Boston, St. Louis, and Los Angeles.

In summary, the remedy section of the halo effect segment of our research tested four subordinate hypotheses:

Hypothesis II: If a health claim is advertised for a food with a high level of a risk-increasing nutrient:

- a. Disclosing the absolute amount of the risk-increasing nutrient in conjunction with the health claim will on net increase the number of consumers who understand that the product is high in a risk-increasing nutrient.
- b. An additional disclosure that expresses the amount of the risk-increasing nutrient as a percentage of the Maximum Daily Value will on net increase the number of consumers who understand that the advertised product is high in a risk-increasing nutrient.
- c. An additional disclosure that explains the health significance of the risk-increasing nutrient will on net increase the number of consumers who understand that the advertised product is high in a risk-increasing nutrient.
- d. Consumers will be most likely to understand that the advertised product is high in a risk-increasing nutrient if the quantitative disclosures used to test hypothesis II(c) above are replaced with an express warning that the advertised product is high in the risk-increasing nutrient.

B. Food Products and Advertisements Tested

Two fictitious food products were chosen as subjects for the Halo Effect component. The first was a canned minestrone soup ("Rich and Hearty") that was high in fiber and sodium. The specific sodium level was 840 mg, which is 35 percent of the maximum suggested daily value and therefore well over the risk-increasing threshold of 20 percent. The tested health claim for this soup linked diets high in fiber to a reduction in the risk of some forms of cancer. The strongest remedy disclosure that was tested connected diets high in sodium to an increased risk of high blood pressure and heart disease.

The second test was for packaged swiss cheese slices ("Matterhorn") that were high in calcium, but contained 7 grams of saturated fat (which also is 35 percent of the daily value). The claimed health benefit was a reduced risk of osteoporosis, while the strongest remedy disclosure was for an increased risk of heart disease.

Six ads--1 control ad and 5 test ads--were developed for both the soup and cheese products. These ads are reproduced in Appendix A. Tables I and II present, respectively, the differences in the main text portion of each of the soup and cheese ads designed to test the hypotheses described above. In most cases, successive quotations in the tables include only additional information not found in the previous ad.

The disclosures in the three remedy ads were designed to meet the criteria the Statement specifies for clarity and prominence.¹⁶ The disclosures appear in a box immediately below the health claim in a font size approximately half that of the main text.

C. Questionnaire Design

The questionnaires administered to viewers of the soup and cheese ads were virtually identical and conformed to the "funneling" format customary in consumer survey research of this type. The questioning began in a very general, completely open ended manner and gradually focused more narrowly on the perceived healthiness of the products and, finally, on the perceived sodium or saturated fat content.

The questionnaire ended with demographic questions concerning income and education level. The ads and survey instruments for the halo effect component of the research were pre-tested twice--once in Baltimore and, following revisions, a second time in Philadelphia. The full questionnaire for the soup respondents is presented in Appendix B.

¹⁶ Statement, p. 20, note 82.

Table I
Principal Differences in Text of Soup Ads

Nutrient Content Claim

And Rich and Hearty is High in Fiber, Too!

Health Claim

That's Good News! Eating Diets That Are Low in Fat and High in Fiber May Reduce the Risk of Some Forms of Cancer.

Absolute Disclosure

Sodium Per Serving: 840 mg

Relative Disclosure

Sodium Per Serving: 840 mg
% of Maximum Daily Value: 35%

Strong Disclosure

Sodium Per Serving: 840 mg
% of Maximum Daily Value: 35%
Diets high in sodium may increase the risk of high blood pressure and heart disease

Revised Strong Disclosure

Rich and Hearty is high in sodium.
Diets high in sodium may increase the risk of high blood pressure and heart disease.

Table II

Principal Differences in Text of Cheese Ads

Nutrient Content Claim

And Matterhorn is High in Calcium, Too!

Health Claim

That's Good News! Eating Foods
That Are High in Calcium May
Reduce the Risk of Osteoporosis.

Absolute Disclosure

Saturated Fat Per Serving: 7g

Relative Disclosure

Saturated Fat Per Serving: 7g
% of Maximum Daily Value: 35%

Strong Disclosure

Saturated Fat Per Serving: 7g
% of Maximum Daily Value: 35%
Diets high in saturated fat may increase
the risk of heart disease.

Revised Strong Disclosure

Matterhorn is high in saturated fat.
Diets high in saturated fat may increase
the risk of heart disease.

D. Copy Test Results

1. Overview

Stated succinctly, the results from both the soup and cheese test advertisements failed to support most of the hypothesized outcomes. Little evidence developed that was consistent with the primary halo effect theory. The positive nutrient content and health claims concerning fiber had only a small impact on respondents' perception of the overall healthiness of the advertised soup, and the calcium information had no effect in the corresponding cheese cells. More important, these positive messages did not appear to alter respondents' perception of the level of sodium or saturated fat in the advertised products, although low response rates in key test cells prevented firm conclusions on this issue.

Further, the three remedy disclosures originally tested were misinterpreted by many consumers, and, on net, did not increase the number of respondents who understood the high level of sodium or saturated fat present in the advertised soup or cheese. Respondents were particularly confused by the dietary advisory in the original strong disclosure. Many viewers apparently misconstrued the warning about the health risks of diets high in sodium or saturated fat as a positive indication that the advertised products were low in these nutrients and therefore did not present the described dietary risk. Taken together, the results from these original remedy cells indicate that, at the time of our test, consumers had a poor understanding of metric nutrient content measures, and that, when in doubt, consumers will tend to attach a positive connotation to information disclosures in advertising.

The revised strong disclosure provided an exception to the otherwise poor performance of the remedies. Virtually all respondents in the cheese cell understood the more direct verbal message, and the corresponding sodium warning avoided most (though not all) of the confusion associated with the original quantitative strong sodium disclosure. Due to resource constraints, it was not possible to determine how much of the effectiveness of the revised disclosure could have been achieved without the heart disease warning contained in the second sentence. Future research might explore productively whether a shorter and less burdensome alert to high levels of a problem nutrient would prove as successful as the more complete warning tested here.

2. Results in Detail

Given that the ad format and questionnaires for the soup and cheese products were essentially identical, the two sets of results will be discussed together.¹⁷ Results for the original strong disclosure cells are provided only for the key close-ended question, since the revised disclosure was conceptually superior and clearly outperformed the initial disclosure in the field.

¹⁷ A complete question-by-question analysis of the results is available from the principal author. The discussion that follows generally omits or treats in cursory fashion questions that failed to yield new insights concerning the principal issues under investigation.

a. Main Points of the Ad, First Exposure

The unaided responses to the first open-ended question, which was asked after the first exposure to the test ad (and the clutter ads), give an initial read on what respondents found to be the primary messages. The most interesting issues promised to be the proportion of respondents that mentioned nutrient content or health messages, and whether respondents in the remedy cells noticed and interpreted correctly the various sodium or saturated fat disclosures.

With the exception of the tombstone control ads, which contained no nutrient information, approximately half of the respondents seeing the various soup or cheese ads mentioned nutrients, usually fiber or calcium. With respect to risk-increasing nutrients, 17 percent of respondents in the absolute and relative disclosure cells, and 38 percent in the strong disclosure cell, made some reference to sodium. For cheese, about one-third of the respondents in the first two remedy cells mentioned fat or saturated fat. This figure rose to 45 percent in the strong disclosure cell.

A closer examination of the sodium and saturated fat responses yields early indications that the absolute and relative remedy disclosures did not perform as hypothesized. No respondents seeing the absolute disclosure cheese ad reported that the cheese was high in fat or saturated fat, and only 5 percent in the relative disclosure cell reported such a message. The outcome was virtually identical in the corresponding soup cells. Results for the strong disclosure ads were much more encouraging, however. Almost one third of the soup respondents recalled high sodium as a main message, and over 40 percent of the cheese respondents recalled as a main point that the advertised product was high in fat or saturated fat.¹⁸

Finally, from 40-50 percent of respondents seeing the health claim and remedy disclosure ads for soup and cheese reported some form of health message as a main point. (The corresponding figures for the control and nutrient content ads were 15 and 8 percent, respectively.) These health responses tended to be very general, such as the soup "is good for you" or "is healthy."

b. Main Points, Second Exposure

After respondents had given their first impressions of the soup and cheese ads, they were allowed to read the ads again, after which they were asked: "Now what do you think the main points are." With one exception discussed below, the pattern of responses to this second query changed very little over the first, although, not surprisingly, there was an across-the-board increase in the proportion of consumers reporting a given point. For example, from 63 percent (health claim) to 78 percent (relative disclosure) of respondents listed calcium content has a main point of the cheese ads, vs. 39-53 percent after the first exposure.

¹⁸ Most cheese respondents recalled correctly that the warning concerned saturated fat rather than total fat.

There were similar increases in the proportion of respondents mentioning a health claim after viewing the health claim and remedy disclosure ads a second time. Unlike the first exposure, however, most of these responses made specific reference to lowered risk of cancer or osteoporosis.

The second exposure did not change the pattern of responses to the three remedy disclosures appreciably. No respondents in the cheese absolute disclosure cell, and only 3 percent in the corresponding soup cell, volunteered as a main point that the products were high in any type of fat or in sodium. For cheese, 7 percent of respondents seeing the relative disclosure mentioned fat or cholesterol in a negative light. This figure rose substantially to 42 percent in the strong disclosure cell. For soup, negative mentions of sodium in the relative and strong disclosure cells were, respectively, 10 percent and 33 percent.¹⁹

c. Purchase Interest

The next question in the interview was intended to give a summary measure of respondent's overall evaluation of the advertised product. The question read:

Q3. "Based on the information in the ad, how interested would you be in buying the product?"

Respondents were shown a card with five possible answers, ranging from "not at all interested" to "extremely interested."

¹⁹ The relatively high playback of negative characterizations of sodium and saturated fat in the strong disclosure cells raises the issue of whether the stern warnings in these ads interfered in any way with communication of the positive messages concerning fiber, calcium, and their associated health benefits. There is no evidence that the strong disclosure ads were any less successful than the simple health claim ads in communicating positive nutrient content information. There were slightly more mentions of fiber and calcium in the soup and cheese strong disclosure cells than in the corresponding health claim cells.

There was, however, a modest falloff in the number of positive health messages reported by respondents seeing the strong disclosure ads. About half of the respondents in the soup health claim cell mentioned cancer reduction benefits as a main point of the ad. This proportion fell to 40% in the strong disclosure cell, although roughly equal reductions occurred in the other soup remedy cells. Similarly, the incidence of osteoporosis messages reported in the cheese health claim cell (57%) was 12 percentage points higher than in the strong disclosure cell. Again, however, the absolute and relative disclosure ads also produced fewer mentions of osteoporosis (51% and 34.4%), even though very few respondents in these cells cast the remedy disclosures in a negative light in the verbatims. Thus, it may simply be the greater quantity of information in the remedy ads, rather than their negative content, that is responsible for the decreased playback of specific health claims.

If the healthiness of a product were important to respondents, and if the ads successfully communicated nutrient and health information that respondents did not already know, we would expect the mean purchase interest rating for the nutrient content cell to be higher than for the control, and the rating for the health claim cell to increase over that of the nutrient content cell. The ratings should then decline steadily as greater detail is added to the remedy disclosures. We cannot predict a priori whether the ratings for the remedy cells would be lower than for the control. This would depend in part upon how respondents weighted the positive and negative nutrient attributes of the soup and cheese products.

The results did not fit the predicted pattern very well. For soup, the highest purchase interest was expressed by viewers of the control ad, which contained no health or nutrient information. Although the lowest purchase ratings were registered in the remedy cells, none was significantly below the health claim cell which, a priori, should have produced the highest rating. For cheese, all of the purchase interest ratings were statistically indistinguishable except, marginally, the revised strong disclosure rating, which was lower than the health claim rating at a .15 level of significance. Overall, the mean purchase interest ratings ranged from 2.19 (where 2.0 represents "not very interested") in the cheese revised disclosure cell to 3.13 (slightly above "somewhat interested") in the soup control cell.

Following the purchase interest question, respondents were asked two very general questions that were intended to direct attention to the nutrient profile health attributes of the featured foods. These questions were:

Q4. Did the ad say or imply anything about what is in the soup (swiss cheese)? What?

Q5. Did the ad say or imply that the soup (swiss cheese) would have any effect on your health? What?

Neither of these questions generated any new information that was not already provided by the responses to the initial open-ended questions concerning the perceived main points of the ads.

d. Ad Believability

Question 6 asked respondents, "How believable is the ad?" Such an inquiry is pro forma in consumer research of this type, and serves as a check on possible glitches in the wording and appearance of the test ads. In this case, the question also tested whether adding increasingly detailed and explicit nutrient content and health information would increase the credibility of an ad. Respondents were instructed to select one of 5 possible ratings, ranging from "not at all believable" to "extremely believable."

There were no large differences in the believability ratings for any of the ads. Although the control scored highest of any of the soup ads, the control cheese ad was the poorest performer. In any event, there are no statistically significant differences between any of the ads.

Overall, respondents found the cheese ads slightly more believable than the soup ads. In all cases, however, respondents scored the ads above the midpoint "somewhat believable" rating.

The mean score for all of the soup ads was 3.21; the corresponding cheese score was 3.36. Considering that the ads were computer mock-ups and that the products were fictitious, this is a fairly strong result. There were not very many derogatory comments about the appearance of the ads in any of the verbatims for any of the questions.

e. Overall Healthiness

Question 7 measured how the test ads affected respondents' appraisal of the overall healthiness of the soup and cheese products. The question asked:

Q7. Based on what the ad says or implies, how healthy would you rate this product?
Why?

The five response choices ranged from "very unhealthy" to "very healthy." Figures 3 and 4 illustrate the results.

Although there is not a great deal of variation in outcomes among the soup ads, the overall pattern does at least conform to a priori predictions. The highest rating of 4.05 is obtained in the health claim cell, after which the scores decline as more explicit remedies are disclosed about the problem nutrients. The final rating of 3.48 for the revised strong disclosure is below the health claim score by an amount that just misses statistical significance at the .10 level ($p=.11$), and is also the lowest recorded in any of the cells. This neat pattern all but collapses in the cheese results. The distribution of scores among the first five cells appears purely random. Again, however, the revised strong disclosure does perform as predicted, lowering the healthiness score by about a full point in comparison with most of the other cells. Even the smallest difference--that between the strong disclosure score of 2.76 and the control score of 3.57--is significant at the .05 level.

In their responses to the follow-up probe in Question 7, about 10% of participants openly misinterpreted the absolute and relative remedy disclosures in the soup and cheese ads, stating that the advertised foods were low in the problem nutrient. This proportion fell to 7% in the soup revised strong disclosure cell, and 3% in the equivalent cheese cell.²⁰

A higher proportion of respondents, however, volunteered explanations that contained correct interpretations of these remedies. For soup, 13% (relative) to 33% (revised strong) cited high sodium levels as a basis for their rating. Negative mentions were considerably higher among the cheese justifications. The advertised cheese was identified as high in saturated or total fat by 21% of respondents in the relative cell, 39% in the absolute cell, and 60% in the

²⁰ There were virtually no mentions of low sodium or fat in any of the other test or control cells.

FIGURE 3

HEALTHINESS OF ADVERTISED SOUP

MEAN RESPONSE

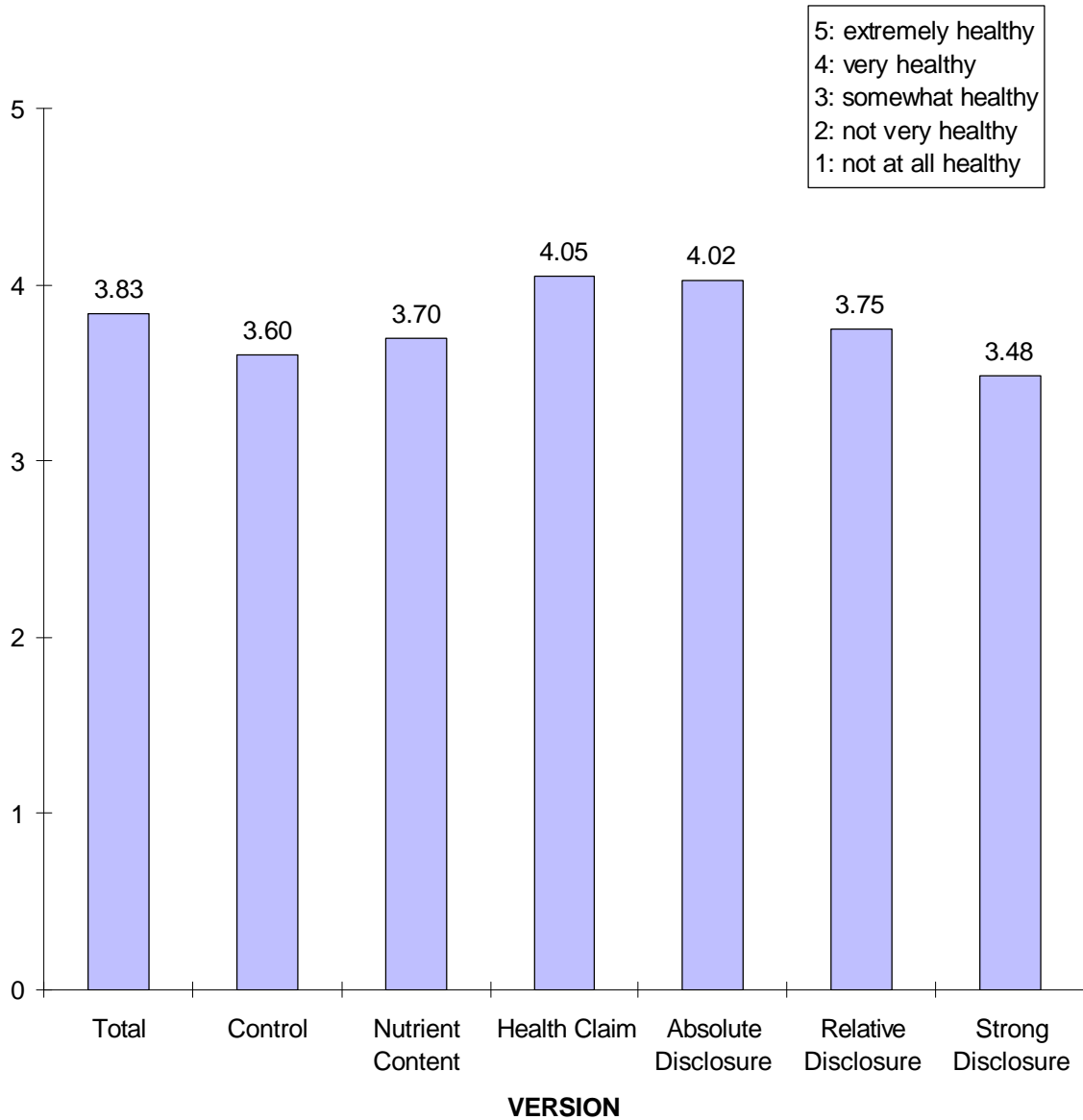
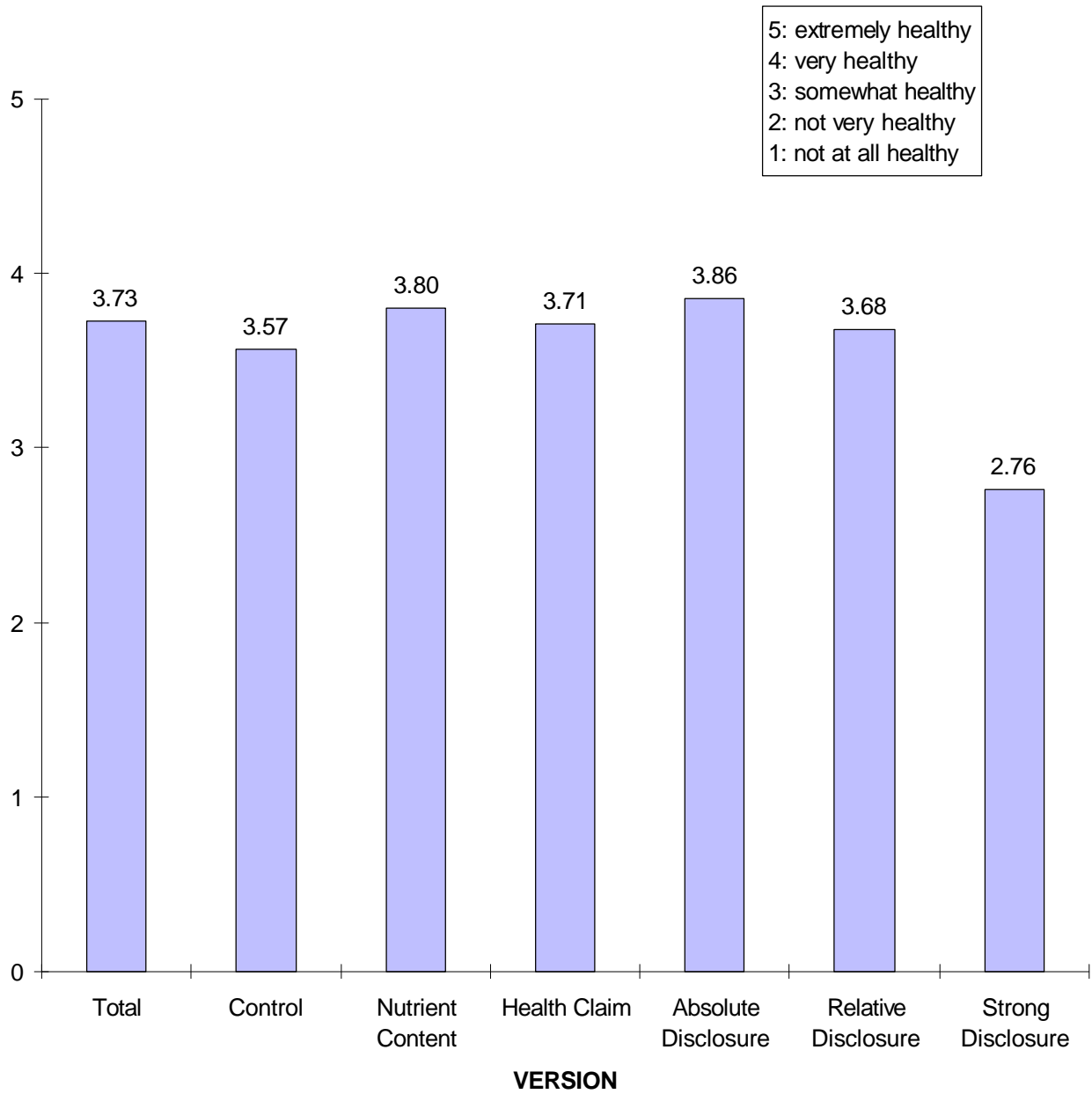


FIGURE 4

HEALTHINESS OF ADVERTISED CHEESE

MEAN RESPONSE



strong cell.²¹

f. Sodium and Saturated Fat Content

Question 8 provides the most focused inquiry into respondents' perceptions concerning the risk-increasing nutrients in the advertised products, and is therefore the most direct test of the hypothesized halo effect and hypotheses IIa-IId, which relate to the predicted efficacy of the alternative remedy disclosures. This question asked respondents to rate the sodium or saturated fat content of the foods along a 5-point scale labeled, "low," "somewhat low," "neither high nor low," "somewhat high," and "high." The average ratings for the various cells are reported in Figures 5 and 6.

These tables include results for the original strong remedy disclosure, which coupled the relative disclosure with a warning concerning risks of diets high in sodium or saturated fat. The revised strong disclosure, it will be recalled, replaced the quantitative information in the relative disclosure component with a simple statement that the food was high in the relevant problem nutrient.

The results for the first three cells in Figures 5 and 6--control, nutrient content, and health claim--provide evidence on the hypothesized halo effect. The results would support this theory if the mean sodium and saturated fat ratings were highest in the control cell, and then fell progressively in the nutrient content and health claim cells.

In the case of soup, the mean sodium rating for the control ad, which contained no health or nutrient content messages, is quite low at 2.43. Over 60% of respondents, however, stated that the ad did not say anything about sodium or that they did not know. The mean sodium rating then rises unexpectedly in the nutrient content cell. Again, however, only a minority of respondents estimated sodium content, and the difference is not statistically significant. The results for the health claim cell must also be interpreted with caution, since the ratings reflect the opinion of only about half of the respondents. The rating of 2.27 is lower than the control, and therefore consistent with Hypothesis I (halo effect), but the difference is not significant.

Figure 6 reveals much less variation in the cheese saturated fat ratings over the first three cells. The mean scores in the control, nutrient content, and health claim cells are within .06 points of each other. It should be noted, however, that response rates were even lower than in the equivalent soup cells, ranging from 45 percent in the control cell to only 33 percent in the health claim cell.

Taken together, the soup and cheese results provide no evidence of a halo effect from the health claims. The analysis is, however, clouded by low response rates, which greatly reduces the power of the test. It is possible that a significant difference in perceived sodium or saturated fat

²¹ The control and health claim ads also elicited an appreciable number of mentions of high fat or saturated fat in the probe to Question 7. Thirteen percent of respondents in these cells commented negatively on the fat content of the advertised cheese.

levels between the control and health claim ads might have emerged if a method had been found to elicit a rating by a majority of respondents in the control cells.²² It is not likely, however, that any such effect would have proven substantial, since the responses to Question 7 revealed no significant difference between the perceived healthiness of the advertised products in the control and health claim cells.

The results for the last four cells reported in Figures 5 and 6 reveal the impact of the remedy disclosures on consumer perceptions of the level of problem nutrients in the advertised products. For soup, the performance of the absolute disclosure is consistent with Hypothesis IIa, which predicted that disclosing the absolute quantity of a risk-increasing ingredient would improve consumer understanding of the amount of that nutrient in the advertised food. The mean sodium rating for the absolute cell is higher than that of the health claim, although the difference is significant only at the .14 level. The subsequent results for the relative disclosure cell and the original strong disclosure cell fail to support Hypotheses IIb and IIc, which predicted progressively higher sodium ratings for increasingly detailed and explicit disclosures. Consumer perception of sodium actually falls to 2.88 in the relative disclosure cell, and again to 2.68 in the original strong disclosure cell. (These differences are not significant.)

The final result reported in Figure 5 provides support for hypothesis II(d), which predicted that a verbal warning concerning the high level of sodium in the soup would prove more effective than the quantitative measures contained in the original strong disclosure. The mean sodium rating for the revised strong disclosure (3.75) is higher than any other cell, and is significantly higher than the rating recorded in the health claim cell ($p=.004$) and the nutrient content cell ($p=.04$). This disclosure cannot, however, be considered a complete success, since the mean rating of 3.75 (slightly below "somewhat high") indicates that respondents on average still underestimated the true sodium content of the soup.

Figure 6 reveals clearly that the absolute disclosure failed to communicate any additional information on the quantity of saturated fat in the advertised cheese. As with soup, the relative disclosure performed even more poorly than the absolute disclosure, although no statistical significance can be attached to the difference in saturated fat ratings (2.91 for the absolute vs. 2.61 for the relative). The rating of 2.86 for the original strong disclosure indicates that adding a dietary warning to the relative disclosure was on net of little or no value. The revised strong disclosure provides a striking exception to these disappointing results. The direct verbal warning succeeded almost completely in alerting consumers to the high level of saturated fat in the advertised cheese. The mean rating of 4.59 is significantly higher than any other cell ($p<.001$).

²² One alternative would have been to phrase the sodium probe as a belief question rather than as a strict ad communication question. That is, the introduction to Question 8 might have read "Do you think the advertised product is..." rather than "Based on what the ad says or implies...." We can only speculate as to whether this approach would have yielded useful information, or whether instead the shift to a beliefs perspective would have introduced biases or ambiguities that would have made the results difficult to interpret.

FIGURE 5

SODIUM CONTENT OF ADVERTISED SOUP

5: high
4: somewhat high
3: neither high nor low
2: somewhat low
1: low

MEAN RESPONSE

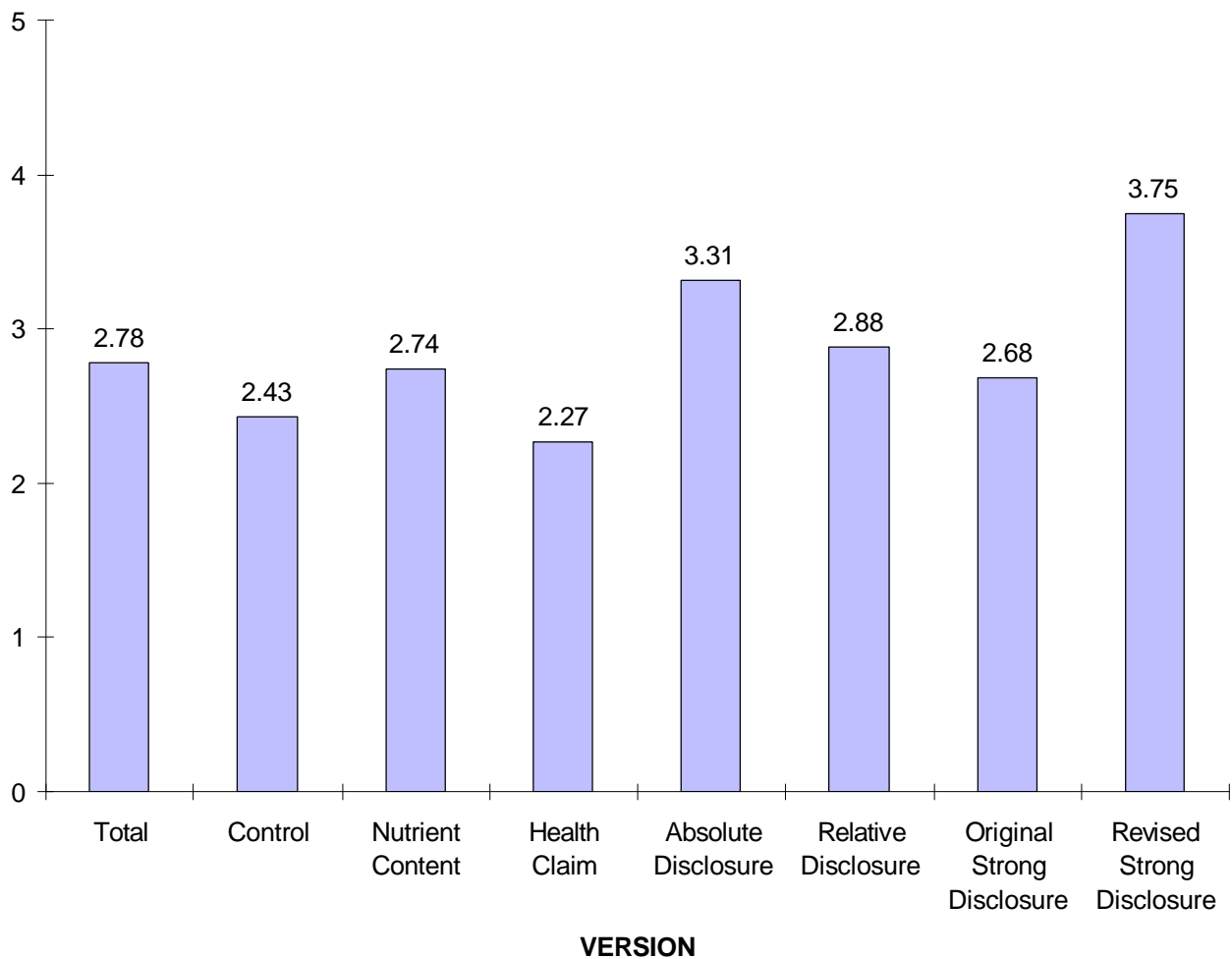
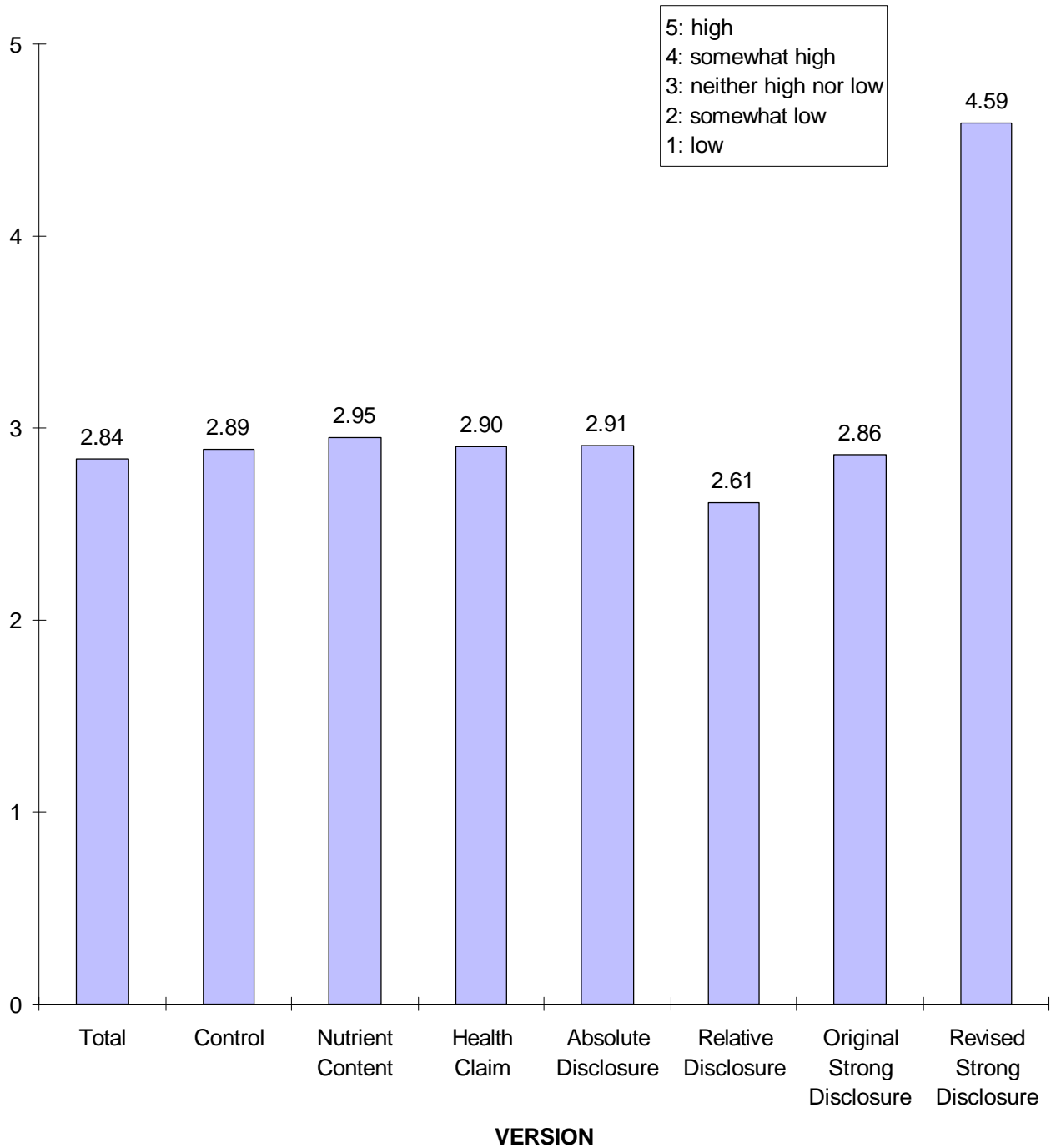


FIGURE 6

SATURATED FAT CONTENT OF ADVERTISED CHEESE

MEAN RESPONSE



The mean ratings for the remedy cells in Figures 5 and 6 obscure important underlying patterns in respondents' perceptions of sodium and saturated fat content that develop as the various remedy disclosures are added to the positive health claims. Figures 7 and 8 show the movement across cells in four categories of responses to Question 8: "Don't Know--Ad Doesn't Say," "Low, Somewhat Low," "Neither High nor Low," and "High, Somewhat High."

Both graphs reveal that the proportion of respondents who answered "don't know" or "ad doesn't say" falls precipitously as the remedies are introduced. This demonstrates that most respondents at least noticed the references to sodium or saturated fat content. Figures 7 and 8 also document that very few respondents interpreted the disclosures in a neutral fashion. In all but one cell (the cheese absolute disclosure), fewer than 10% of the respondents selected the "neither high nor low" option.

The dashed line graphs in Figures 7 and 8 show that, as hypothesized, many respondents interpreted the remedies correctly to indicate that the soup and cheese products were somewhat high or high in sodium or saturated fat. The proportion of responses in this category jumps to about 40% in the soup absolute disclosure cell and to about 25% in the corresponding cheese cell. The soup percentages fall slightly in the next two cells, but the proportion of "high, somewhat high" responses increases again to about 65% in the revised strong disclosure cell. For cheese, the percentage of "high, somewhat high" responses does not change in the relative disclosure, and increases to 45% in the original strong disclosure. This percentage then practically doubles when respondents are shown the revised strong disclosure.

The most surprising and troubling pattern in Figures 7 and 8 is the impact of the original remedies on the proportion of respondents answering "low, somewhat low." As is charted by the solid black line in these graphs, this proportion *increases* steadily as more detail is provided on the quantity and health significance of the sodium or saturated fat in the advertised products. Further, with the sole exception of the soup absolute disclosure, more respondents misconstrue the original disclosures as positive information than correctly understand the negative message that was intended. Thus, Figures 7 and 8 demonstrate even more clearly than the mean ratings in Figures 5 and 6 that the metric disclosures and the hybrid original strong disclosure essentially backfired, and that the extent of their perverse impact varied directly with the amount of information provided.

These results suggest more generally that many Americans are apparently not sufficiently familiar with metric measurements as applied to nutrient levels or sufficiently comfortable with percentage measurements and the concept of a Daily Value to interpret such disclosures correctly. Consumers may also assume that the disclosures are intended to be construed in a positive manner simply because they are presented in the context of an advertisement promoting the health benefits of a product.

FIGURE 7

SODIUM LEVEL RATINGS BY CATEGORY

SOUP

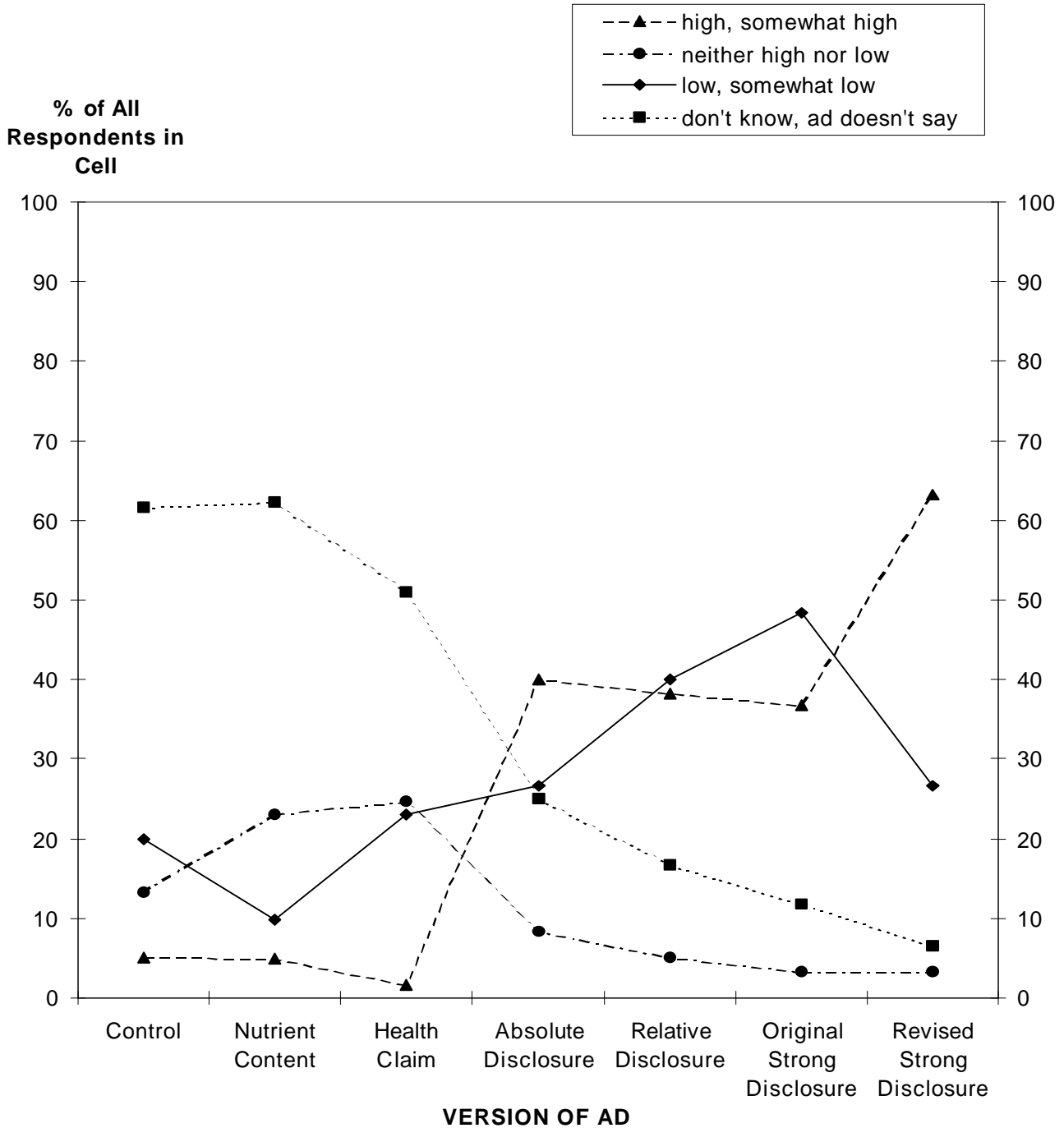
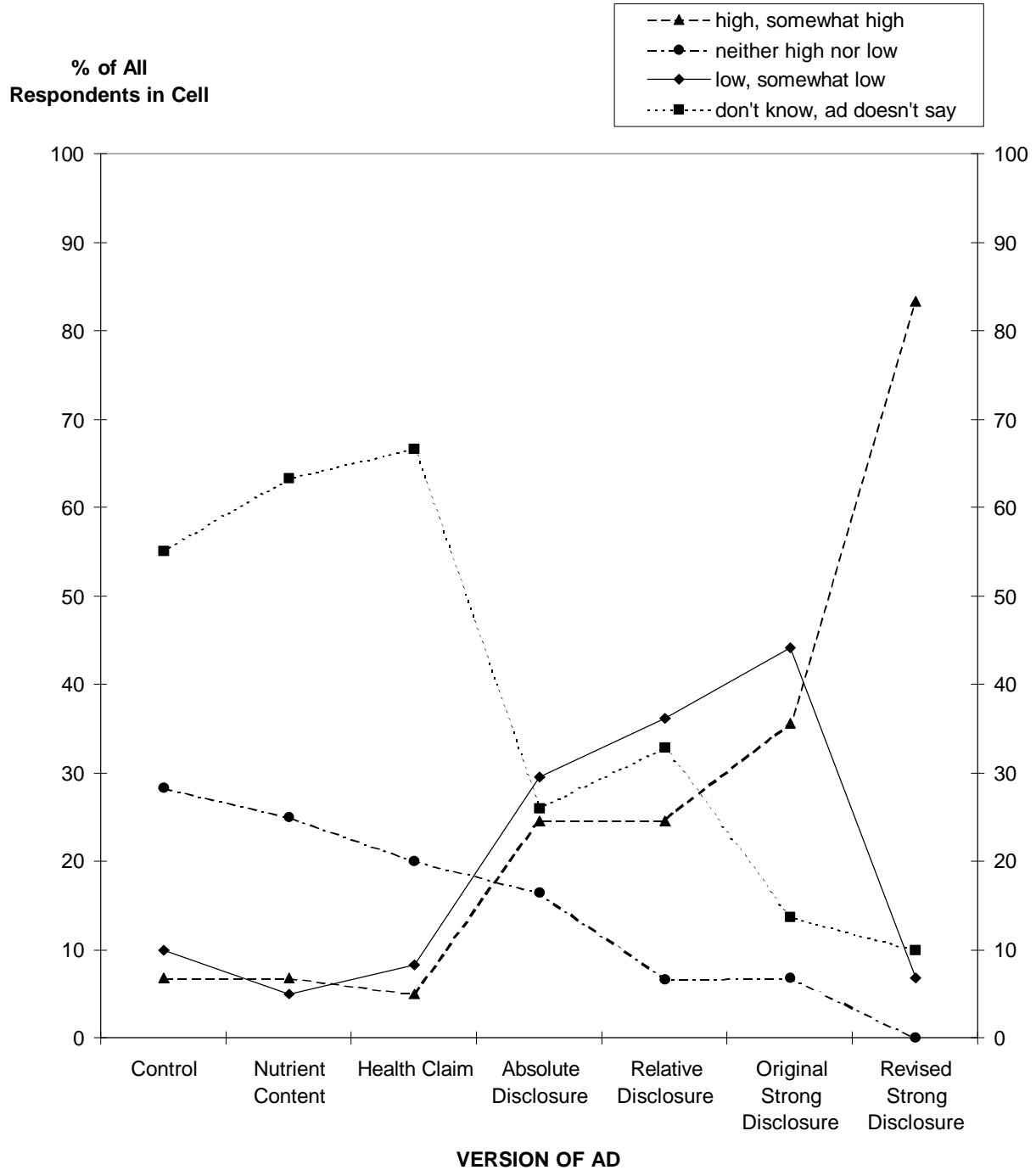


FIGURE 8

SATURATED FAT RATINGS BY CATEGORY

CHEESE



In addition, the particularly poor performance of the original strong disclosure seems to suggest that any consumer confusion concerning absolute or relative nutrient content disclosures will be exacerbated by an additional disclosure that explains the health significance of the problem nutrient in the context of the overall diet. It seems clear that those consumers who fail to understand the numerical disclosures will tend to misinterpret the dietary warning as a positive statement about the healthiness of the advertised product. Such consumers will assume that the products will help them reduce the quantity of sodium or saturated fat in their diets and therefore avoid the health risks mentioned in the warning.

Figures 7 and 8 also reveal that much, though not all, of the confusion caused by the original strong disclosure can be avoided by substituting for the numerical information a simple verbal statement warning that the product is high in the relevant risk increasing nutrient. In the case of cheese, this approach reduced the percentage of respondents taking a low or somewhat low saturated fat claim from 44 percent to only 7 percent.

The improvement is less dramatic though still appreciable for soup. The revised strong disclosure cut the proportion of seriously confused or misled respondents approximately in half, with the percentage of low or somewhat low responses falling from 48 percent to 27 percent. This is considerably below the corresponding figure for the relative disclosure (40%), and exactly equal to the absolute disclosure fraction. It should be noted, however, that on an overall basis the revised strong disclosure outperformed the absolute disclosure. Sixty-three percent of respondents in the new strong disclosure cell replied that the soup was high or somewhat high in sodium, whereas 40% of the absolute respondents gave this answer.²³

²³ We can only conjecture as to why the revised strong disclosure in the cheese ad communicated so much more effectively than the same format in the soup ad. One possibility is that consumers are more concerned about fat and saturated fat than sodium, and therefore focused more attention on the warning in the cheese ad. It is certainly true that the link between sodium and hypertension has received much less attention in the press and health community during the last decade than has the relationship between saturated fat and heart disease.

It may also be important that the positive health message in the soup ad was somewhat more complex than the equivalent portion of the cheese ad. The soup health claim was worded to track the FDA language for fiber-cancer claims in labeling. Because of the suspected link between certain kinds of cancer and high fat intake, this language places the fiber claim in the context of a diet that is low in fat. (It will be recalled that our version of the FDA fiber claim was: "Eating foods that are low in fat and high in fiber may reduce the risk of some forms of cancer.") Thus, the soup ad contained references to three nutrients--fiber, fat, and sodium--while the cheese ad mentioned only calcium and saturated fat. This may partially explain why the saturated fat advisory received greater attention than the sodium advisory.

g. Recall of Nutrient Information

Question 10 tested the degree to which respondents noticed the various nutrient content disclosures presented in the ads. Of particular interest was the proportion of respondents in the remedy cells that remembered seeing information about sodium or saturated fat.²⁴ For the cheese ads, respondents were asked whether they remembered seeing any information about fat, saturated fat, and sodium. Sodium was included to measure yea-saying bias, since this nutrient was never mentioned in any of the ads. Questions about both fat and total fat were included to determine how many respondents had noticed that the remedy disclosure concerned saturated fat rather than total fat.

Soup respondents were asked about fiber, sodium, and fat. The question about fat functioned as a check on yea-saying bias only in the control and nutrient content cells, since fat was mentioned peripherally in the health claim linking diets high in fiber and low in fat to a reduction in the risk of some forms of cancer.

For some unexplained reason, interviewers asked question 10 to only a small minority of respondents in the cheese control, nutrient content, and health claim cells, and to only 24 respondents in the soup control cell. These results are therefore of little value. In the remedy cells, most cheese respondents said they recalled information about both fat and saturated fat. Recall levels were lowest for the relative disclosure, where 78% mentioned total fat and 63% mentioned saturated fat. Corresponding recall levels in the absolute cell were 10 percentage points higher. Fully 90% of respondents in the revised strong cell recalled seeing information about saturated fat. Even when corrected for yea-saying bias, which ran from 7% to 19%, recall of the disclosures about fat or saturated was high enough to conclude that the cheese quantitative remedy disclosures did not perform poorly simply because they were overlooked.

Recall levels in the soup remedy cells were very similar. From 73% (absolute) to 88% (revised strong) remembered seeing sodium information of some kind. Recall of the positive fiber information was even higher, ranging from 92 to 98 percent.

²⁴ Before Question 10 was posed, respondents were asked a question intended to clarify the responses of consumers who thought the advertised products were low in a problem nutrient. In such cases, we wanted to know whether the respondents meant that the product was low enough to meet a formal nutrient content standard, or whether they were attaching a less rigorous interpretation, perhaps thinking that the food was relatively low for products of that type.

Given that the two foods in the halo effect test were unambiguously high in risk-increasing nutrients, and that any rating of low or somewhat low was incorrect however interpreted, this question was of less interest than an equivalent question asked in the Substitution Claim test, where the advertised product was in fact relatively low in sodium when compared with other packaged turkey slices. In any event, there are signs that the logistics involved in asking the question overwhelmed many of the interviewers and/or respondents, and the results are of insufficient reliability and interest to warrant discussion here.

h. Demographic Questions

The last two questions in the interview asked about respondent education and income level. The demographic results are summarized in Figures 9 and 10. The average education level of the cheese respondents as a group was slightly above a high school degree, and significantly higher than the soup respondents ($p=.09$). The average family income of the viewers of the cheese ads was also higher than that of the soup respondents (\$32,750 vs \$30,000), but this difference is not statistically significant. Across cells, there are no significant differences in income or education for either the soup or cheese respondents.

These data allow a test of whether the previously discussed bifurcation in respondents' interpretation of the various remedy disclosures (Q7) could be explained in part by differences in education level. Subsequent analysis, however, did not support this conjecture. A variety of cross tabulations and regressions failed to find any overall relationship between education and perceived sodium or saturated fat content in the remedy cells.

E. Conclusion

The results from the first segment of the food health claims copy test suggest three primary conclusions. First, judging from the responses to the purchase interest question, our respondents did not consider the positive health information presented in the nutrient content and health claim advertisements very important.

Second, and perhaps related, the results provide little or no support for the existence of a strong halo effect from an unqualified health claim. This is particularly true if the effect is construed very generally as an increase in the perceived healthiness of a product. Low response rates in certain test cells prevent firm conclusions on the more focused question of whether or not a health claim will lower consumer perceptions of the level of specific problem nutrients in a food.

Third, there is no support for the hypothesized capacity of either the sodium or saturated fat numeric nutrient content disclosures to improve respondents' overall understanding of the nutrient profile of the advertised products (Hypotheses IIa and IIb). These metric disclosures, whether expressed in absolute or relative terms, generally confused or misled about as many respondents as they helped. Coupling the quantitative measures with a dietary warning actually furthered the confusion. Based on the superior performance of the verbal approach tested in the revised strong disclosure, consumers apparently require nutrient content information to be presented directly in plain English.

FIGURE 9

AVERAGE EDUCATION LEVEL

CHEESE & SOUP RESPONDENTS

EDUCATION LEVEL

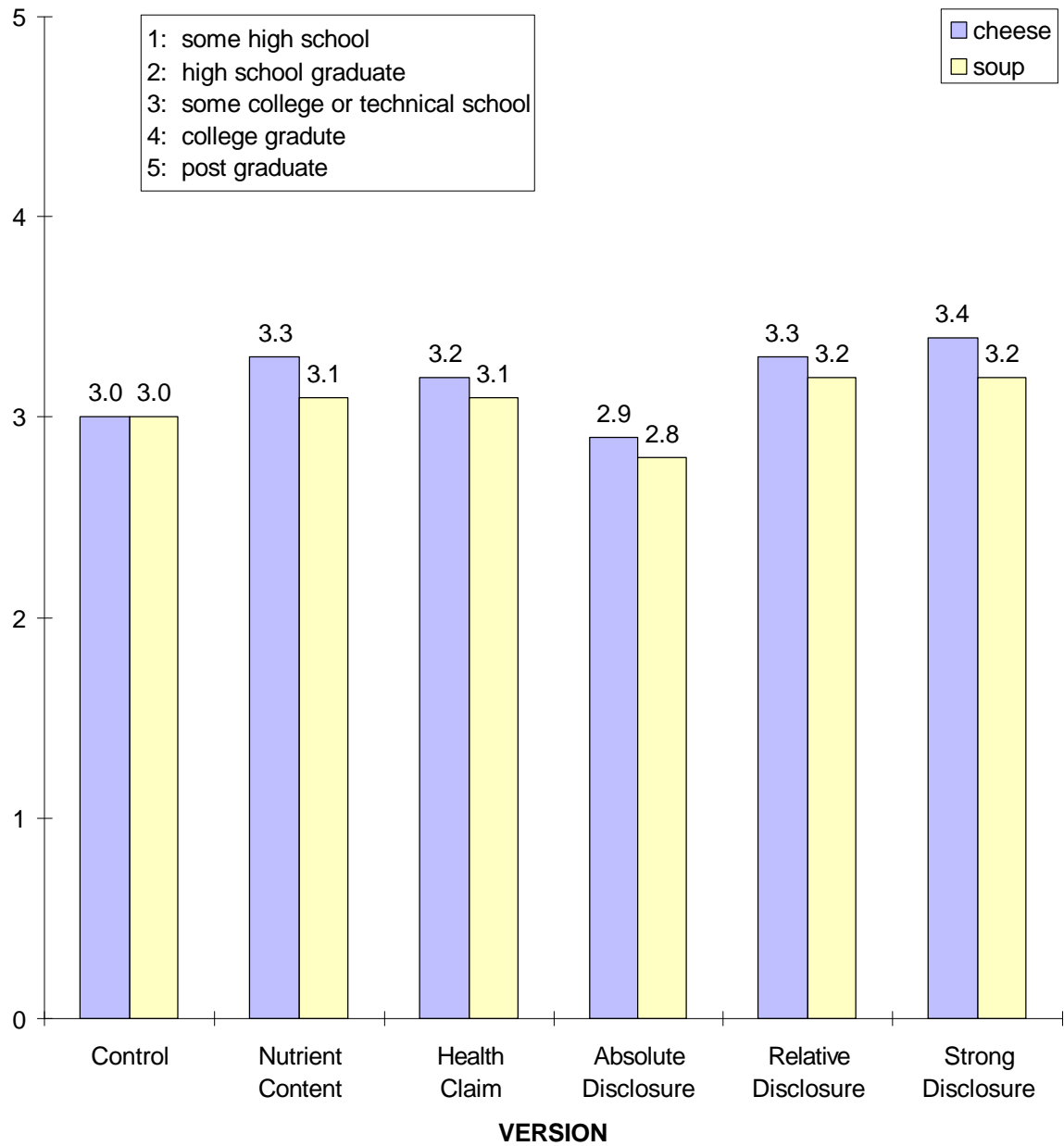
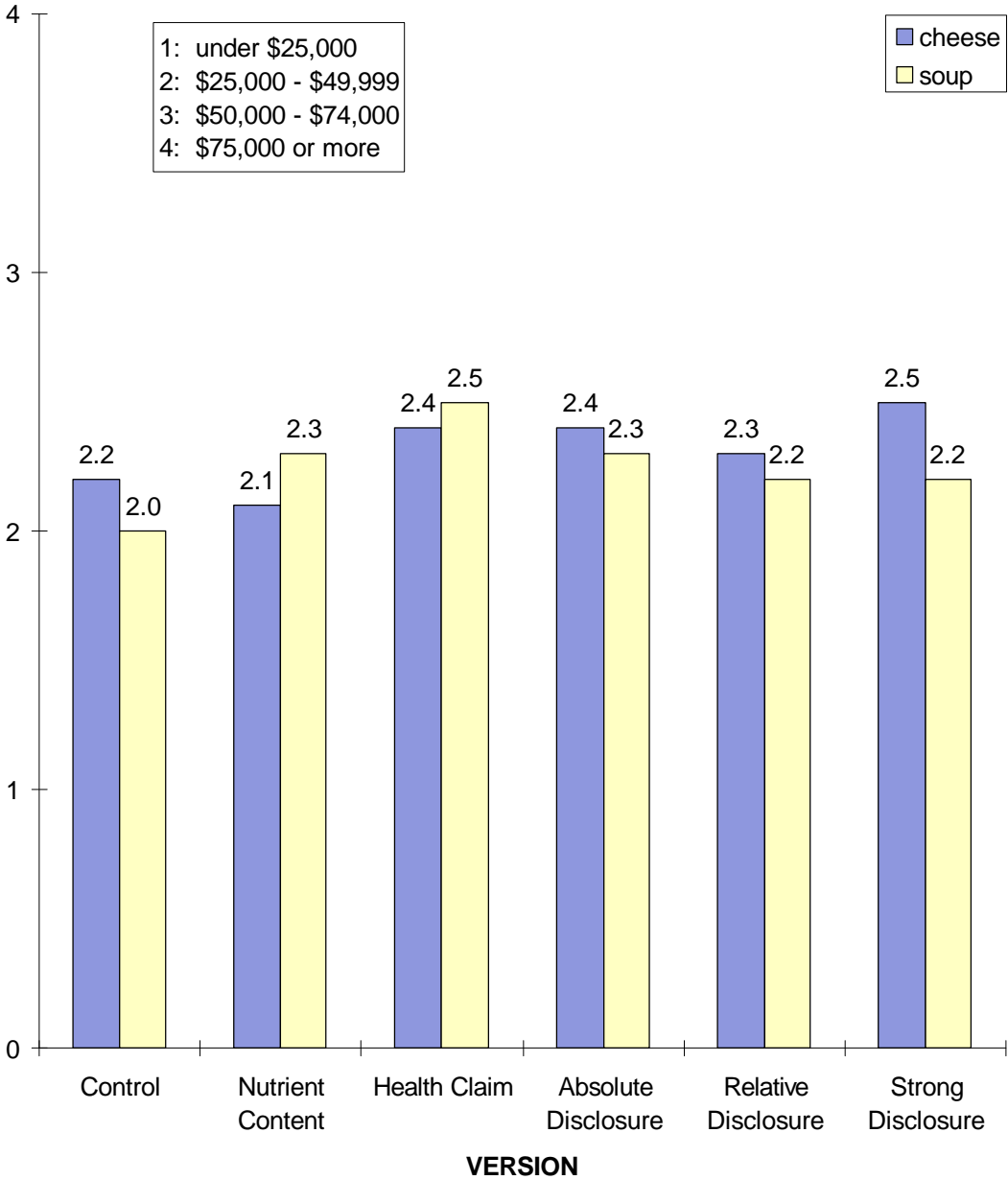


FIGURE 10

AVERAGE INCOME LEVEL

CHEESE & SOUP RESPONDENTS

INCOME LEVEL



IV. Detailed Discussion of Substitution Claim Component

As summarized earlier, the FDA labeling regulations prescribe strict minimum and maximum nutrient content requirements that are specific to each of the eleven health claims currently authorized. By way of illustration, no food product can relate its sodium content to a reduced risk of hypertension unless that food is "low" in sodium (140 mg of sodium or less per serving.) Calcium-osteoporosis claims are allowed only for foods that contain 180mg or more of calcium per serving, which is the standard for "high" calcium. In short, FDA generally reserves health claims for products that are among the very best in the relevant nutrient domain.

In the FTC Statement's less restrictive treatment of this subject, health claims may appear in advertising for products that are substantially "better" in a given nutrient characteristic, though not necessarily among the "best" in the food category. The governing considerations are (1) whether the ad describes a dietary substitution that will in fact appreciably reduce or increase intake of the nutrient(s) at issue, and (2) whether readers of the ad will understand the limited scope of the brand's nutrient superiority. The Statement cautions advertisers that "it may be necessary to disclose the actual level of the nutrient that is the basis for the claim and its significance to prevent deception...."²⁵

A. Hypotheses Tested

The first task of the Substitution Claim component was to determine whether consumers exposed to health claims for the type of product at issue would, in fact, overestimate the relative attractiveness of the food's nutrient profile when no information is provided concerning the absolute level of nutrients in the advertised or compared foods.

In principle, the Statement's treatment of substitution claims applies both to foods that are relatively low in a problem nutrient or relatively high in a beneficial nutrient. In practice, however, claims concerning reduced levels of risk-increasing nutrients such as fat or sodium might be more prevalent, since they would appeal to consumers who, for taste reasons, are unwilling to purchase products that are truly low in these nutrients, but who might still be seeking healthier alternatives to their current choices.

We therefore specified the first hypothesis to be tested as follows:

Hypothesis I: If an advertisement states that a food product contains less of a risk-increasing nutrient than other identified products, and if the advertisement relates this comparative advantage to a specific health benefit, the advertisement will convey to a substantial number of consumers that the advantaged product is low in the risk-increasing nutrient whether or not the product meets the relevant FDA definition of low.

²⁵ Statement, p. 25.

The second objective of this phase of the food copy test was to determine which types of nutrient disclosures or explanations would best correct any misimpressions that consumers received from the advertisement used to test the first hypothesis. Four remedy disclosures were tested. The first, identical to the absolute disclosure in the halo effect component, tested whether a simple statement of the absolute quantity of the risk-increasing nutrient in the advertised product and in the compared products would correct any impressions that the advertised product was low in the problem nutrient.

The remaining remedies implemented in varying formats the Statement's advisory that advertisers might have to disclose both the level of the relevant nutrient and its significance in order to prevent deception. The first attempt to indicate the significance of the nutrient level replicated the format of the relative disclosure in the halo effect segment. This remedy disclosed the level of the risk increasing nutrient terms both in absolute terms and as a percentage of the Maximum Daily Value.

The last two remedies tested more explicit disclosures. In one approach, an advisory was appended to the relative disclosure stating that the advertised product was not low in the relevant risk-increasing nutrient. The other approach tested whether a purely verbal advisory placed in the main body of the text would prove more effective than the separate quantitative or hybrid quantitative-verbal disclosures employed in the first three remedies.

The four tested hypotheses may be summarized as follows:

Hypothesis II: For the advertisement used to test Hypothesis I:

- a. Disclosing the absolute amount of the risk-increasing nutrient will on net increase the number of consumers who understand that the advertised product is not low in this nutrient.
- b. An additional disclosure that expresses the amount of the risk-increasing nutrient as a percentage of the Maximum Daily Value will on net increase the number of consumers who understand that the advertised product is not low in this nutrient.
- c. An additional disclosure stating that the advertised product is not low in the risk-increasing nutrient will on net increase the number of consumers who understand that the product is not low in this nutrient.
- d. In lieu of IIa-IIc, a separate disclosure in the main body of the text stating that the advertised product is not low in the risk-increasing nutrient will on net increase the number of consumers who understand that the product is not low in this nutrient.

B. Food Product and Advertisements Tested

The food chosen for the test ads was a fictitious brand of packaged sliced turkey ("Bradley Right Slices"). The test focused on sodium as the risk-increasing nutrient of interest. The absolute level of sodium in the product was chosen to be 240mg per serving, which is slightly less

than half the sodium content of regular turkey slices (about 600mg) but 100mg higher than the FDA upper limit of 140mg per serving for a low sodium claim.

All versions of the tested ads began with the following comparative nutrient content claim:

New Bradley Right Slices deliver the taste of oven roasted turkey with less than ½ the sodium of other leading brands!

This comparative nutrient content claim was followed by an explanation of the health significance of the lowered sodium content:

And that's important news. Because diets high in sodium can increase the risk of high blood pressure and heart disease!

The control ad, which is shown in Figure 1, provided no other information concerning the specific sodium content of Right Slices. Results from this cell were intended to provide a baseline for determining the proportion of consumers who would interpret the comparative nutrient information and the health claim to mean that Right Slices was in an absolute sense low in sodium.²⁶

The four remedy ads constructed to test Hypotheses IIa-IId are presented in Figures 2-5. The first remedy ("absolute disclosure") presented the sodium content of Right Slices (240 mg) and the corresponding figure for "most leading brands." The average figure for "most leading brands" was set at 600 mg, which is realistic for the most popular brands of packaged turkey slices currently on the market. The comparative sodium content information was presented in a large box immediately below the text discussion of the link between high-sodium diets and high blood pressure.

The second remedy ("relative disclosure") expressed the sodium content of Right Slices and the average for other leading brands both in milligrams and as percentage of the Maximum Daily Value (10% and 25% respectively). The third remedy ("strong disclosure") was identical to the relative disclosure, except that the following advisory was added in close proximity to the numerical disclosures: "Right Slices is not a low sodium food."

²⁶ This ad was not a true tombstone control, since it contained the same sodium comparison and health claim that appeared in the remedy ads. Thus, the results from this cell cannot reveal definitively whether any misimpressions concerning the sodium content of Right Slices arose from the sodium information in the ad, or whether they were formed by prior beliefs. Although it would have been preferable to test a true tombstone control that did not reference sodium in any way, it was decided that resources would better be devoted to an additional remedy ad than to testing the hypothesis that a large number of consumers believe that packaged poultry slices are low in sodium.

As is evident from Figures 2-5, the boxed remedy disclosures were even more prominent than the sodium or saturated fat advisories in the halo effect segment. Thus, these ads provided a best-case opportunity for quantitative nutrient content information to reveal its corrective potential.

The last remedy cell ("verbal disclosure") replaced the boxed sodium content information with a statement in the main body of the text that attempted in purely verbal terms to place the nutrient benefit of Right Slices in proper perspective:

Although Right Slices is not a low sodium food, switching to Right Slices can help you cut back on sodium.

C. Questionnaire Design

In general, the Substitution Claim questionnaire was identical to that used in the Halo Effect segment. In Questions 7-10, however, the substitution claim instrument asked respondents to compare the healthiness or sodium content of Right Slices both with foods in general, as had been the case in the Halo Effect survey, and also more specifically with other leading brands of turkey slices. The objective was to see whether respondents would rank Right Slices more favorably in the limited comparison with high-sodium competitors than in the broad comparison with all other food products. The full questionnaire for the substitution claim segment is presented in Appendix C.

D. Copy Test Results

1. Overview

The results from the Substitution Claim segment of our research indicate that, for the type of food product tested, a substantial minority of consumers will interpret a comparative "less than" claim for a problem nutrient to mean that the advertised product is absolutely low in that nutrient. Consistent with the findings in Halo Effect research, metric disclosures that expressed the level of the problem nutrient in milligrams and as a percentage of the Daily Value had little or no impact on such misimpressions. In contrast to the Halo Effect outcome, however, disclosures that characterized the level of sodium in simple verbal terms proved no more effective a remedy than the metric disclosures. These disparate results may indicate that even stronger or more prominent disclosures are necessary when the information being provided requires consumers to reconcile positive and negative messages about the same nutrient.

2. Results in Detail

a. Main Points of the Ad, First and Second Exposure

The first question was a completely open-ended inquiry into what respondents perceived as the main points of the ad. It was asked twice, once after the initial showing with the two clutter ads, and again after respondents reread the ad in isolation.

Although it was anticipated (and confirmed) that respondents would report a sodium content claim of some kind as the primary message, there were two major areas of interest concerning the pattern of these responses. The first was the proportion of respondents in the control cell that would express the sodium advantage of Right Slices in relative terms (e.g. "lower" or "less" sodium) rather than absolute ("low" sodium). The second important issue was the impact of the various remedy disclosures on this split between relative and absolute characterizations.

Figures 6 and 7 summarize, respectively, the results obtained from the initial exposure (with clutter ads) and the second showing of the ad. The percentages reported are the combined total of all response codings that reasonably could be considered a relative or an absolute sodium content message, and are therefore subject to some interpretational judgment calls. Still, certain patterns seem clear.

First, irrespective of cell, the plurality of respondents took away an absolute low sodium claim from their first cursory examination of the turkey ads. From 40% to almost 50% of respondents mentioned a low sodium claim of some kind; the corresponding range for lower or less sodium interpretations was 24-35 percent. Second, the pattern across cells appears somewhat perverse. The highest playback of low sodium claims generally occurred in the remedy cells, while the second-lowest incidence of such claims occurred in the control cell. It should be noted, however, that the control cell also posted the second-highest incidence of lower sodium claims.

Third, these response patterns essentially reverse after respondents had a chance to read the turkey ads a second time without the distraction of clutter ads. Figure 6 shows that the plurality of respondents now express the main point of the ad in relative rather than absolute terms. Further, with the exception of the strong disclosure cell, the proportion of relative characterizations increases steadily from a low of 40.3% in the control cell to a high of 51% for the verbal remedy.

The pattern for absolute low sodium characterizations, however, appears completely random after the second viewing. The proportion of respondents reporting such absolute messages seesaws back and forth across cells, with the greatest disparity occurring between the strong disclosure (40.7%) and the verbal disclosure (29%).

FIGURE 6

MAIN POINTS, FIRST EXPOSURE

TURKEY

PERCENT

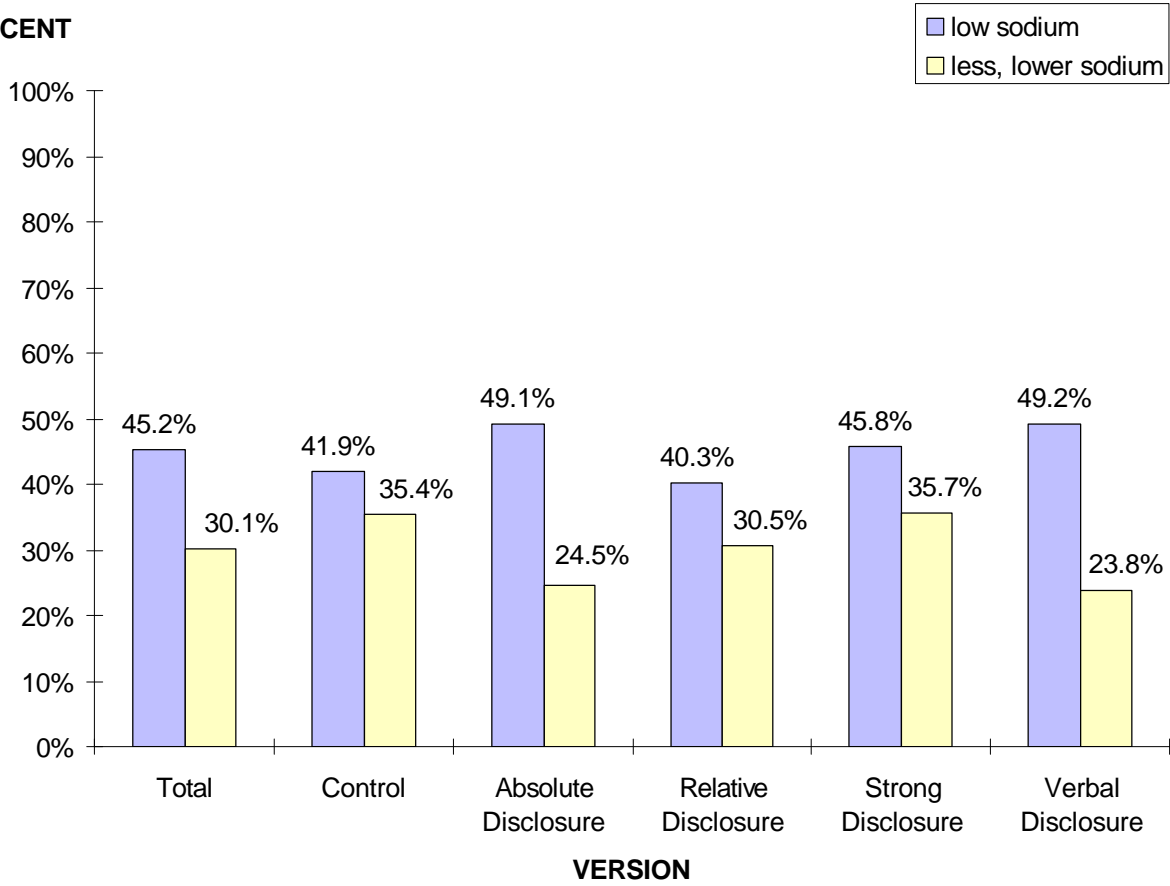
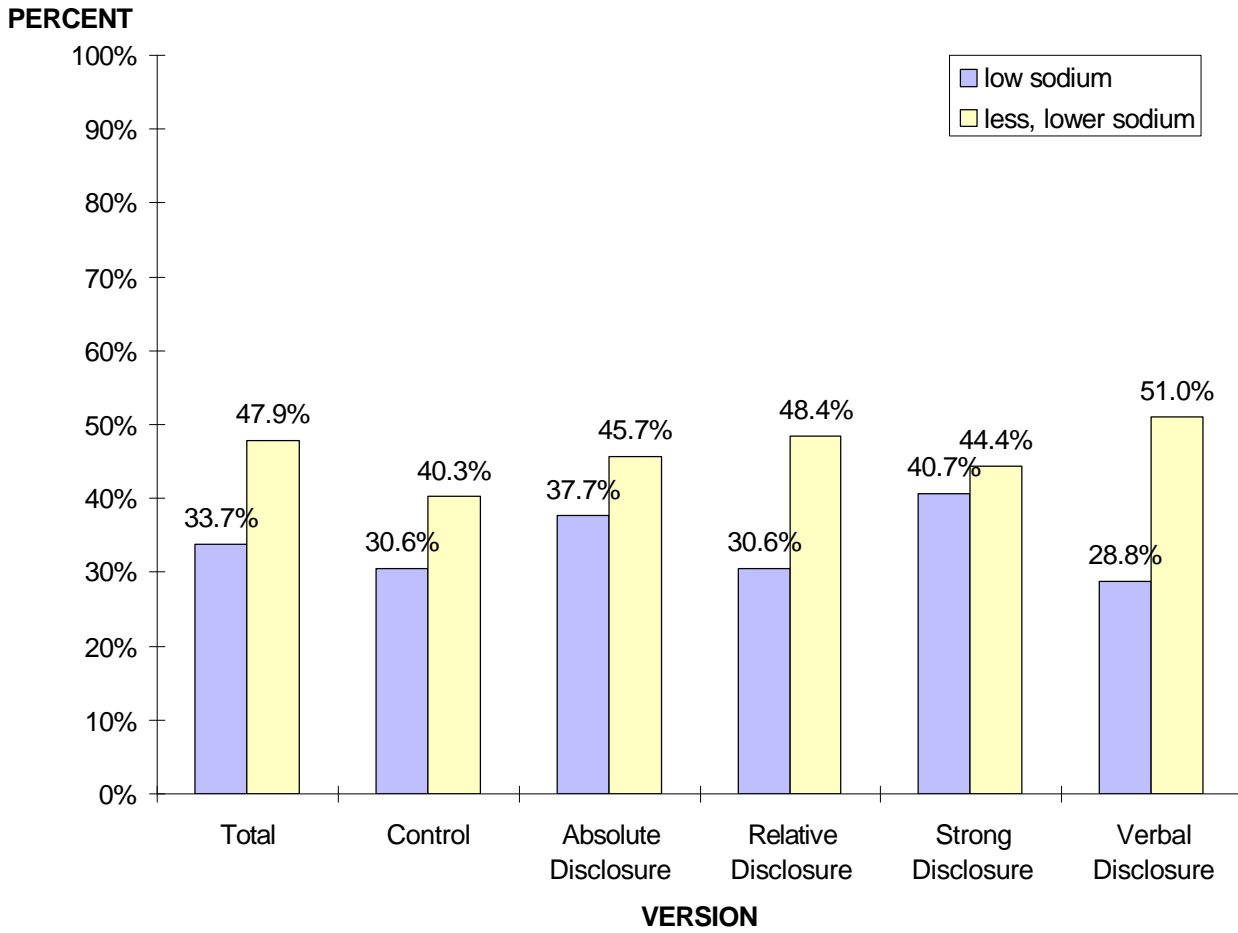


Figure 7
MAIN POINTS. SECOND EXPOSURE
TURKEY



As a whole, the second exposure results suggest that the verbal disclosure was most successful in communicating the true nutrient profile of the advertised turkey, since the greatest number of consumers viewed the sodium advantage in relative terms and the smallest number volunteered an absolute low sodium response. For no apparent reason, however, the worst performing remedy based on these same criteria appears to be the strong disclosure, which contained an explicit advisory that Right Slices was not low in sodium. About the only generalization possible from the initial open-ended results is that none of the remedies prevented at least a significant minority of respondents from reporting a low sodium message as a major point of the ad.

b. Purchase Interest

The first close-ended question in the survey explored the general appeal of the turkey slices as presented in the various ads. The question asked respondents to use a five-point scale to indicate how interested they would be in purchasing the advertised product. The pattern across cells would be very difficult to predict even if we assumed that respondents saw and understood the remedy disclosures. If respondents primarily were concerned about the health benefits of reduced sodium intake and were seeking a truly low sodium product, purchase interest might decrease in the remedy cells as it became increasingly clear that Right Slices did not meet this criterion. Conversely, purchase interest might increase in the remedy cells if participants were more concerned about the possibly bland taste of a low sodium food.

The actual results revealed essentially no patterns of any kind. As a group, the remedy cells registered a slightly higher rating than the control, but the differences are quite small and statistically insignificant. Within remedies, the strong and verbal disclosures scored lower than the absolute or relative, but again the differences are too small to deserve serious attention. The overall purchase interest rating of 3.04 was somewhat higher than that observed in the Halo Effect segment for soup and cheese (2.87 and 2.69 respectively).

c. Contents of Advertised Food

Question 4 was what proved to be a failed attempt to focus the respondents on the nutrient profile of the product in question. Its very general wording was as follows:

Q4. Did the ad say or imply anything about what is in the turkey slices?

The question perplexed about two-thirds of the respondents, who answered either "no" or "don't know." Of the minority giving an affirmative response, most referenced sodium in some manner, and of these a larger fraction expressed sodium content in comparative rather than in absolute terms. Mentions of "low sodium" ranged from 5%-10% of all respondents, while relative characterizations ranged from 6 to 17 percent. No clear patterns emerged across cells, and the response rate was, in any event, too low to justify detailed analysis.

d. Effect on Health

Question 5 asked respondents whether the ad stated or implied that the turkey slices would have any effect on their health; a follow-up probe asked respondents to identify any such effects. The results failed to provide any significant new insights and will not be discussed here. Two additional questions asked respondents to rate the general healthiness of Right Slices in relation to other sliced turkey products and to foods in general. Again, the results were not of sufficient interest to warrant presentation.

e. Ad Believability

In all components of the food copy test, respondents were asked to rank the believability of the tested ads on a 5-point scale. The purpose was to check for any extreme negative reactions respondents may have had to some aspect of the test ads' layout or wording. It was also of interest to see whether the various disclosures and advisories in the remedy cells would have any impact on credibility, although we had no strong conjectures as to the direction of any such effect.

This question generated fairly high believability ratings for the turkey ads. The mean score was 3.38, which is between "somewhat believable" and "very believable."). This rating is equal to or higher than the corresponding figures for the soup and cheese ads (3.21 and 3.36 respectively), and equal to the rating given the real vitamin supplement ad that was tested in the Strength of Science component. There were no significant variations across cells.

f. Sodium Content

The next two questions in the survey focused specifically on respondent perceptions of the sodium content of Right Slices. Respondents first rated the sodium content of Right Slices in relation to other leading brands of turkey, and then were asked to broaden the comparative context to foods in general. If respondents realized that foods in general tended to be lower in sodium than the leading brands of turkey referenced in the ads, the sodium ratings for Right Slices should rise when the basis for comparison is broadened.

Questions 9 read as follows:

Q9. Based on what the ad says or implies, compared to other leading brands of packaged sliced turkey, is the advertised product:

1	2	3	4	5
Low in sodium	Somewhat low in sodium	Neither high nor low in sodium	Somewhat high in sodium	High in sodium

Question 10 replaced the reference to other turkey brands with a comparison to "other foods."

The format of these questions was less than ideal, in that the response categories were absolute measurements, while the question was strictly comparative. It would have been more consistent to have phrased the response choices in relative terms, such as "much lower in sodium," "somewhat lower," etc. This would, however, have generated responses that could not be compared directly with the absolute nutrient content descriptor "low sodium" that was of primary interest. In any event, there were no signs in the pretests or in the verbatims that consumers were confused by this nonparallel construction. The results for Questions 9 and 10 are shown in Figures 8 and 9

As hypothesized, respondents as a whole rated Right Slices lower in sodium when the frame of reference was other brands of turkey rather than foods in general. The mean response to Question 9 was 1.48, vs. 1.65 for the broader Question 10 ($p=.008$). This pattern was uniform across cells except for the verbal disclosure, where the scores for the two questions were identical.

In both Questions 9 and 10, respondents seeing the absolute disclosure remedy gave Right Slices the lowest sodium rating (1.34 and 1.51 respectively). In Question 9, the scores generated by the other three remedies were virtually identical to the control rating of 1.50. In Question 10, only the relative disclosure produced a higher sodium rating than the control, and the difference is clearly inconsequential and insignificant. Indeed, there are no significant differences between any of the cells in either Question 9 or 10.

Among those respondents who replied "low" or "somewhat low," the most common justification offered in the follow-up probe was "it says so" or "that's what it said or implied." Interestingly, this explanation was most common for the strong disclosure ad, which included the "not a low sodium food" advisory. Fully half of the relevant subgroup of strong disclosure respondents offered this justification, although some of these explanations may have related to a "somewhat low" response and therefore may not have been inconsistent with the advisory.

The mean scores reported in Figures 8 and 9 do not permit a direct test of Hypothesis I or Hypotheses IIa-IId, which all turn more narrowly on the proportion of respondents that rated Right Slices explicitly as a low sodium food. Figure 10 breaks out from the Question 10 results the percentage of respondents in the various cells that rated Right Slices as "low in sodium" compared with other foods (which is the frame of reference embodied in the FDA nutrient content descriptors).

FIGURE 8

SODIUM CONTENT COMPARED TO OTHER BRANDS

TURKEY

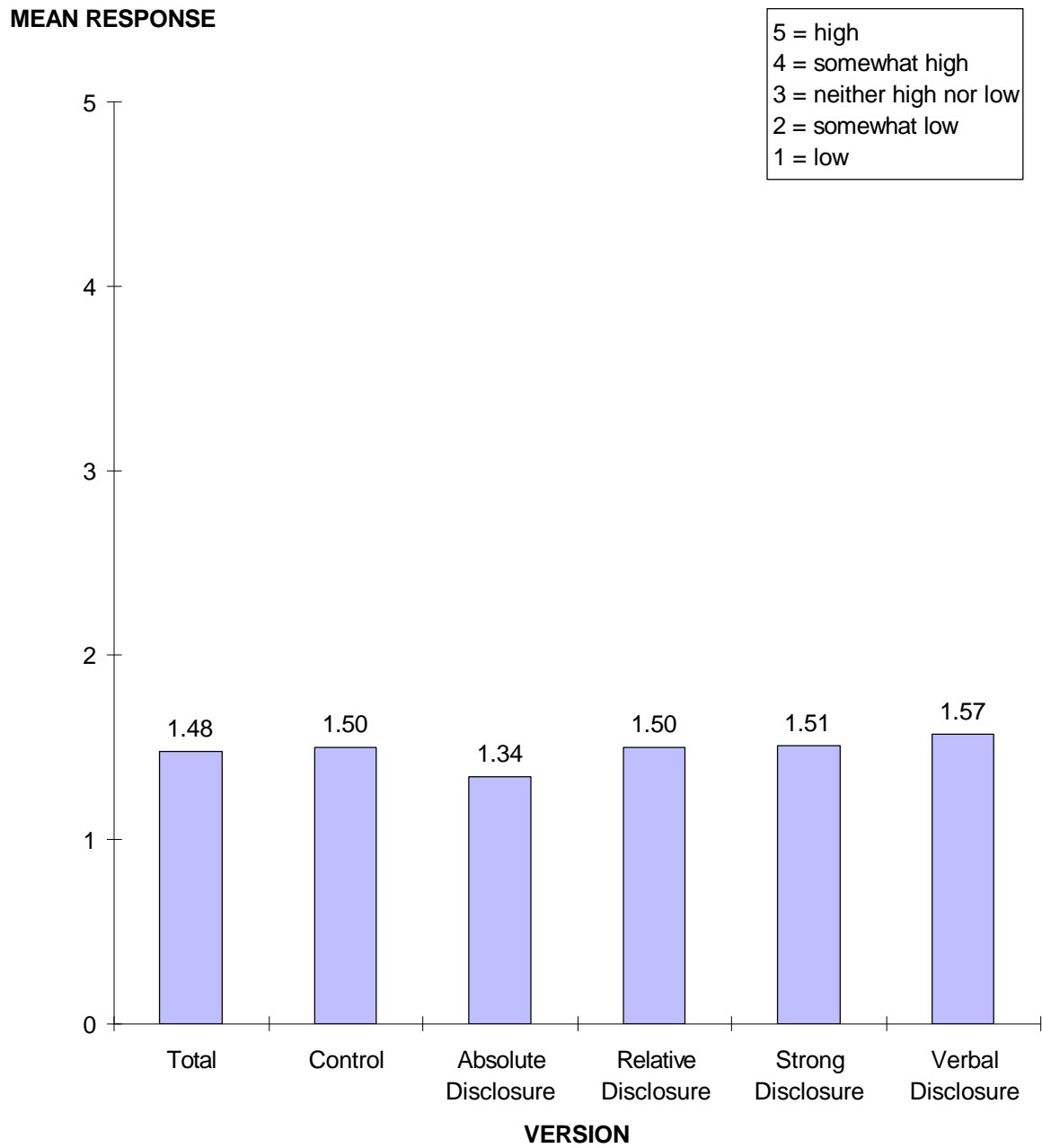


FIGURE 9

SODIUM CONTENT COMPARED TO OTHER FOODS

TURKEY

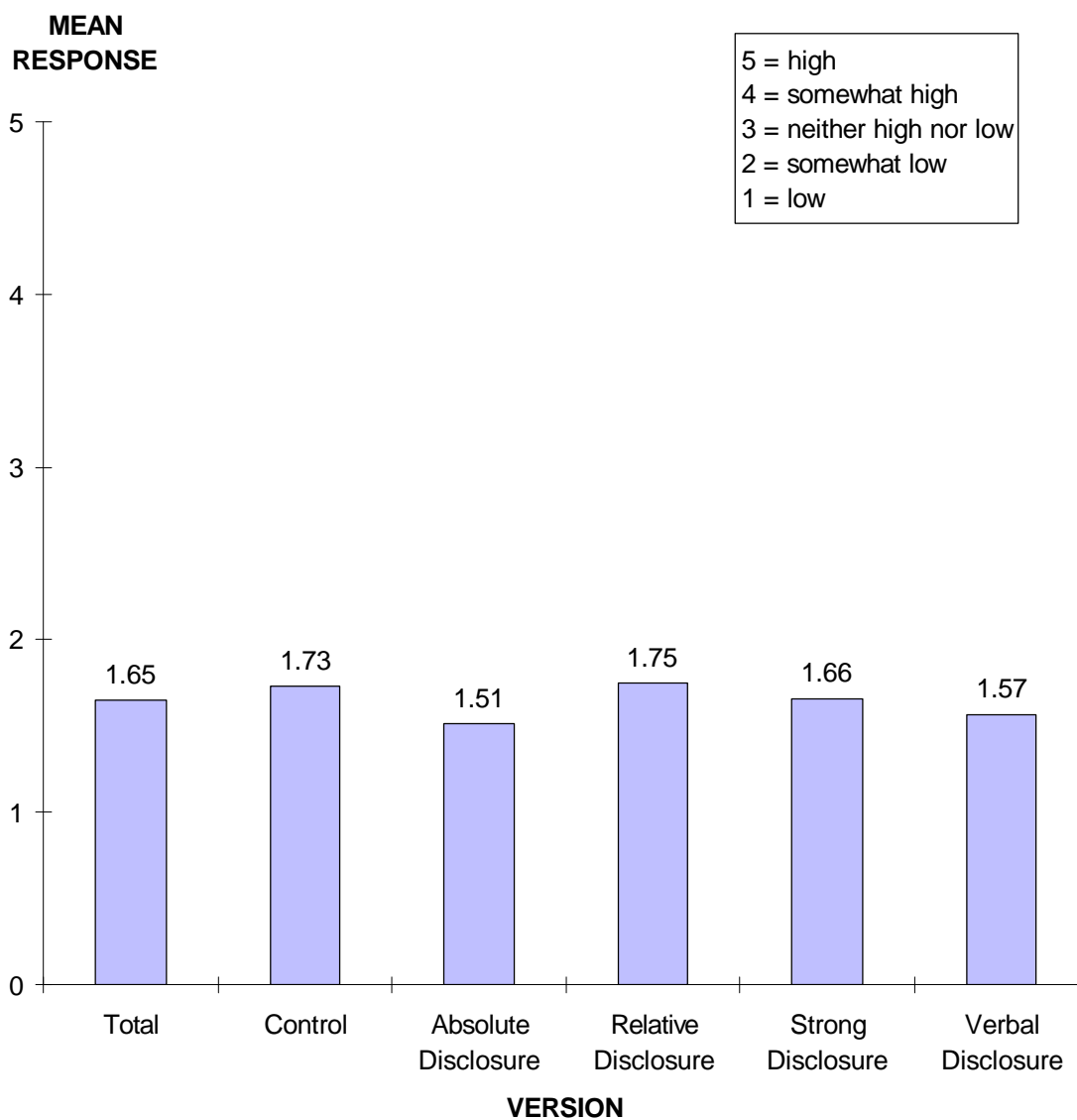
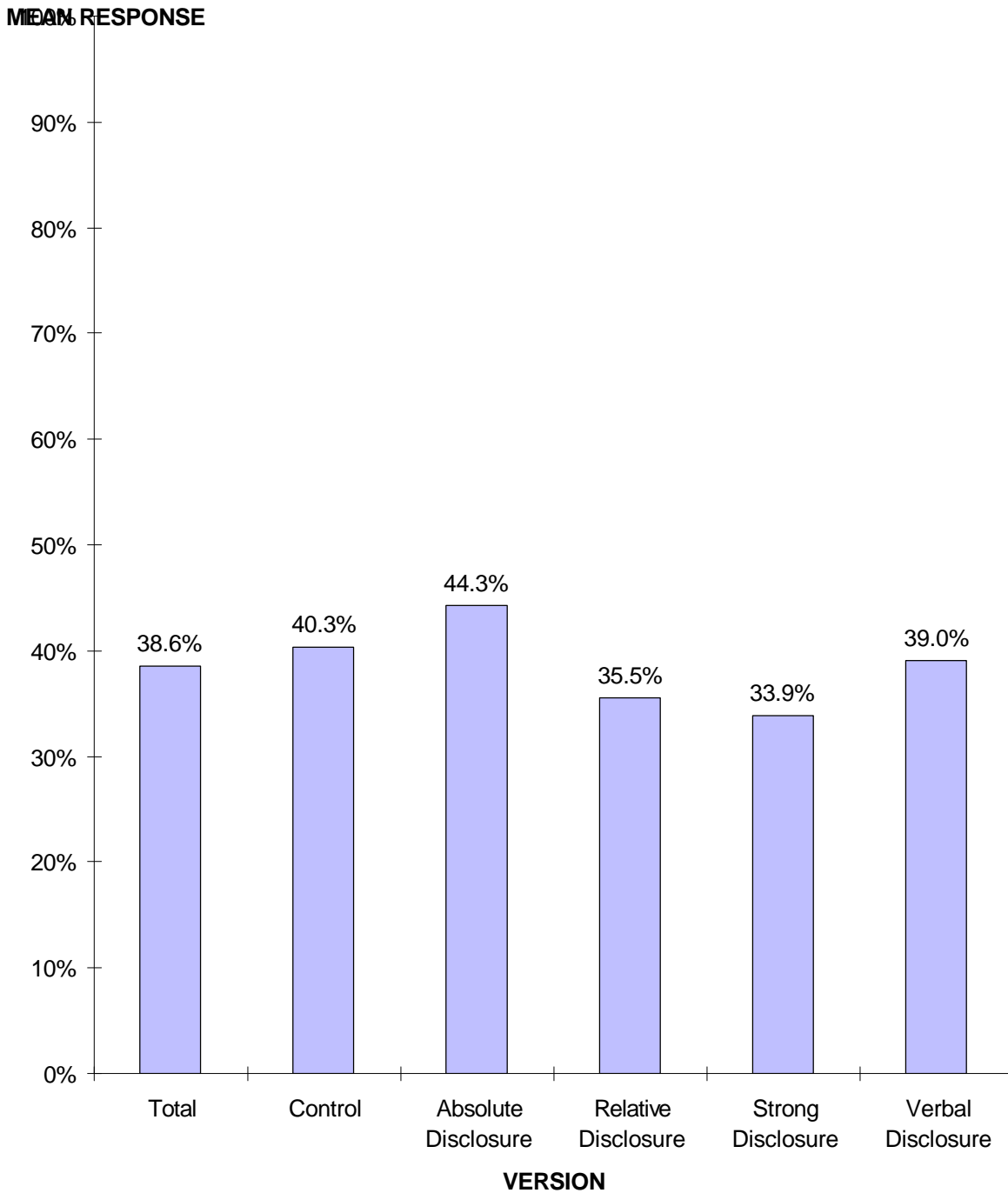


FIGURE 10
SODIUM CONTENT COMPARED TO OTHER FOODS
% RESPONDING LOW
TURKEY



Hypothesis I predicted that the control ad would convey a low sodium claim to a substantial number of respondents. Figure 10 confirms this prediction. Approximately 40% of the viewers of the control ad reported the claim. This represents the upper bound on the proportion of respondents misled by the ad, since there was no tombstone control for prior beliefs, and, in any event, we have no firm basis for concluding that these respondents thought that Right Slices was low enough in sodium to meet a formal definition of "low sodium."

Hypotheses IIa-IId predicted that the proportion of low sodium responses would continuously decline across the four remedy cells arrayed in Figure 10. The results reveal a more complex pattern. There is clearly no support for Hypothesis IIa, which predicted that the absolute disclosure would generate fewer low sodium responses than the control. The opposite occurred, although the difference is not significant.

The percentages for the relative and strong disclosures, on the other hand, are at least below the 40% figure for the control and therefore not inconsistent with Hypotheses IIb and IIc. These differences do not even approach significance, however. Finally, there is no support for Hypothesis IId, which predicted that a clear "not a low sodium food" advisory in the main text would prove the most effective remedy. The verbal disclosure generated essentially the same impressions among respondents as did the control ad.

The poor performance of the verbal disclosure is perhaps the most surprising element in these results. It is difficult to explain why virtually no readers of that ad saw (or at least attached any significance to) the very prominent and explicit statement in the text that Right Slices was not a low sodium food. A review of the explanations accompanying the follow-up probe to Question 10 reveals only one mention of the statement from respondents in the verbal disclosure cell.

In general, the unimpressive showing of the remedy disclosures was unaffected by educational or gender status, nutrition knowledge (as reflected in responses to the self-administered quiz), or interest in nutrition issues (as measured by responses to the 10 self-administered nutrition attitude questions.) A variety of probit regressions showed that the probability of a respondent rating Right Slices as low in sodium was not related in any of the cells to the respondent's quiz score or to the response for the most relevant single quiz question (which asked respondents to choose the correct Maximum Daily Value for sodium). Similarly, this probability showed no relationship to the combined score on the attitude questions or to the question that asked specifically whether a respondent was concerned about the amount of sodium in his or her diet. The only association that emerged from these regressions was a slightly higher tendency for females in the control cell to rate Right Slices low in sodium.

g. Comparison With Low Sodium Product

The questionnaires for both the Halo Effect and Substitution Claim components introduced at this point a comparative exercise intended to determine more definitively whether a substantial proportion of respondents in any of the cells thought the advertised food met a formal definition of "low sodium" or, for the sliced cheese product, "low saturated fat." In all cases, respondents were asked how they thought the level of the risk-increasing nutrient in the advertised product compared with that of a second fictitious product, a picture of which was

shown prior to questioning. The labeling for this product stated explicitly that the brand was "low" in the problem nutrient.

For the Substitution Claim component, the fictitious comparison product was "Ridgefield's Low Sodium Turkey Slices." It was hoped that respondents would interpret the nutrient content claim in the title as a reference to an official descriptor for low, and then assess whether they thought that Right Slices also qualified for this descriptor. Question 11 asked:

Q11. Please look at this picture I have just handed you. Compared to the brand in the picture you just looked at, would you say the advertised product we talked about first has....(Respondents selected one of 5 response categories ranging from "much less sodium" to "much more sodium.")

Since Right Slices exceeded the upper limit for a low sodium food by 100 mg, the most correct answer would have been response category 4--"somewhat more sodium." Although we had no reason to believe that respondents knew the specifics of FDA's "low sodium" descriptor requirements, it was nevertheless of interest to determine whether viewers of the strong and verbal disclosures, which explicitly disclaimed a "low sodium" designation for Right Slices, would rate the sodium content of the advertised product higher than the "Ridgefield's" slices.

Question 11 was worded to ensure as best we could that respondents would rate Right Slices in relation to Ridgefield's, and not vice versa. It was readily apparent from an inspection of the verbatims to Question 11, however, that many respondents had compared the sodium content of Ridgefield's with that of Right Slices. This meant that, in many cases, a response of "much less" or "somewhat less" should have been recorded as "much more" or "somewhat more." Despite our best efforts to identify the reversed comparisons from the verbatims and to adjust the responses accordingly, we no doubt failed to rectify the problem fully. The adjusted results, which are displayed in Figure 11, should therefore be interpreted with caution.

Taken at face value, the results indicate that respondents regarded Right Slices as exceedingly low in sodium. In most cells, including the verbal disclosure, the average score for Right Slices was between "somewhat less" and "about the same" sodium as the "low sodium" Ridgefield's turkey slices. Only in the strong disclosure does the mean rating even equal that of the low sodium reference product. Although we cannot be certain that respondents took the desired descriptor connotation from the Ridgefield's label, these results at the very least provide no support for the theory that consumers understood that Right Slices was relatively, but not absolutely, low in sodium.

h. Demographic Analysis

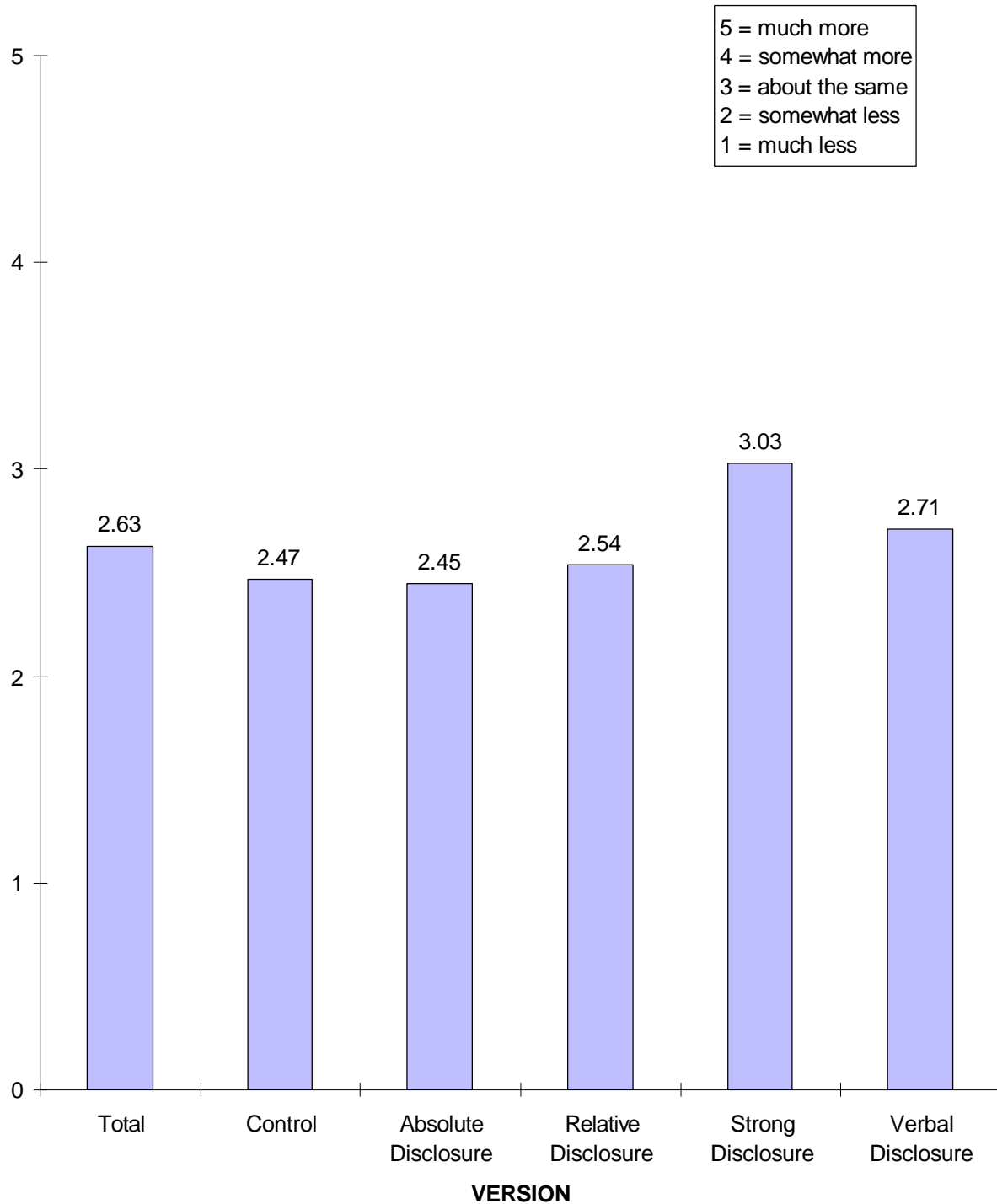
The Substitution Claim survey closed with the customary questions about respondent education and income level. The demographic profile of these respondents closely resembled that of the Halo Effect participants. On average, respondents scored 3.03 on the 5-point

FIGURE 11

COMPARISON WITH LOW SODIUM PRODUCT

TURKEY

MEAN RESPONSE



education scale (some college or technical school training), which is virtually identical to the educational status of the Halo Effect participants. Across-cell differences were not significant.

The average income of the Substitution Claim respondents was about \$29,000, which is slightly below the soup (\$30,000) and cheese (\$33,000) respondents. Again, there were no significant differences in income across cells.

E. Conclusion

The results from the Substitution Claim component provide fairly strong support for Hypothesis I, which predicted that a substantial number of respondents would take a low sodium claim from the control ad's sodium comparison between Right Slices and other leading brands of packaged turkey slices. About 30% of the control cell respondents volunteered a low sodium interpretation of the ad in answering the first open-ended question (although a higher fraction expressed the sodium claim in relative terms). This figure rose to 40% in the relevant close-ended question.

Although logistical difficulties compromised our efforts to explore the precise meaning respondents attached to the phrase "low sodium," the results are consistent with the view that the participants who described Right Slices in these terms really meant that the product was in an absolute sense a low sodium food. All of these conclusions are subject to the caveat that the Substitution Claim segment did not employ a true tombstone, and we therefore cannot determine the extent to which the responses were shaped by prior beliefs rather than claims conveyed in the test ad.

Results from the remedy cells are consistent with some, but not all, of the findings of the Halo Effect segment. One result seems quite robust. Most consumers did not know how to interpret a disclosure of the absolute amount of a nutrient expressed in grams or milligrams. The absolute disclosure consistently performed poorly, and even perversely, in both the Halo Effect and Substitution Claim components. Although the relative disclosure at times appeared to communicate more effectively than the absolute disclosure in the turkey ads, the differences were insignificant either from a practical or statistical standpoint. Thus, both the Halo Effect and Substitution Claim results point strongly to the conclusion that, in the context of advertising, metric disclosures such as those found on food product labels are ineffective vehicles for correcting misimpressions consumers may have concerning saturated fat or sodium content.

In other respects, results from the remedy cells appear inconsistent with corresponding findings in the Halo Effect segment. It will be recalled that by far the most effective sodium or saturated fat disclosure for the soup and cheese ads was a verbal statement warning that the advertised food was high in the problem nutrient (followed by an explanation of associated health risks). In the substitution claim segment, however, the performance of the closest analogues to these remedies--the strong and verbal disclosures--was statistically indistinguishable from that of the absolute or relative disclosures.

The principal difference between the verbal advisories in the two segments was that the Halo Effect statements warned of the presence and significance of high levels of a bad nutrient, whereas the substitution claim disclosures merely advised that the turkey slices were not low in

sodium. Further, the disclosures in the Substitution Claim segment were imbedded in ads that otherwise promoted the healthiness of the product's sodium profile. In the Halo Effect ads, the warnings concerned nutrients that were not the subject of the positive health claims. Thus, one possible explanation for the seemingly disparate results is that the prominent and highly positive discussion of the lower sodium content of the turkey slices dominated perceptions and distracted respondents from what may have seemed unimportant or contradictory messages.

V. Detailed Discussion of Strength of Science Component

A. Introduction

The Strength of Science segment tested consumer responses to a series of advertisements containing qualified and unqualified health claims for two disease relationships that have not yet met FDA's "Significant Scientific Agreement" standard of proof, and therefore have not been sanctioned for labeling. As explained earlier, the Statement allows advertisers to discuss unapproved diet-disease relationships, provided that such ads are carefully qualified to convey the actual level of scientific support for the relationship, and there is no larger body of scientific evidence that contradicts the claim.

The two nutrient-disease relationships for our research were (1) antioxidant vitamin supplements and a reduction in the risk of cancer, and (2) trans fatty acids and an increase in the risk of heart disease.²⁷ Print ads for a fictitious vitamin supplement product ("ACE Vitamins") provided the vehicle for testing health claims concerning the cancer reduction benefits of antioxidants. The possible heart risks of trans fatty acids were presented in ads for an imaginary new margarine ("Better Blend") that claimed to be free of trans fatty acids.

These nutrient subject areas were selected in part for their topicality, but more principally because we believed the level of scientific support in the two areas had advanced to the point where it might be possible to craft qualified health claims that could provide a meaningful test of the guidance provided in the Statement. In both cases, the hypothesized relationships were supported by a level of scientific evidence that could not be characterized as purely speculative or even preliminary. At the same time, limitations in the types or sheer quantity of research prevented firm conclusions on the efficacy of antioxidant vitamin supplementation or the importance of the heart disease risks of trans fatty acids.

More specifically, at the time our research was conducted, an extensive scientific literature had developed suggesting a positive relationship between dietary intake of certain antioxidant vitamins (principally beta carotene and vitamins C and E) and a reduction in the risk of several cancers.²⁸ There were, however, no completed long-term, placebo-controlled clinical trials using

²⁷ Trans fatty acids are found most commonly in commercial baked products and margarines that contain "partially hydrogenated vegetable oil." This substance is formed by bubbling hydrogen through vegetable oil, which adds enough hydrogen atoms to straighten the carbon chains that form the fatty acids in the oil. The new configuration forms a solid crystal at room temperature.

²⁸ This literature included a long record of animal tests with positive results for antioxidant vitamin supplementation, numerous epidemiologic studies showing an inverse relationship between dietary intake of carotenoids and cancer risk, and matched case studies and serum level studies supporting a protective effect of beta carotene intake for lung cancer and oral cavity cancers. Epidemiologic evidence and serum level studies also pointed to the efficacy of Vitamin
(continued...)

vitamins C or E, and only one such clinical using beta carotene. That study, involving 29,000 male smokers in Finland, reported a *higher* incidence of lung cancer for smokers in the beta carotene treatment group, and no reductions in the risk of any other forms of cancer.²⁹ Thus, absent additional clinical evidence, scientists could not reject the possibility that the cancer reduction benefits reported in non-clinical studies were attributable to some other unidentified element in foods that contained high levels of antioxidant vitamins.

Although the literature on trans fatty acids and heart disease was more limited, it did include several well conducted clinical studies indicating that diets high in trans fatty acids could increase serum cholesterol levels, and thus the risk of heart disease. The principal scientific controversy concerned the magnitude of this adverse serum cholesterol impact, rather than its existence per se.³⁰

As detailed below, the remedy test ads used in the Strength of Science component contained qualifying language that attempted to describe in varying degree the extent and nature of the scientific debate concerning the health effects of antioxidant vitamins and trans fatty acids. For comparative purposes, our research also employed an unqualified version of these ads that claimed cancer or heart benefits had been proven for the advertised products.

B. Experimental Design

This section of the food copy test posed by far the most difficult analytical challenges and required compromises in the precision of the tested hypotheses and the analytical rigor with which they were developed. The most fundamental challenge was determining the level of certainty that consumers should attach to a health claim about antioxidant vitamins or trans fatty acids.

²⁸(...continued)

C in reducing the risk of oral cancers and cancers of the esophagus and stomach. The results of studies investigating Vitamin E were less consistent. For a review of the scientific literature on antioxidant vitamins and cancer risk, see J. Dorgan and A. Schatzkin: *Antioxidant Micronutrients in Cancer Prevention*, Hematology Clinics of N. America: 5: Feb., 1991.

²⁹ O.P. Heinonen, J.K. Huttunen, et al.(N Engl J Med 1994;330:1029-35). The seemingly anomalous results were viewed with some skepticism when the Finnish study first appeared. Recently, however, a clinical study of over 18,000 U.S. smokers and workers exposed to asbestos was terminated before completion due to preliminary results that also found an increased incidence of lung cancer among recipients of beta carotene supplements. (See Omen GS, Goodman GE, et al. N Engl J Med 1996;334:114-9.)

³⁰ The most important clinical intervention trial was conducted by researchers at the Department of Agriculture (Judd, et al., 1993.) This study found a rise in serum cholesterol when trans fatty acids were consumed at levels typical of the American diet. Several contemporaneous trials reported similar results (Wood, R. et al., 1993; Lichtenstein, et al., 1993; Mensink, et al., 1992). A major epidemiologic study found that U.S. nurses who reported consuming about 6 grams of trans fatty acids per day had a 50% higher risk of coronary heart disease than nurses who reported consuming about 2.5 grams per day (Willett, et al., 1993).

Reasonably intelligent but lay readers of the literature on antioxidant vitamins and cancer might take away widely varying impressions of the current level of scientific certainty in this field. Although the perceived range of probabilities would be narrower for the existence of a positive link between trans fatty acids intake and serum cholesterol levels, this range would widen considerably if the issue were posed in terms of the practical importance of the relationship, particularly in comparison with the proven link between saturated fat intake and cholesterol levels.

One approach to the problem would have been to ask leading researchers in these areas to rate their own certainty of the relevant nutrient-disease relationship on a scale of 1-100. We might then have asked respondents viewing our test ads to indicate on the same scale how certain they thought scientists were about the relationship. This would have anchored the analysis in a manner that would have allowed a meaningful test of the success with which a given advertisement conveyed the proper level of scientific uncertainty.

We did not pursue this polling of medical researchers due to resource constraints and a concern that these individuals would be reluctant to engage in such a speculative exercise, particularly for a government agency. We did, however, retain the basic construct of the exercise in formulating the key close-ended interview question, which asked viewers of the various ads to indicate on a five-level scale "how sure" they thought "scientists" were about the nutrient-disease relationship at issue.

Since we lacked a concrete reference point for assessing the absolute accuracy of respondents' evaluations of qualified and unqualified health claims, our research attempted only to test for *differences* in the level of certainty that respondents would attach to health claims with varying degrees of qualification. Our basic objectives were to determine (1) whether consumers would in fact distinguish between qualified and unqualified health claims, and, if so, (2) what minimum level of qualification was needed to achieve a significant difference in the certainty scores that respondents assigned to the two types of claims.

The range of possible qualification levels was constrained by a need to portray the existing state of scientific evidence in a truthful manner, without exaggerating the extent of controversy or impeaching the significance of valid research. For reasons of both public policy and law, the Commission would never require advertisers to include disclaimers that were not truthful in their own right. Further, to the maximum extent possible, we wished to test qualifications that did not undermine the positive messages advertisers might validly convey. Our objective was to determine whether the guidance in the Statement could be implemented in a manner that did not constitute an effective ban on advertising claims concerning diet-disease levels that have not yet achieved significant scientific agreement.

An analysis of the relevant scientific literature suggested two levels of qualifications that would be appropriate for both the antioxidant vitamin supplement and margarine ads, and a third qualification level specifically appropriate for the vitamin supplement product. The first level represented what we considered to be the least intrusive disclaimer that might reasonably alert consumers to the lack of scientific certainty in the two areas of interest. The ads incorporating this disclaimer maintained a positive overall tone, and should be regarded as a very lenient implementation of the Statement's guidance.

For the vitamin product, the first level of qualifying language was limited to a simple advisory that scientists would not be sure about the cancer benefits of antioxidant vitamin supplements until "longer term research" was completed. The margarine disclosure stated that scientists would not know "how big a problem" trans fatty acids were until further research was conducted.

The second qualification level provided additional detail concerning the extent and nature of current scientific uncertainty and represents a strict, though reasonable, implementation of the Statement's call for advertisers to "carefully qualify" the level of scientific support for health claims of this type. Specifically, the antioxidant vitamin qualification disclosed that some recent studies had failed to find a cancer benefit for supplements. This was an oblique reference to the Finnish study's negative findings for beta carotene and to other research that has not always reported cancer reduction benefits for one or more antioxidant vitamins. The corresponding margarine ad disclosed that there was an ongoing scientific debate concerning the amount by which trans fatty acids raise serum cholesterol. In addition, the second level of qualification emphasized the incomplete status of research in both fields by cautioning readers that "It's too early to tell for sure."

The third and strongest qualification, limited to antioxidant vitamin supplements, disclosed the Finnish study's finding that high doses of an antioxidant may actually have increased the risk of cancer among smokers. Although this statement arguably went beyond mere qualification to include disclosure of a negative material fact, and might also prove unpalatable to advertisers of these vitamin products, we included the warning to determine whether disclosures of this severity are needed to affect consumers' perception of scientific certainty.

Two other variants of the vitamin supplement and margarine ads completed the experimental design for the Strength of Science component of our research. To control for prior beliefs concerning antioxidant vitamin supplements or trans fatty acids, one group of respondents was shown an ad that described the composition of the vitamin supplements or the margarine without making an explicit health claim.³¹ Finally, another test cell was devoted to an advertisement that presented the health benefits of the vitamin supplement or margarine as established scientific fact. This unqualified "proof claim" served as the frame of reference for gauging the impact of the various levels of qualifications in the remaining ads.

As will be even clearer shortly, there were too many differences in the wording of most of the ads to permit testing of narrowly drawn hypotheses. Had sufficient resources been available, it would have been preferable to test a long series of ads, each one of which differed only in the

³¹ The complement of ads lacked a true tombstone control that did not describe the vitamin or nutrient composition of the advertised products in any manner. Although such a control was not needed to measure the marginal impact of the various qualified and unqualified health claims, its absence did prevent us from gauging the effect of the simple "no trans fatty acids" claim on respondents' perceptions concerning the general healthiness of Better Blend or, more specifically, the total fat content of the product.

addition of or change in a single key phrase. With the exception of the most highly qualified vitamin ad, which did test the impact of a specific negative disclosure, the ads that finally emerged allow only a much rougher analysis of what (if any) degree of qualification is needed to convey scientific uncertainty about a health claim that FDA has not approved for labeling.

C. Complete Description of Test Ads

The vitamin supplement and margarine ads developed for the Strength of Science component are displayed in Figures 1-10. The two control ads--Figures 1 and 7--are basically nutrient content claims with additional references to convenience, taste, and/or cost. No information is provided on the health significance of antioxidant vitamins or trans fatty acids.

The test ads with health claims all adhere to the same format. The ads begin with the banner and headline that appeared in the controls. Below this point, the text is presented in two subsections.

The first of these, which is introduced with the headline "What We Know About Antioxidants and Cancer" or "What We Know About Trans Fatty Acids and Heart Disease", contains brief background information concerning antioxidant vitamins or trans fatty acids, and closes with an explicit health claim. In the unqualified vitamin ad, for example, the health claim states that "scientists have now proven" that antioxidant supplements reduce the risk of cancer. This ad, and the corresponding margarine ad, will be referenced hereafter as the "proof claim," and are shown in Figures 2 and 8. The corresponding health claims in the various qualified ads state that medical studies "report" or "suggest" the relevant diet-disease relationship.

The second subsection, which appears under the heading "What This Means To You," showcases the varying levels of qualifications concerning the certainty of the scientific evidence for the health claim. The ads with the mildest disclaimers--hereafter referenced as the "qualified claim"-- are shown in Figures 3 and 9. As discussed earlier, a deliberate attempt was made to cast these ads in a positive light in order to determine whether a strong sales message could still convey a significant degree of qualification. Accordingly, this section of the vitamin ad is introduced with the phrase "it looks promising," followed by the caution that longer term research is still needed. The corresponding opening for the margarine qualification is "It looks like you should try to avoid foods with trans fatty acids."

Figures 4 and 10 present the ads with the next level of qualification. These "highly qualified claims" are much more emphatic about the degree of scientific uncertainty or controversy and, at least for trans fatty acids and heart disease, are near the limit of what reasonably could be expected of advertisers given the state of scientific knowledge at the time of the test.

The strictest qualifying language, specific to antioxidant vitamin supplements, is shown in Figure 5. This "very highly qualified claim" adds the previously discussed reference to the Finnish study results that showed a positive relationship between beta carotene supplementation and the risk of lung cancer among smokers.

Finally, Figure 6 presents an actual print ad for "Nature Made Antioxidant Formula" that functioned as a kind of reality check in our study. This ad allowed us to compare consumer interpretations of and reactions to the fictitious and somewhat rough test ads with those obtained using a professionally produced commercial advertisement that made similar kinds of claims for antioxidant vitamins.

At the time of the test, very few advertisers of vitamin supplements were making explicit health claims. The few such ads that did exist all employed lengthy small-print texts that did not promise to hold the dedicated attention of respondents intercepted in shopping malls. Of this limited group, the Nature Made ad was considered the most appropriate for our purposes, since it tracked the general message presented in our ACE vitamin ads (albeit in far more detailed fashion).

Fortuitously, the ad specifically references the opinions of "scientists," which provided a convenient premise for our close-ended question that asked respondents to rate how sure scientists were about the health benefits of antioxidant supplements. Like our ads, the Nature Made ad also cites the benefits of foods rich in antioxidants, and cautions that "studies haven't proven that vitamin supplements offer the same protection as foods rich in antioxidants." The principal difference between our qualified ads and the real Nature Made selection is the latter's explanation of the hypothesized role free radicals play in cell damage and the development of heart disease and cancer, and the beneficial role antioxidants may play in helping to neutralize free radicals.

D. Questionnaire Design

The survey instrument employed in the Strength of Science segment was very similar to the questionnaire in the other two components. The first portion of the interview gathered information on general ad recall after the first and second exposure, purchase interest, perceived healthiness of the advertised products, and ad believability. These inquiries led to the principal close-ended question that asked respondents to rate how sure they thought scientists were about the cancer reduction benefits of antioxidant vitamin supplements or the cholesterol-raising impact of trans fatty acids.

The vitamin supplement and margarine surveys both contained questions that tested whether viewers of any of the ads took away an exaggerated impression of the overall healthiness of the products. For vitamins, the issue was whether respondents would tend to view the supplements as "magic pills" that could, at least in part, replace fruits and vegetables and greatly reduce the risk of cancer. Respondents were asked to rate the healthiness of the advertised vitamin supplement in comparison with fruits and vegetables, and were also asked to indicate the amount by which they thought the supplements would reduce the risk of cancer.

In the case of margarine, the concern was that respondents would interpret the "no trans fatty acids" claim as an indication that Better Blend was also free of or low in total fat. Although some soft "tub" margarines may in fact be low fat or even fat free, the pictured product was in stick form and therefore should not necessarily have been construed as markedly different in total fat content than other stick margarines. This issue was probed by asking respondents to rate the

healthiness of Better Blend in relation to other margarines and in relation to other foods in general, and to indicate the level of fat contained in the advertised margarine.

E. Copy Test Results

1. Overview

The Strength of Science component of our research generated clear evidence that the types of disclosures included in the Highly Qualified vitamin and margarine ads can alter respondents' perceptions of the level of support for a health claim that has not yet attained significant scientific agreement. Respondents in these cells gave significantly lower ratings of scientific certainty than did viewers of the unqualified health claims. Although mean certainty ratings among respondents exposed to the milder disclaimers in the Qualified ads also were lower than in the corresponding Proof cells, these differences did not achieve statistical significance.

In absolute terms, the results revealed that relatively few consumers in any of the qualified cells took away a clearly deceptive impression of the level of scientific agreement concerning the relevant health claim. For example, only about 12% of the vitamin respondents exposed to the qualified claim (which was still fairly positive in tone) interpreted the ad to mean that scientists were highly certain about the ability of antioxidant vitamin supplements to lower the risk of cancer.

Our research did not find that the vitamin or the margarine ads had a significant impact on actual consumer beliefs concerning the health benefits of antioxidant vitamin supplements or products that are free of trans fatty acids. Nor could we find any evidence that demographic factors or nutrition knowledge affected respondents' reactions to the test ads.

2. Results in Detail

Due to the close correspondence between the vitamin and margarine questionnaires, results for the two sets of ads will be presented simultaneously for each question.

a. Main Points of the Ad, First Exposure

All respondents in the Food Copy Test were given two opportunities to report what they considered to be the main points of the test ads--once after an initial exposure with two clutter ads, and a second time after respondents read the ad again in isolation. For the Strength of Science ads, the key issues at this early point in the test were the proportion of respondents that would recall a health claim, and, more important, the number of respondents within this group that would describe the claim in qualified terms rather than as an established fact. For the margarine ads, an additional issue was whether readers would correctly recall the specific term "trans fatty acids," or whether they would misconstrue the ads as referencing fat in general.

Responses to the first exposure should be approached with lowered expectations, since most of the test ads contained fairly complex messages and respondents were unlikely to focus intently on all three of the ads they were initially shown. Nevertheless, a considerable majority of

the vitamin respondents appear to have volunteered some form of health claim as a major ad message of all but the control ad. The precise percentages cannot be easily determined, since individual respondents may very well have identified more than one health message.³² At the upper limit, however, from 85% (Proof claim) to 98% (Extremely Qualified claim) of the vitamin respondents mentioned a health claim, and of these most tied the health benefit specifically to cancer. Half of the viewers of the real Nature Made ad reported a health claim, but only one-fifth of these referenced cancer in any way. There is, in fact, only one explicit mention of cancer in that ad, and it occupies a very small fraction of the lengthy text.

Health claims were reported less frequently by initial readers of the margarine ads. With the exception of the control ad, which generated only very general references to health from about 13% of respondents, from 40% to 50% of the margarine respondents cited health in some manner; about half of the mentions were specific references to the adverse impact of trans fatty acids on cholesterol.

Most of the margarine respondents (from 53% to 70%) merely cited the absence of trans fatty acids or some other fat as the main point of the margarine ads. Slightly more than half of these respondents identified trans fatty acids correctly by name. The remaining mentions were evenly split between "fat" and "fatty acids." About 12 percent of respondents stated explicitly that Better Blend was low in fat or contained no fat.

Figures 11 and 12 provide a preliminary indication of the impact of the various levels of qualifying language on respondents' perception of the level of proof supporting the vitamin and margarine health claims. To generate these bar charts, coding categories relating to all but the most general health claims were classified as either unqualified or qualified. A response category was classified as unqualified if it pertained to definite statements such as trans fatty acids "raise" cholesterol, "cause" heart disease, or, more generally, "are unhealthy." A response category was considered qualified if it related to indefinite statements, such as antioxidant vitamins "may" or "might" fight cancer, or to explicitly negative characterization of the evidence, such as "studies are contradictory," or "nothing is proven yet."

³² Since the major codings of the verbatim responses for all of the open-ended and follow-up questions in the Food Copy Test were performed by the project contractor, it was not possible to determine the authorship of the verbatims in any of the coding categories, and therefore not possible to tell whether the same respondent had played back multiple messages that appeared in separate categories.

FIGURE 11

MAIN POINTS, FIRST EXPOSURE

VITAMINS

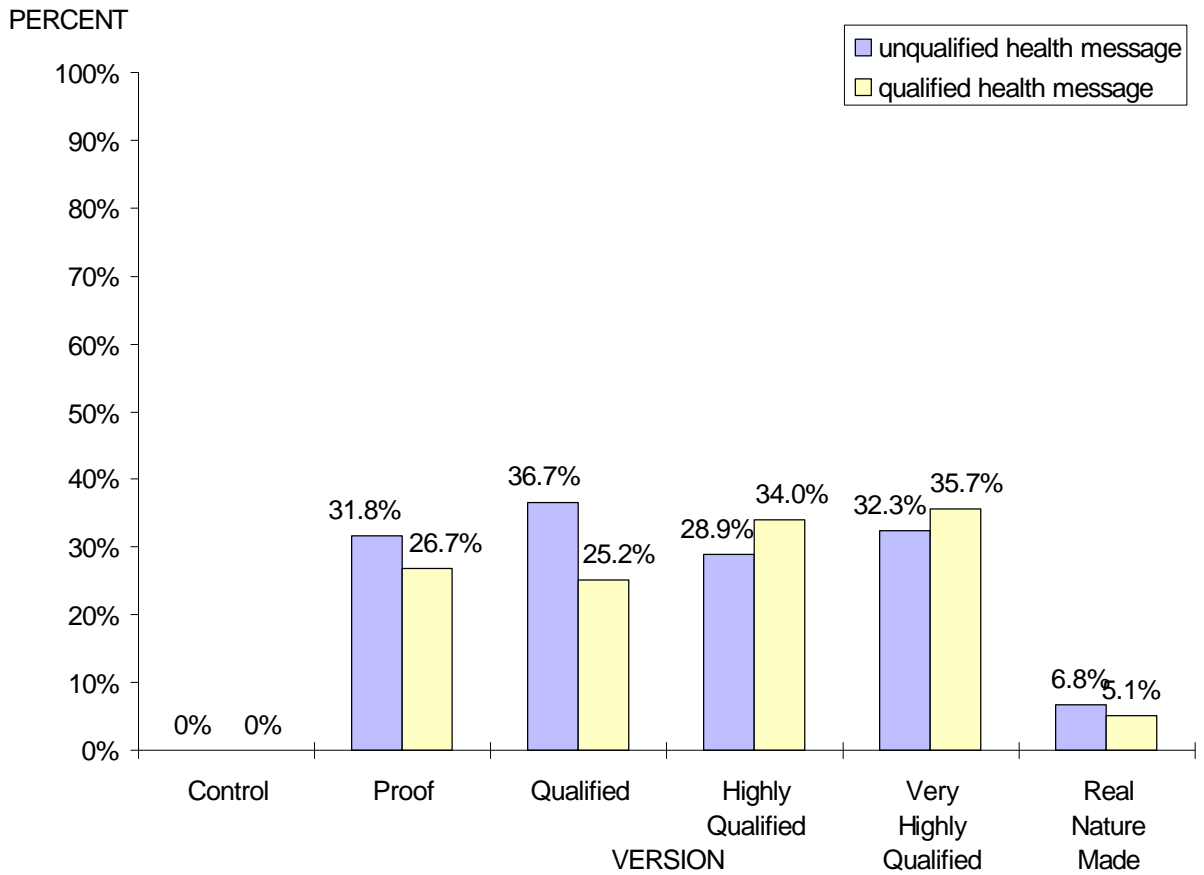
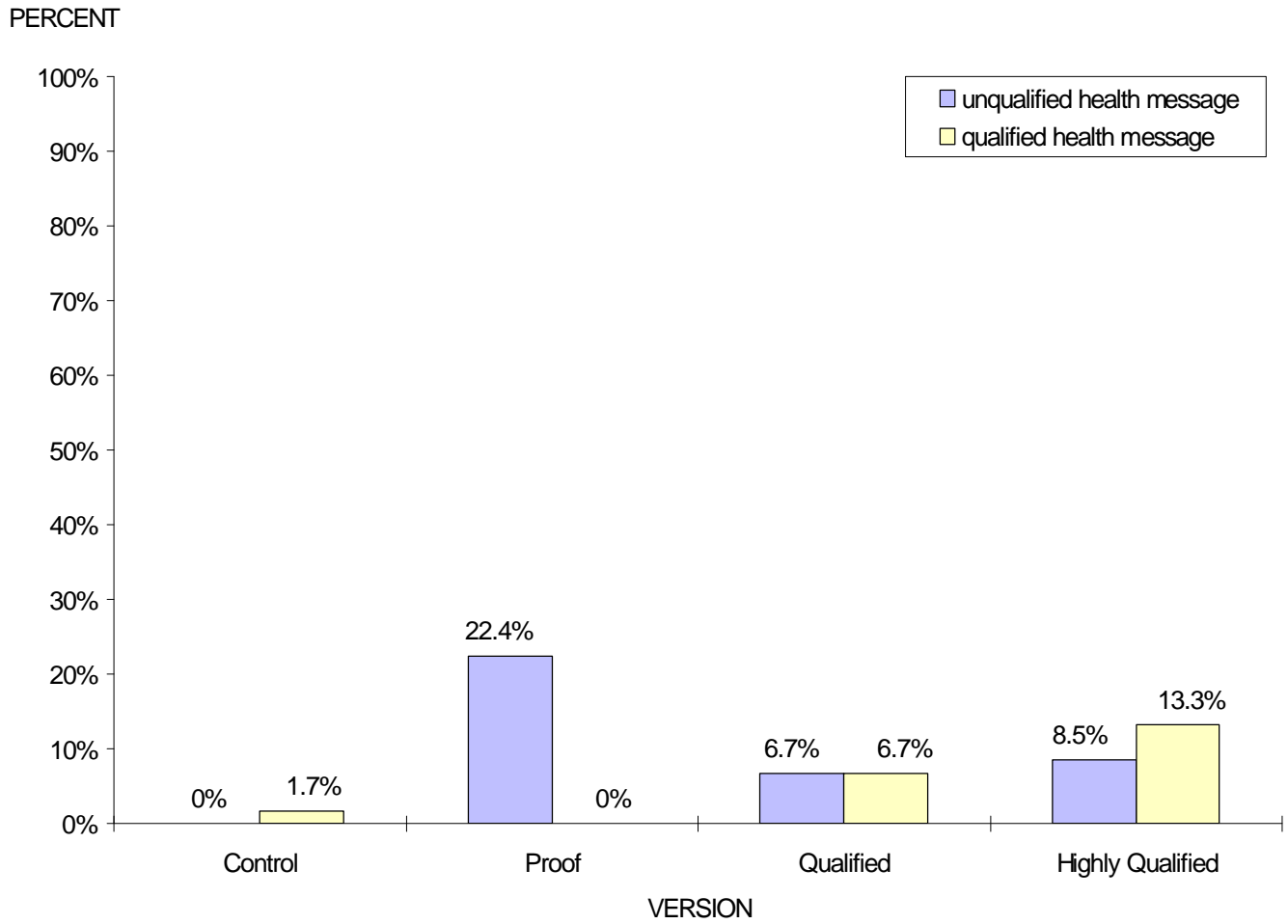


FIGURE 12
MAIN POINTS, FIRST EXPOSURE
MARGARINE



The base for the percentage calculations is the entire cell (approximately 60 respondents). Thus, the percentages vary across cells both because of differences in the distribution of qualified vs. unqualified claims, and because of differences in the proportion of respondents that volunteered a health claim as a main ad message. Note that the percentages for the Nature Made Vitamins cell and the control ads are dramatically lower than any of the other cells, since so few readers of these ads mentioned a specific health claim after the initial exposure.

The only pattern in Figure 11 (vitamins) that is consistent with the expected impact of the qualifiers is a general rise in the percentage of qualified claims as the level of qualification in the ads increases. There is, however, no corresponding fall in the percentage of unqualified claims.

Figure 12 (margarine) is more consistent with expectations. Although the absolute number of reported health claims is lower than in vitamins, there is a large and steady decline in the incidence of unqualified claims as we move from the proof claim to the qualified and highly qualified claims. Interestingly, no respondents reported a qualified health message in the proof claim cell; these claims become slightly more evident in the two qualified cells.

b. Main Points of the Ad, Second Exposure

Not surprisingly, health claim messages figure even more prominently in the verbatims that followed the second reading of the vitamin and margarine ads. With no adjustment for double counting due to multiple responses, the percentage of respondents reporting a health claim for the ACE vitamin test ads varied from 90% to 100%. (This percentage was 18% in the ACE control cell.) Specific references to cancer ranged from 65% in the qualified cell to 77% in the proof cell. Seventy-five percent of respondents viewing the real Nature Made ad a second time cited a health claim as a main message, although once again the claims were very general and almost never cited cancer specifically.

Unlike the initial verbatims, more margarine respondents reported a health claim after the second reading than offered a simple ingredients claim, such as "no trans fatty acids." Health claims ranged from 72% - 80% in the three test cells. (The control cell figure was 15 percent.) Within the category of ingredients claims, the plurality of respondents--41%--referenced the complete phrase "trans fatty acids," rather than "fatty acids"(24%), or "fat" (only 8%).

Figures 13 and 14 illustrate cell-to-cell variations in the incidence of qualified vs. unqualified expressions of the health benefits of the vitamin and margarine products after a second exposure. Two elements stand out in the vitamin results. First, as would be predicted, Figure 13 shows that the highest proportion of unqualified claims--about 37%--occurs in the Proof cell. Although the occurrence of these claims does not fall steadily thereafter as the level of qualification increases, the lowest proportion of unqualified claims does occur in the most qualified cell (17%).

Second, again as predicted, the incidence of qualified claims for vitamins is, by a considerable margin, highest in the very highly qualified cell. Almost half of the viewers of this ad tempered their remarks with some form of qualification. An examination of the codings for this group shows that most of the increase in qualified statements occurs in the sub category

FIGURE 13
MAIN POINTS, SECOND EXPOSURE
VITAMINS

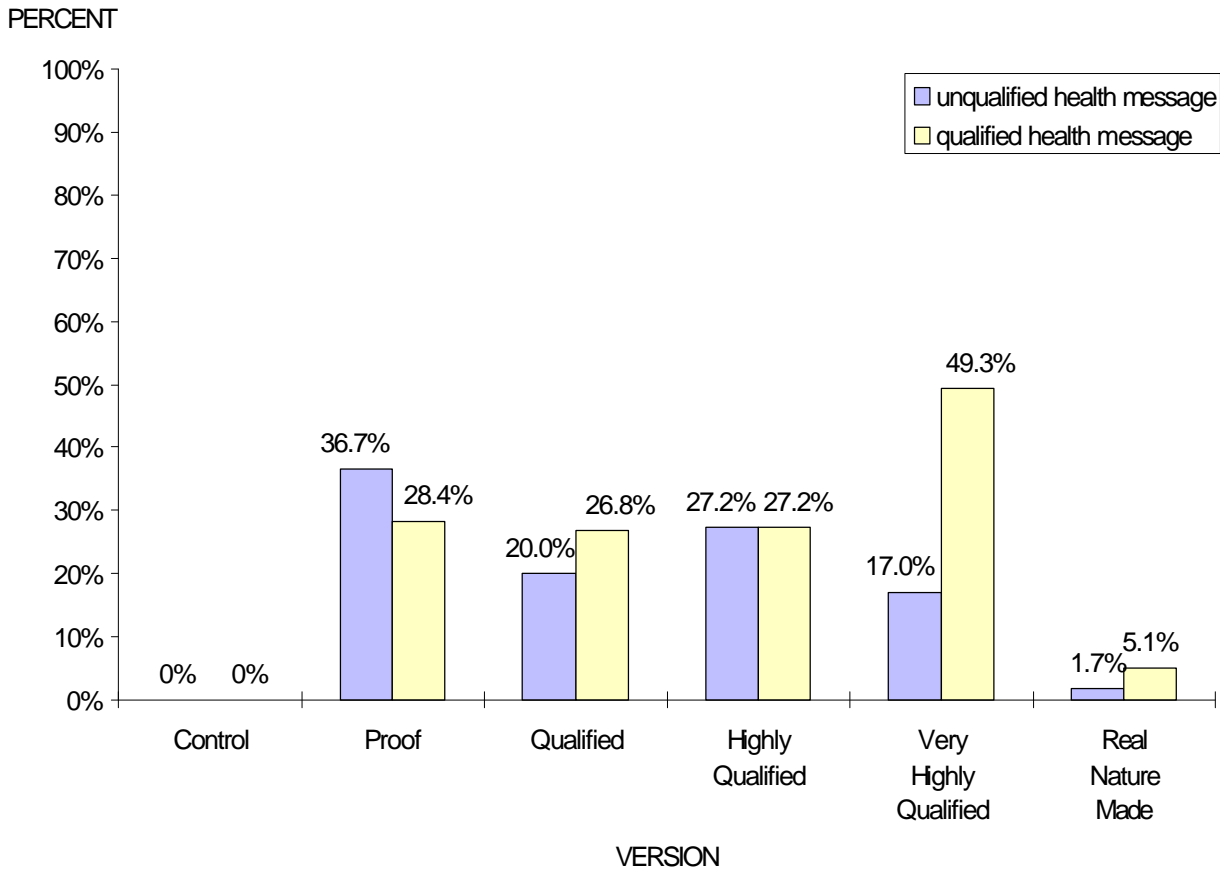
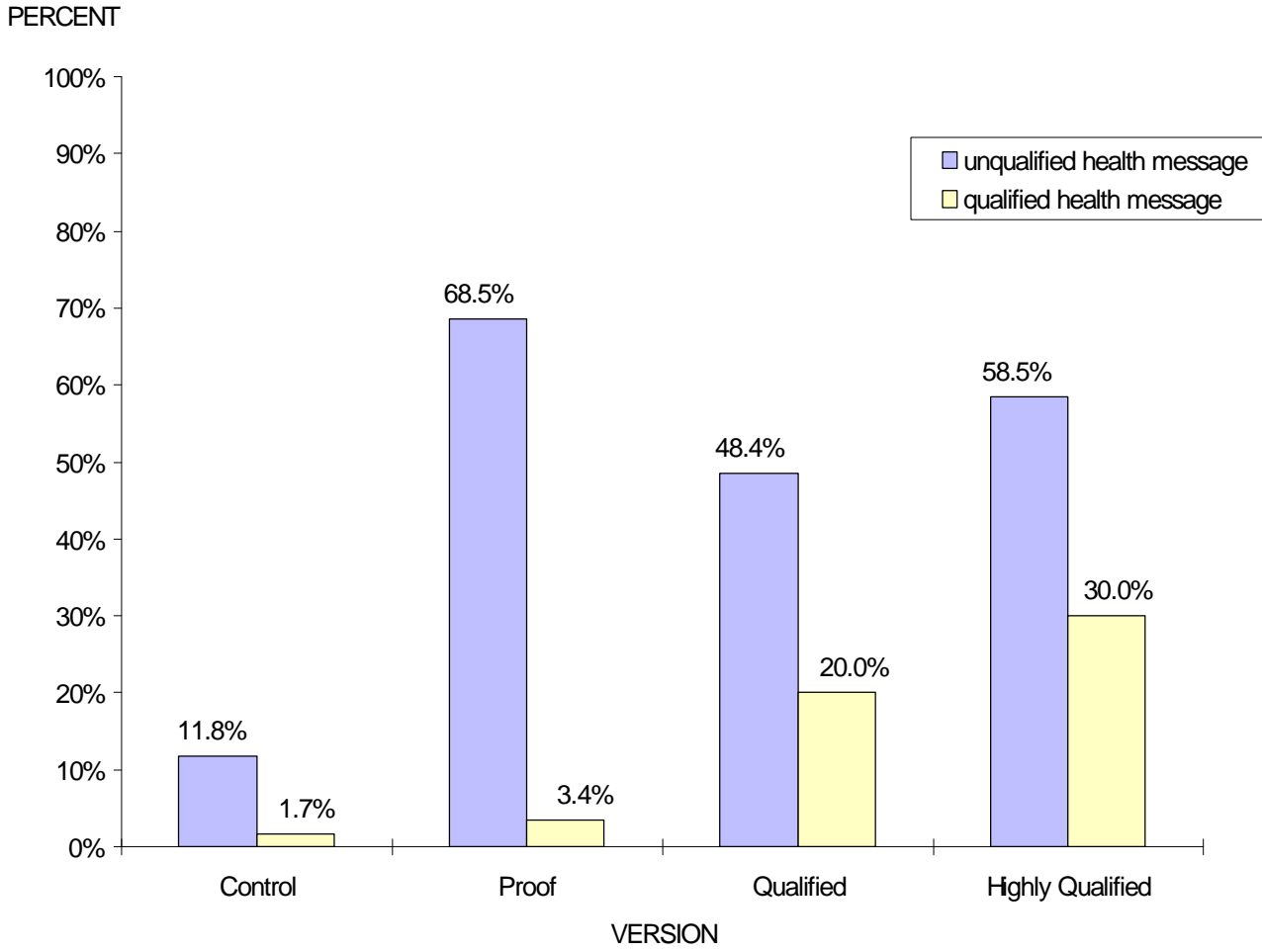


FIGURE 14
MAIN POINTS, SECOND EXPOSURE
MARGARINE



"studies are contradictory." This is probably a reference to the perverse result for Finnish smokers that was mentioned only in the very highly qualified ad. There is no clear pattern in the percentage of qualified claims across the proof, qualified, and highly qualified cells.

The pattern of qualified and unqualified statements is quite similar in the margarine verbatims (Figure 14). Again, the highest playback of unqualified claims occurs in the Proof cell. Anomalously, however, more viewers of the highly qualified ad gave unqualified responses than did respondents in the qualified cell. The pattern of qualified claims conforms completely to expectations. Almost no one in the Proof cell qualified their response. The incidence of qualified claims then increases to 20% in the qualified cell and to 30% in the highly qualified cell.

c. Purchase Interest

This close-ended question, which appeared in all of the Food Copy Test questionnaires, provides a convenient summary measure of the overall impression, either positive or negative, that respondents received from the various ads. The question asked respondents to rate their degree of interest in purchasing the advertised product on a 5-point scale. Presumably, the strongly worded proof claim that promised established health benefits for ACE Vitamins and Better Blend Margarine would increase purchase interest when compared with the ratings generated by the control ads, which contained no health claims. Mean purchase interest should then decline across the remaining cells as the claimed benefits appear less and less certain.

Figures 15 and 16 reveal that this pattern did in fact materialize, although most of the differences between cells are not statistically significant. For vitamins, only the difference between the proof rating of 2.92 and the very highly qualified score of 2.38 comes close to significance ($p=.16$). The results are stronger for the margarine ads, where the difference between the proof and highly qualified scores is significant at the .02 level, and the corresponding difference between the control and highly qualified scores verges on significance ($p=.11$).

One interesting outcome, common to both sets of ads, is the slight decline in purchase interest between the control and qualified cells. This suggests that consumers may be at best indifferent between products presented in an ad that provides no specific health information and one that is more informative but raises even an element of doubt about the underlying evidence. Considering that both the vitamin and margarine qualified ads were intended to strike a positive tone, this result raises questions about the incentives advertisers have to provide qualified health claims.

FIGURE 15

PURCHASE INTEREST

VITAMINS

MEAN RESPONSE

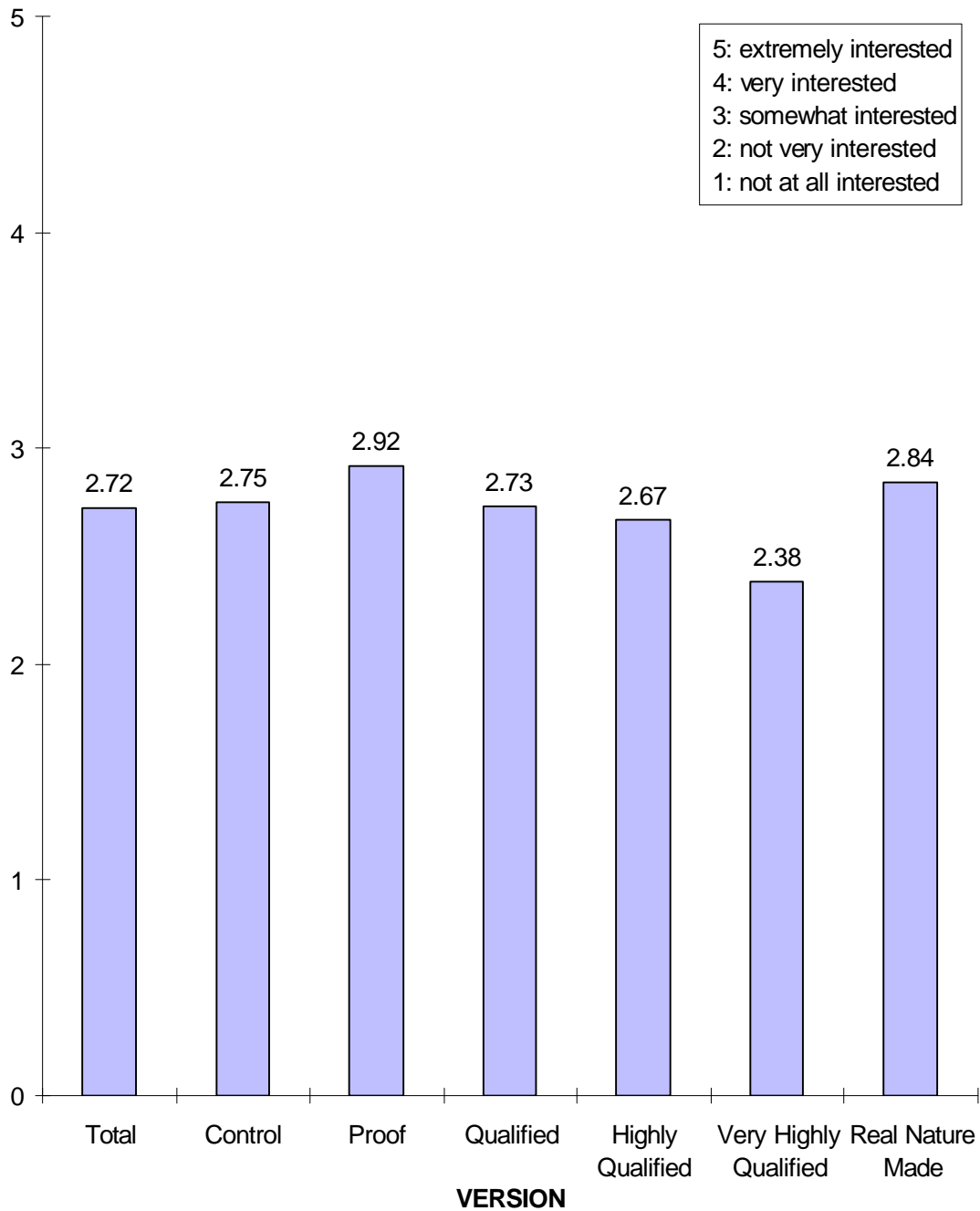
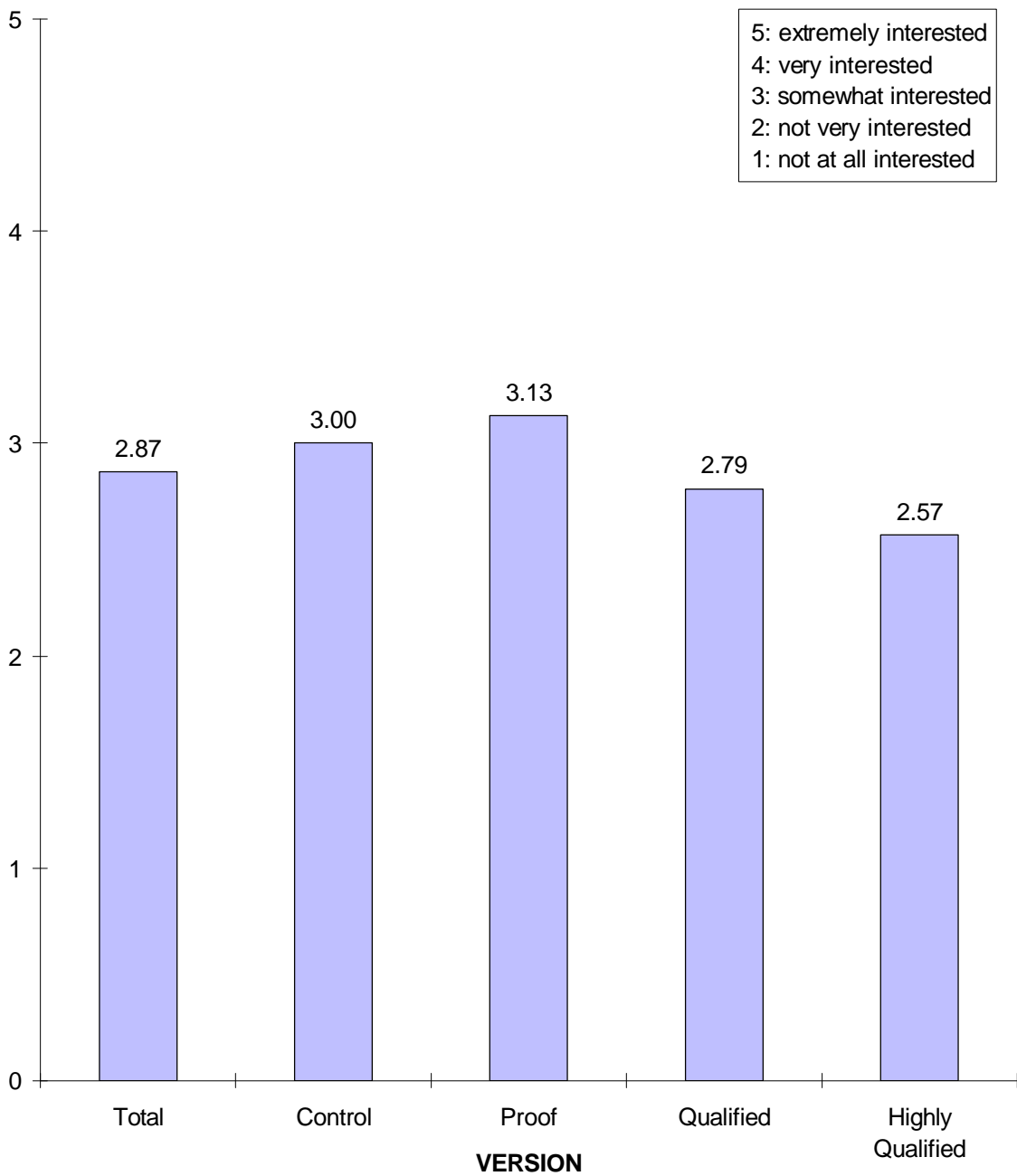


FIGURE 16

PURCHASE INTEREST

MARGARINE

MEAN RESPONSE



In terms of the absolute level of purchase interest displayed by respondents in general, it is clear from Figure 15 that viewers of the vitamin ads were less than enthusiastic about ACE. The overall mean rating of 2.69 tied with Matterhorn Swiss Cheese Slices for the lowest rating accorded any of the products in the Copy test. This does not appear to be due to any glitch in the name or presentation of ACE Vitamins, since the real Nature Made ad generated a rating of only 2.84 (half way between the Proof and Qualified ratings for ACE). One explanation for the lackluster showing, cited in 13 percent of the verbatims, is that respondents were already satisfied with the antioxidant vitamins they were taking.

Respondents were more interested in sampling Better Blend Margarine. Its overall purchase interest score of 2.87 equaled the rating for Rich and Hearty Soup, although it was still lower than the 3.04 registered for Right Slices Turkey.

d. Effect on Health

The fourth question, which asked respondents whether the advertisement stated or implied that the vitamins or margarine would have any effect on health, was intended to focus attention on the health claims and qualifications in the test ads. The primary item of interest was the response to the follow-up probe, which asked respondents to explain what effect the product would have on health. The verbatims supporting the affirmative responses revealed a pattern of qualified and unqualified characterizations very similar to that portrayed previously in Figures 13 and 14, and will not be discussed further here.

e. Ad Believability

Respondents in all of the Food Copy test segments were asked to rate the believability of the test ads on a 5-point scale that ran from "not at all believable" to "extremely believable." This question served two very specific purposes in the Strength of Science segment. As indicated, one of the vitamin test ads was a real ad for an actual vitamin supplement. The results of the believability question therefore allowed us to determine whether respondents would consider a professionally produced ad for a possibly familiar product significantly more credible than the cruder computer mock-ups and fictitious products used in the remainder of the study. This question also allowed a test of the hypothesis that consumers would find qualified claims more credible than an absolute proof claim.

The results are presented in Figures 17 and 18. There is a rather remarkable lack of variation across cells within either the margarine or vitamin groupings. This uniformity is sufficient to reject the hypothesis that qualification per se will lend credibility to an ad. In addition, the similarity between the believability scores of the fictitious vitamin ads and the real Nature Made ad indicates (reassuringly) that respondents saw nothing particularly problematic in the format or content of the our computer mock-ups.

For reasons that are not readily apparent, respondents as a group found the vitamin ads more believable than the margarine ads. The mean vitamin score of 3.35 essentially replicated the ratings recorded for the Matterhorn Swiss Cheese and Right Slices Turkey ads, and exceeded the 3.21 score for Rich and Hearty Soup. The overall score of 3.09 for Better Blend Margarine

FIGURE 17

AD BELIEVABILITY

VITAMINS

MEAN RESPONSE

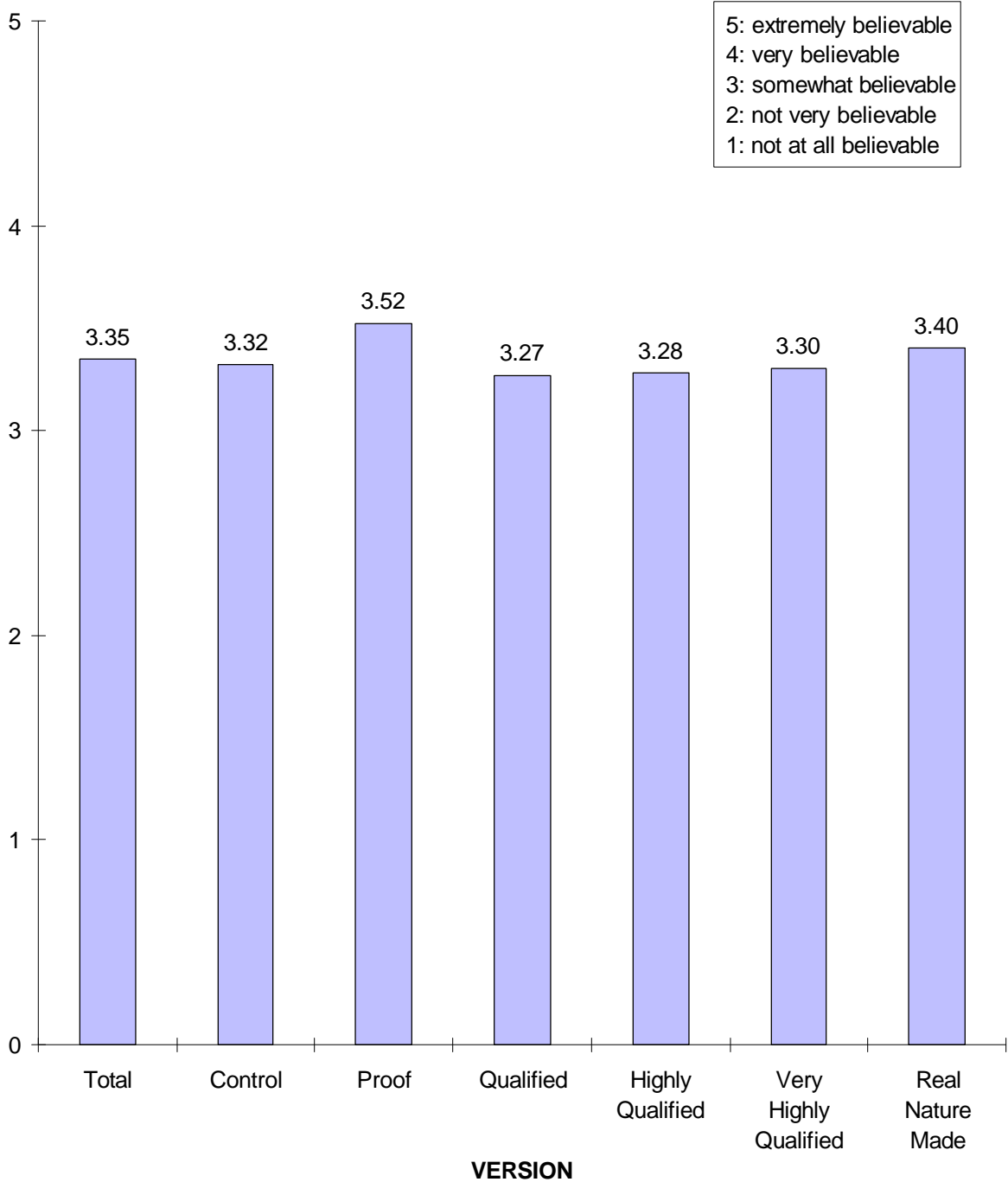
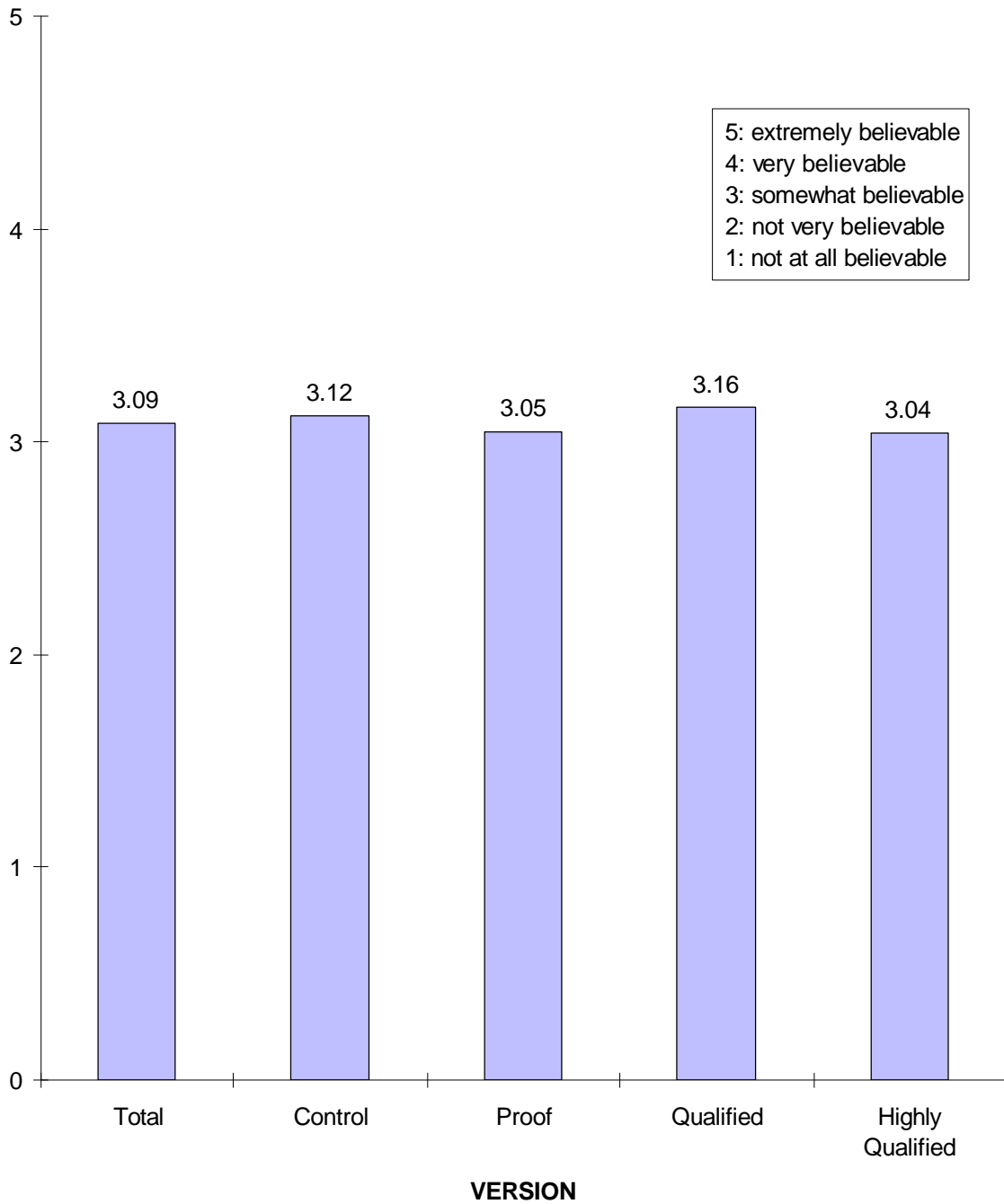


FIGURE 18

AD BELIEVABILITY

MARGARINE

MEAN RESPONSE



was the lowest in the food Copy test. The verbatims explaining respondents' responses do not point to any particular deficiency in the presentation of the Better Blend ads.

f. Healthiness of Advertised Products

The next section of the vitamin and margarine questionnaires contained a series of questions that asked respondents to rate the healthiness of the advertised products in a variety of comparative contexts. The vitamin respondents first rated the absolute healthiness of ACE on a 5-point scale with no frame of reference specified. This question was intended to reveal whatever impact the varying levels of qualification had on respondents' perceptions of the overall healthiness of ACE.

Respondents were then asked to evaluate the healthiness of ACE in relation to fruits and vegetables using a 5-point comparative scale (ranging from "much less healthy" to "much more healthy"). As discussed earlier, this question was an attempt to determine whether the ads were leading respondents to regard ACE as a "magic pill" that could allow a substantial reduction in the dietary intake of fruits and vegetables.

The corresponding margarine sequence began by asking viewers of the Better Blend ads to rate the product's healthiness in relation to other brands of margarine, which was the frame of reference presented in the ads. This question explored the impact of the differing levels of qualification on respondents' appraisal of the healthiness of Better Blend as a margarine product. Respondents were next asked to rate the healthiness of Better Blend in relation to other foods. If viewers of the ads thought that Better Blend contained appreciable levels of total fat, the ratings obtained from the previous question should now be displaced downward to reflect the broader comparative framework.

The first set of healthiness questions produced a very similar pattern of responses for vitamins and margarine. When respondents were asked to rate the overall healthiness of ACE or the healthiness of Better Blend in relation to other margarines, the healthiness ratings were highest in the proof cell, and then declined slightly (and insignificantly) in the qualified and again in the highly qualified cells. The principal distinction was the additional and more precipitous decline that occurred in the very highly qualified vitamin cell. The difference between the healthiness rating for that ad (3.06) and the rating in the proof cell (3.69) is highly significant ($p=.003$). The results for the vitamin ads are illustrated in Figure 19.

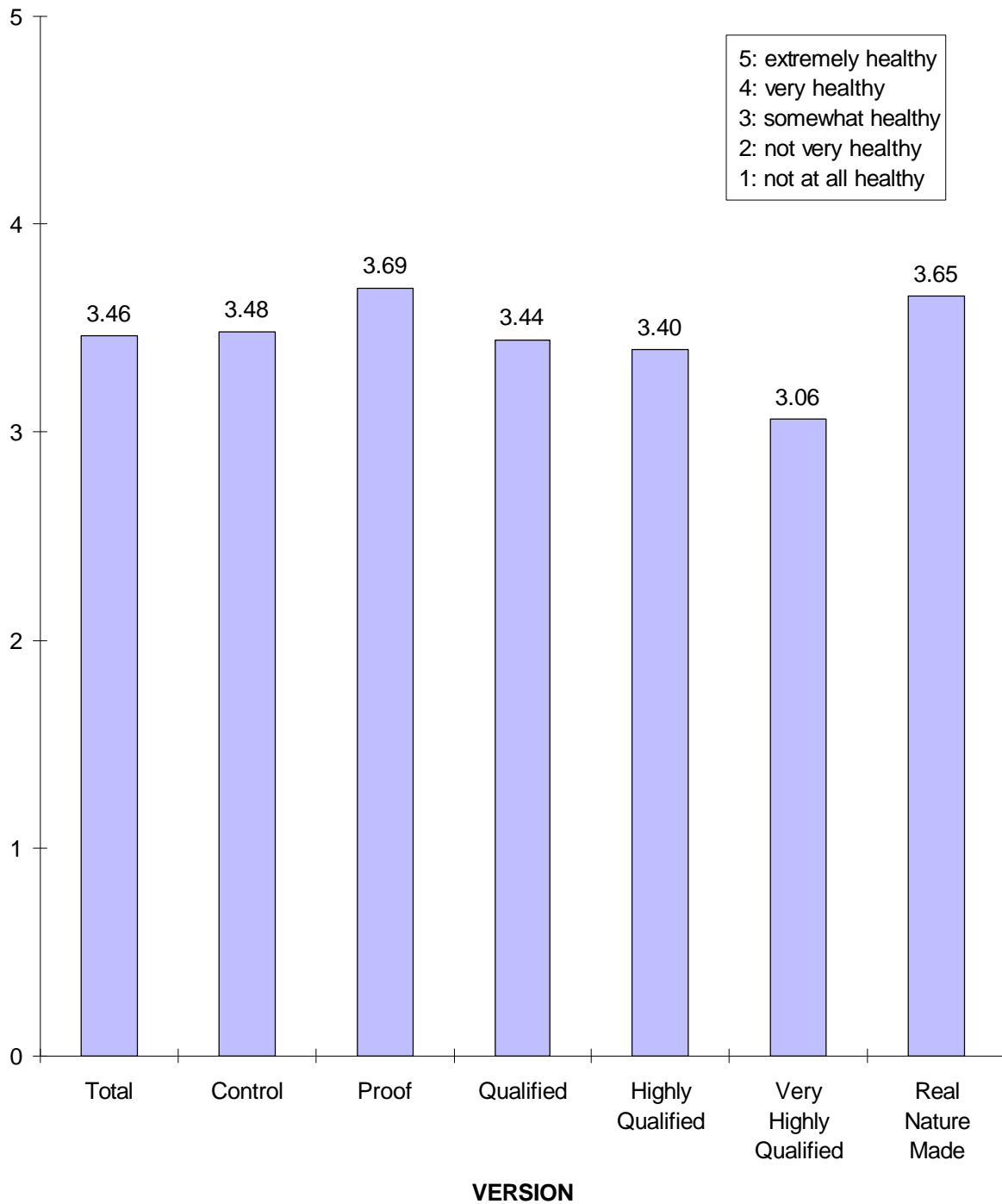
The next question in the vitamin sequence did not find strong support for the theory that respondents would regard ACE as a magic pill. When asked to compare the healthiness of the vitamin supplement and fruits and vegetables, respondents on average placed ACE midway between "less healthy" and "about as healthy."

Results from the second healthiness question in the margarine sequence showed that Better Blend was not considered as healthy in relation to foods in general as it was in relation to other brands of margarine. The mean rating fell from 3.64 to 3.21. This score still places Blend as above average in healthiness, which might indicate that respondents did not think that the product was very high in total fat. This conjecture was confirmed in a later question that asked

FIGURE 19

HEALTHINESS OF ADVERTISED VITAMINS

MEAN RESPONSE



respondents to rate the fat content of Better Blend on a 5-point scale, ranging from "fat free" to "high." The mean response to that question was 2.56, which is half way between "somewhat low" and "neither high nor low." Although it is plausible, and even likely, that the "no trans fatty acids" claim influenced this outcome, the absence of a true tombstone control that contained no ingredient claims prevents a definitive conclusion on this point.

g. Certainty of the Science

Following the queries concerning the perceived healthiness of the advertised products, Question 8 in the vitamin and margarine questionnaires focused on the primary issue in the Strength of Science segment. The specific wording of this question was as follows:

Q8(Vitamins): Based on what the ad says or implies, how sure are scientists that taking antioxidant vitamin supplements will reduce the risk of certain kinds of cancer?

Q8(Margarine): Based on what the ad says or implies, how sure are scientists about whether eating foods that contain trans fatty acids will increase the risk of heart disease?

The 5-point response scale ranged from "very unsure" to "very sure."

The results of the two key questions are displayed in Figures 20 and 21. The pattern of responses among the vitamin and margarine respondents is virtually identical across the 4 common cells. For both products, the certainty ratings are highest in the Proof cell, and then fall continuously in the Qualified and Highly Qualified cells. For vitamins, the differences between the ratings in the Highly Qualified and Proof cells and the Very Highly Qualified and Proof cells are highly significant ($p=.002$ and $.006$, respectively). There are, however, no other significant differences between any of the four scores.

In the case of margarine, the difference between the ratings in the Proof and Highly Qualified cells is also highly significant ($p=.003$). The corresponding difference between the Proof and Qualified cells just misses significance at the 10% level ($p=.13$). There are no other significant differences between any of the margarine cells.

A comparison of the mean ratings for the two control ads, which did not contain explicit health claims, suggests that the margarine respondents were more confident about the relationship between trans fatty acids and heart disease (mean = 3.67) than the vitamin respondents were about the power of antioxidant vitamin supplements to reduce cancer risks (mean = 3.17.) Only about half of the respondents in these cells gave certainty ratings, however, and the difference is not statistically significant. Still, the ratings for all of the margarine cells are consistently higher than the corresponding vitamin ratings, which suggests that respondents found the claimed health benefits for Better Blend somewhat more plausible than the claimed

FIGURE 20

HOW SURE ARE SCIENTISTS?

VITAMINS

MEAN RESPONSE

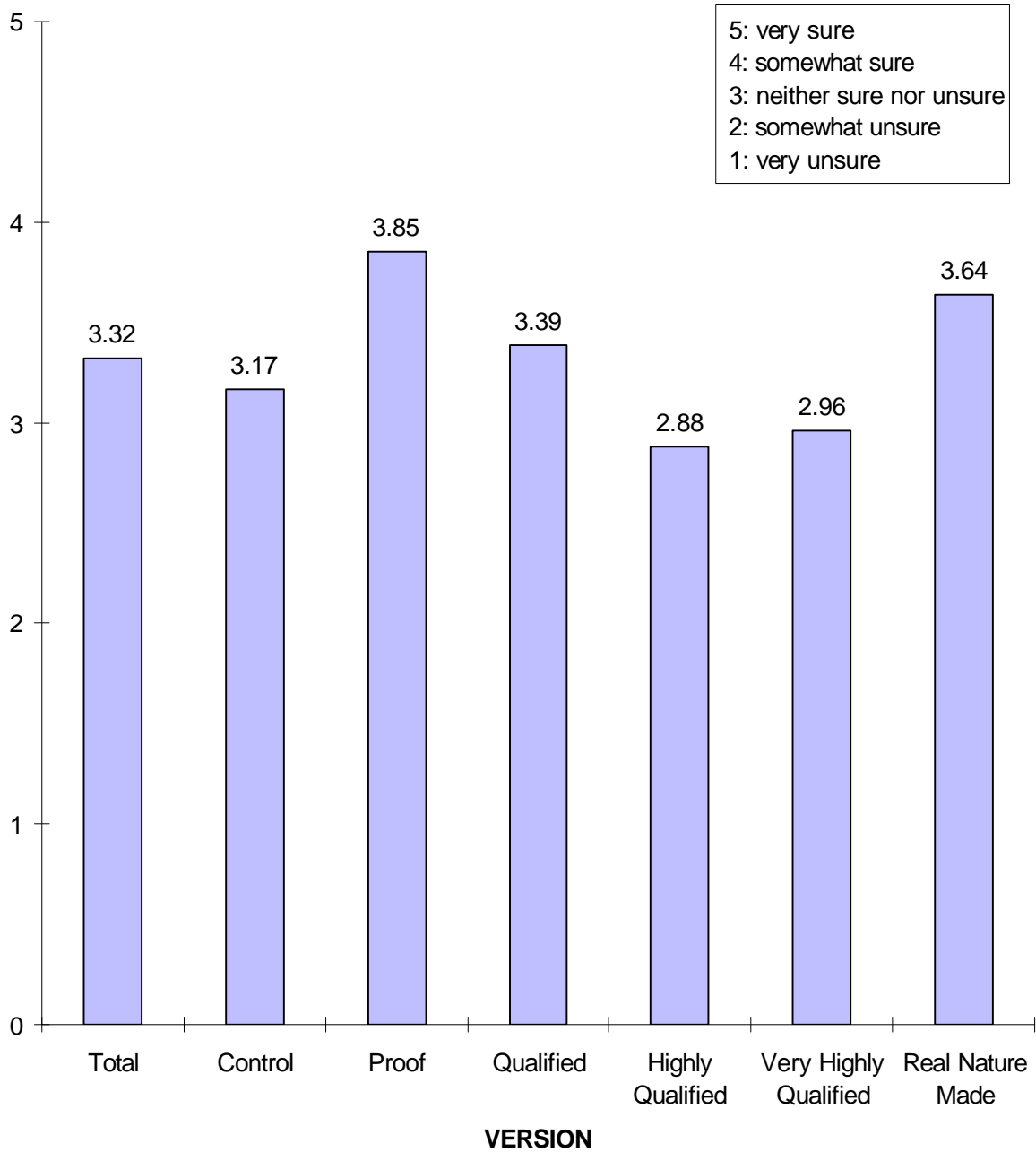
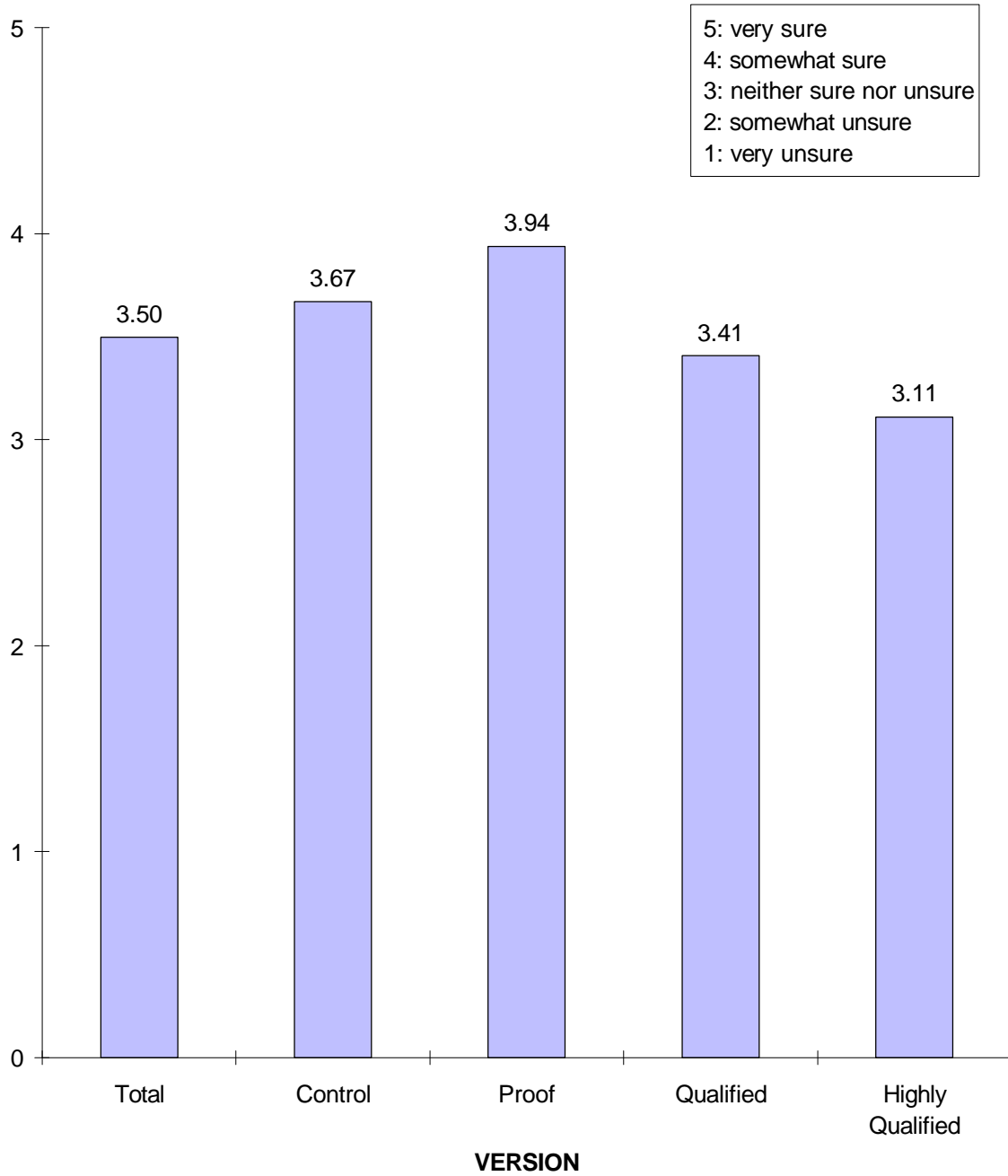


FIGURE 21

HOW SURE ARE SCIENTISTS?

MARGARINE

MEAN RESPONSE



cancer benefits for ACE. This was an unanticipated result, given the considerable attention that antioxidant vitamins have received in the popular press

Another interesting result is the relative reluctance of respondents in either of the proof cells to assign very high levels of certainty to the science supporting the antioxidant-cancer or trans fatty acids-heart disease relationships. In both cases, the mean rating is slightly below "somewhat sure," which suggests that many respondents either did not believe the strongly worded establishment claims, or perhaps were merely exhibiting an aversion to choosing the most extreme options in the provided ratings menu. Only 27% of the vitamin respondents and 31% of the margarine participants chose the "very sure" rating.

The verbatims for the respondents in the vitamin Proof cell suggest that general skepticism about advertising claims explains at least part of the observed outcome. About 20% of the vitamin participants complained that the ad offered no proof or stated that they do not believe everything they read in ads. Practically none of the margarine respondents, on the other hand, volunteered negative comments of this type.

With respect to the verbatims supporting the ratings in the Qualified and Highly Qualified vitamin and margarine cells, the most frequently volunteered remark was some form of reference to the need for more research. Thirty percent of the respondents in the Qualified vitamin and margarine cells and Highly Qualified vitamin cell offered such an explanation for their rating. This figure rose to 39% in the highly qualified vitamin cell.

Figure 20 reveals interesting outcomes for the two vitamin cells that were not included in the margarine sequence--the Very Highly Qualified ad and the real ad for Nature Made Vitamins. Although the Very Highly Qualified ad had scored lowest of all the vitamin ads in the purchase interest and overall healthiness ratings, it registered an (insignificantly) higher scientific certainty score than the Highly Qualified ad. This suggests that respondents did not interpret the advisory concerning the Finnish smokers' higher cancer rates as a caveat about the state of scientific agreement, but rather as a warning about a specific subgroup that might wish to avoid antioxidant vitamins.

Finally, the relatively high certainty rating accorded the real Nature Made ad suggests that the qualifiers in this ad were overwhelmed by the extensive discussion of free radicals and the mechanism by which antioxidants are thought to help fight cancer. The score of 3.64 is significantly higher than the 2.88 rating for the Highly Qualified ACE ad ($p=.05$) and half way between the scores for the Proof and Qualified ACE cells.

The results in Figures 20 and 21 do not lend themselves to definitive conclusions. If the object of our experiment is viewed strictly as determining what degree of qualification is needed to lower consumers' perception of scientific certainty below the level generated by an establishment claim, our results from the vitamin ads indicate that the degree of qualification would have to be quite high. The margarine results suggest that somewhat less strict qualification might be needed in the area of trans fatty acids health claims, although the lower certainty score in the Qualified cell did not quite achieve statistical significance.

If, on the other hand, the goal of our research is viewed as determining what degree of qualification is needed to avoid consumer deception, the results point to much less strict remedies. In the Qualified vitamin cell, only about 12 percent of respondents stated that scientists were very sure that taking antioxidant vitamins would reduce the risk of cancer. Only one respondent in the Highly Qualified cell gave this response. The mean score in the Qualified cell is only about one-third of the way between "neither sure nor unsure" and "somewhat sure," a rating that does not appear to have seriously overstated the actual level of confidence of scientists in this health area at the time the interviews were conducted.

The potential for consumer deception is even less evident in the margarine results. Although the ratings are as a group higher than in vitamins, the actual level of agreement among researchers in the field of trans fatty acids is at least correspondingly higher. Although, as discussed, these scientists may not agree on the magnitude of the problem, there is relatively little doubt that dietary intake of trans fatty acids can increase serum cholesterol significantly. Thus, the mean rating of 3.41 in the Qualified cell does not appear problematic. The upper limit on the incidence of deception in this group would appear to be the 18% of respondents who rated scientists as "very sure."

h. Impact on Beliefs

Although all of the interview questions in the Food Copy test were phrased to measure what respondents thought was said or implied in the ads ("ad conveyance"), a self-administered questionnaire given to all of the respondents following the interview allowed a unique opportunity to determine whether the health claims in the Strength of Science component had altered respondents' *beliefs* about the diet-disease relationships at issue. Specifically, all respondents were asked to indicate on a 7-point scale how strongly they agreed with the statements, "I believe that taking antioxidant vitamin supplements will reduce the risk of cancer," and "I believe that eating foods that contain trans fatty acids increases the risk of heart disease." If any of the ads in the vitamin or margarine sequence changed respondents' opinions about these health relationships, the average belief ratings of the respondents in these cells should diverge from the average ratings of all of the remaining respondents who were not exposed to a health claim concerning antioxidant vitamins or trans fatty acids.

Figure 22 displays for each of the vitamin cells the average agreement rating for the statement concerning antioxidant vitamins and a reduced cancer risk. Also presented is the average agreement rate for all of the remaining non-vitamin respondents in the Copy test. The results show that the vitamin respondents as a group indicated slightly higher agreement with the antioxidants-cancer statement than did respondents not seeing these ads (4.72 vs. 4.60, respectively). The highest degree of agreement was, as would be expected, recorded by the proof cell participants. Even this rating, however, indicates that respondents only "somewhat agreed" with the statement. The ratings then fall in the various qualified cells, though not continuously. Interestingly, none of these ratings falls below the rating of 4.6 recorded by the remaining participants in the Food Copy test.

The mean score of 4.47 for the real Nature Made score is puzzling. Although the explicit reference to cancer is not very prominent, it is not evident why viewers of this ad would be less likely to believe in the cancer benefits of antioxidant vitamins than respondents who saw no claims of any kind about these substances.

The margarine results are, at best, anomalous. As shown in Figure 23, none of the agreement ratings, with the exception of the highly qualified cell, conforms to reasonable prior expectations. Most curiously, the proof claim rating of 5.17 is lower than the mean rating of 5.55 recorded by all other Copy test respondents.

The subsequent rise in agreement among respondents in the Qualified cell is also a bit curious, particularly since this first level of qualification barely missed statistical significance in reducing the scientific certainty ratings from those of the Proof ad. Finally, the fall in agreement in the Highly Qualified cell is of the expected direction, but its magnitude seems large given the more modest cell-to-cell variations in the other vitamin and margarine scores.

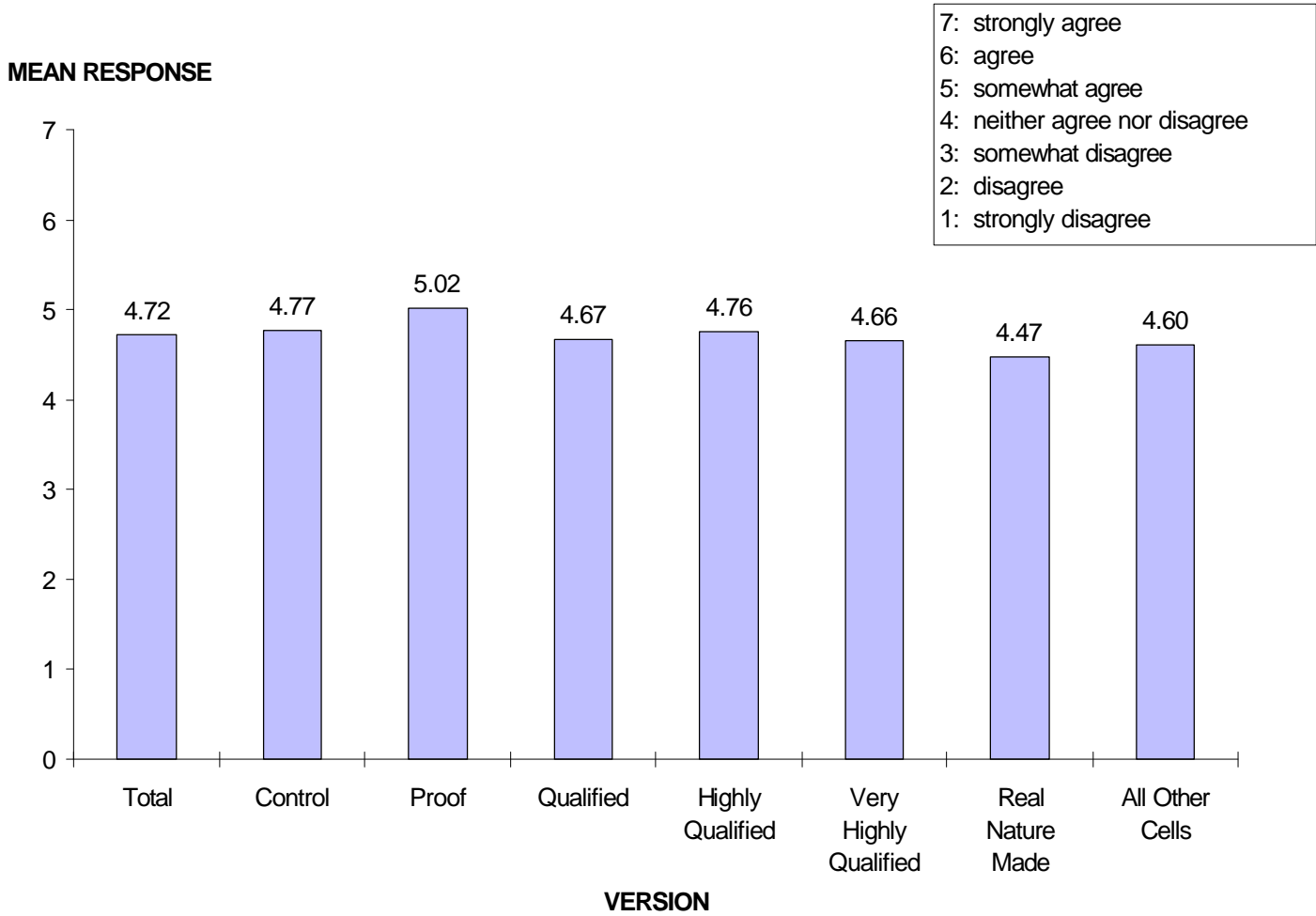
i. Demographic Analysis

There was relatively little variation across cells in the mean education and income levels of either the vitamin or margarine respondents. On an overall basis, the mean education levels of the vitamin and margarine participants were all but identical (slightly above some college or technical school training), as were the corresponding income levels (about \$30,000.)

In addition to the demographic data on education and income, all of the data sets for the Food Copy test also contain information concerning respondents' knowledge and interest in nutrition issues as measured by the self-administered nutrition quiz and attitudes questionnaire that respondents completed following the main interview. The demographic and questionnaire information allowed a test of whether the answers to the key interview question concerning the relative certainty of scientists (Question 8) were related to respondent demographics or nutrition knowledge and interest.

Specifically, a series of ordered logit regressions were run using--both individually and collectively--sex, income, education, and scores on the nutrition quiz and attitudes survey to predict the level of certainty that respondents chose in response to Question 8. In theory, we would expect that more educated, knowledgeable, and interested consumers would be more likely to recognize and understand the various qualifications in the remedy ads. We should therefore observe a negative correlation between the certainty ratings and education, knowledge, and perhaps income and nutrition interest, in the various qualified cells. The results showed no significant relationships of any kind, other than one negative correlation between education and the certainty ratings for the qualified margarine ad.

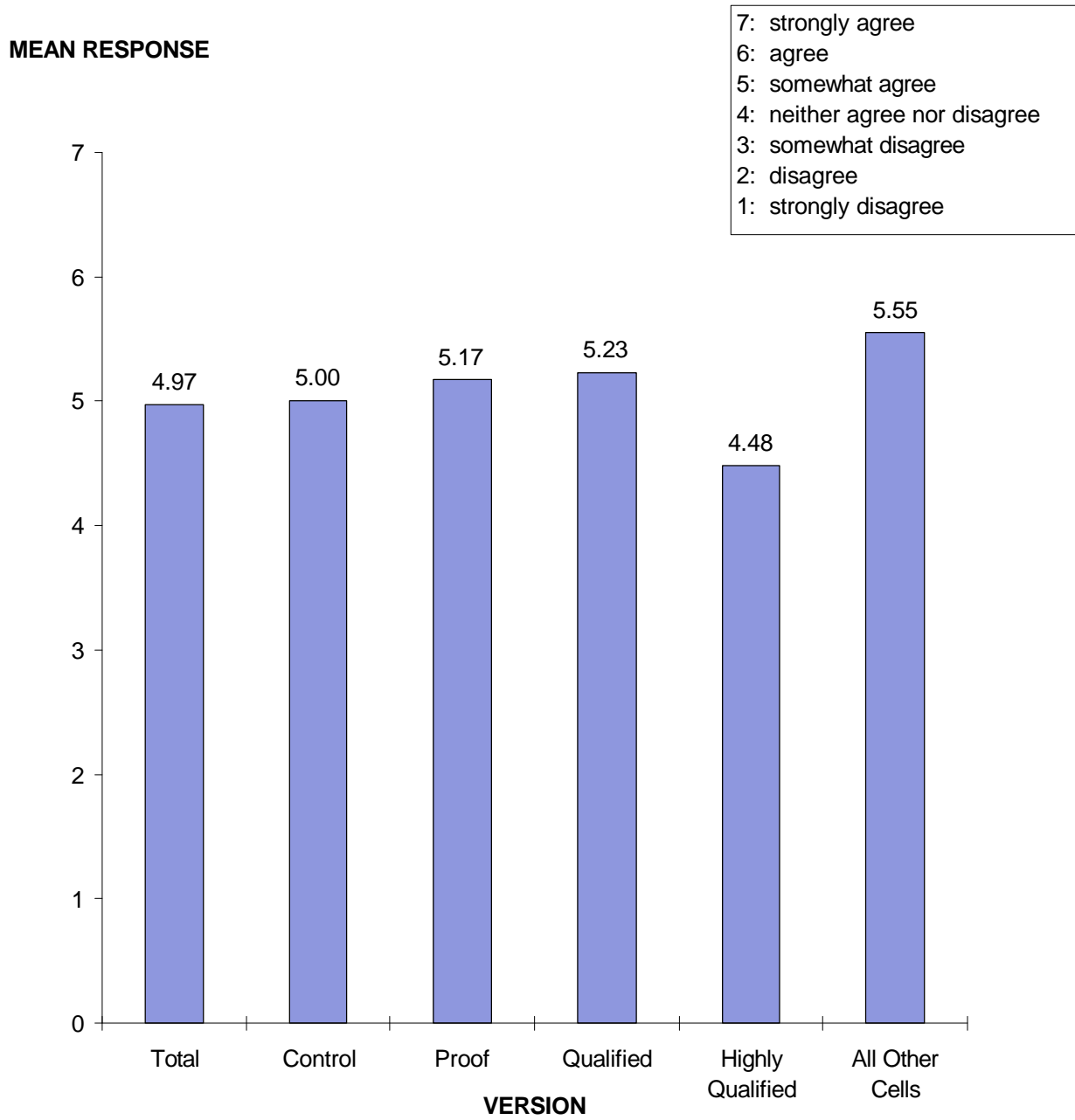
FIGURE 22
ANTIOXIDANT VITAMINS REDUCE CANCER RISK
 AGREEMENT RATES
FIGURE 23



TRANS FATTY ACIDS INCREASE HEART
DISEASE RISK
AGREEMENT RATES

F. Conclusion

The results of the Strength of Science component demonstrate that advertising disclosures can have a significant impact on consumer perceptions of the degree of certainty underlying health claims that have not yet attained a high level of scientific agreement. If we use as a baseline the



mean certainty rating that consumers accord a "proof" or "establishment" claim for a nutrient-disease relationship, our results reveal that these ratings will be reduced significantly by explicit qualifications concerning inconsistencies in study results and the existence of ongoing scientific debate.

The performance of the tested qualifications appears even stronger when viewed within a more traditional deception framework that focuses on the proportion of consumers that receives a deceptive message concerning the degree of scientific agreement for an advertised health claim. In the case of antioxidant vitamin supplements and cancer, only about 12% of respondents viewing ads with the mildest level of qualification rated scientists as "very sure" about the relationship (compared with about 27% in the proof cell). This percentage fell to only 2% in the highly qualified cell. For the trans fatty acids-heart disease relationship (about which scientists are fairly confident), the percentage of respondents reporting a "very sure" rating fell from 31% in the proof cell to 18% in the qualified cell, and then to 10% in the highly qualified cell.

Although our research confirmed the feasibility of qualifying health claims to conform with the guidance provided in the Commission's Policy Statement on Food Advertising, we did not find strong indications that such qualifying language, even in its mildest form, would have significant sales appeal. Viewers of the control ads, which contained no explicit health claims, registered a purchase interest that was at least as high as that recorded in any of the qualified cells.

Finally, the impact of the qualifying disclosures did not appear to depend upon the demographic characteristics of respondents. We found almost no evidence that education, income, sex, interest in nutrition issues, or knowledge of nutrition facts affected respondents' ratings of the degree of scientific certainty supporting the various health claims.