including those from her own observations, those generated by the Trenton staff, and those from the statistical staff in CX 448 (Marcelli, Tr. 17388).

1115. Dr. Marcelli, a graduate industrial pharmacist at the time, was assigned to the 223 Test while working for Dr. Tainter as a Special Project Assistant for the Sterling Research Board. Dr. Marcelli had overall responsibility and supervision for all aspects of the study except the sample pickup. She was responsible for determining the attributes to be tested, conducting analysis of the data, and writing the report (Marcelli, Tr. 17401–02, 17404–05).

1116. It is respondent's position that the 223 Test was a properly conducted study and that its results are valid and reliable. Respondent contends that the study demonstrates that [**275**] Bayer Aspirin was pharmaceutically superior to other aspirin brands on the market, and that said study was, therefore, properly relied upon as a basis for claims of pharmaceutical superiority in the Blue Book campaign. Furthermore, respondent contends that, if an implied representation of therapeutic superiority is found in this case (contrary to its view), the 223 Test, in addition to other evidence, shows a reasonable basis for such representation.

1117. The samples of aspirin tested in the 223 Test were assembled as a result of a survey conducted in 1967 and a pickup of samples in 1968 (CX 448K-L; Alberts, Tr. 8952; Mattimore, Tr. 15336–38). In early 1967, the sales administration manager of respondent's Glenbrook Laboratories Division, Mr. Mattimore, was asked by his superior to conduct a survey to identify the various brands of aspirin which were available for sale in retail outlets in the United States. This was carried out by requesting every Glenbrook Laboratories salesman to report all brands of aspirin they encountered in a one-week period. The salesmen were provided with a form and instructed to record the name of every brand of 5-grain aspirin, and the name and address of the store in which it was found (Mattimore, Tr. 15336; CX 448K).

1118. Mr. Mattimore estimated that approximately 100 men went into approximately 3,350 stores, 35 per week per man. The salesmen were asked to go into units of all the major chains so that their coverage would reflect the brands in all outlets of such chains. In Mr. Mattimore's view, the survey reflected what was being sold in at least 70 or 80% of the stores in the United States at that time (Mattimore, Tr. 15336–38).

1119. In 1968, Mr. Mattimore was asked to obtain samples of the brands which had been identified in the survey. This was carried out in three stages—collection of samples of minor or regional brands, collection of samples of major brands, and collection of Bayer samples (Mattimore, Tr. 15338; CX 448K-L).

1120. In the first stage, in March 1968, the collection of samples of minor or regional brands, Mr. Mattimore assigned salesmen to pick up specific brands by going through the 1967 report forms and identifying salesmen who had reported such brands. Mr. Mattimore testified that this was the only practical way to proceed because store or private-label brands are commonly available only in particular outlets and that it was necessary to assign the task to the personnel who had located the brand initially. Otherwise, many of the brands identified the previous year would have been missed. Where a brand was clearly available in more than one location, *e.g.,* certain chain store brands, several salesmen were asked to provide samples (Mattimore, Tr. 15338–41). In the sample collection, each salesman was asked to pick up six samples of a particular brand, of different control numbers where possible (Mattimore, Tr. 15338–40). [**276**]

1121. In the second stage, salesmen were asked to pick up samples of Sterling's major competitors—Anacin, Excedrin, Bufferin, and brands which are known as nationally distributed 5–grain aspirins, including Squibb, McKesson, Norwich, and St. Joseph. Nationally distributed brands are those which are probably available in all 48 states which Sterling serviced (Mattimore, Tr. 15343). Since the objective was to get a national sample, Mr. Mattimore asked a sales representative in each of the Glenbrook 14 sales districts to obtain samples (Mattimore, Tr. 15343–44).

1122. In the third phase, Mr. Mattimore instructed salesmen to pick up samples of Bayer Aspirin, also on a national basis. This request was made to the 14 sales districts, following the same procedure as in obtaining samples of the combinations and the nationally distributed 5–grain aspirin tablets (Mattimore, Tr. 15344).

1123. When samples were received by Mr. Mattimore at his New York office, he checked them against his requests to the salesmen. Follow-ups were made if no samples were submitted for a brand, or if insufficient number of samples were received. In almost every instance, the explanation was that the representative could not locate the branded product (Mattimore, Tr. 15346). This was not surprising since there are changes in store and private-label brands (Alberts, Tr. 8953–54; Mattimore, Tr. 15348–49).

1124. The samples were generally shipped in a corrugated container, packaged with paper or some type of resilient material, sealed and shipped by mail. The Glenbrook salesmen from whom Mattimore requested samples had previous experience in picking up samples, a procedure followed with competitive products. It was Mr. Mattimore's view that in connection with the sample pick-up for this study the sales representatives did carry out his instructions properly (Mattimore, Tr. 15342–43, 15347).

1125. Mr. Mattimore testified that CX 448, pages K and L, called "Identification and Location of Brand and Acquisition of Samples," accurately describes the procedures which were followed in the 1967 survey and 1968 collection of samples (Mattimore, Tr. 15348–49).

1126. Mr. Mattimore testified that when the brand survey was conducted in 1967, and the samples were collected in 1968, neither Mr. Mattimore nor the sales representatives knew the specific purpose for which this was being done (Mattimore, Tr. 15344, 15374).

1127. In determining the characteristics to be tested in the 223 Study, Dr. Marcelli relied upon her own experience in the [277] Pharmacy Division, and also considered an earlier and less complete pharmaceutical study from the early 1960's, the results of which were available in a draft report, CX 445, "The Quality of Aspirin Tablets," by Jerome Winig and Gail Prince. On this basis, Dr. Marcelli prepared a list of characteristics which were considered important to pharmaceutical quality. The tentative list was checked with the Pharmacy Research Division of the Sterling Winthrop Division for their suggestions and confirmations. It was also discussed with Mr. Winig and Mr. Mannix at the Glenbrook Laboratories Trenton plant. Dr. Tainter approved the basic list in March 1968 (Marcelli, Tr. 17407–08).

1128. The 223 Test, CX 448, reports tests and observations of the following physical and chemical characteristics:

1. Aspirin content - USP requirements
2. Aspirin content - Bayer standard
3. Tablet weight - USP
4. Absence of capping
5. Disintegration time - USP method
6. Disintegration time - Bayer method
7. Free salicylic acid - USP limits
8. Free salicylic acid - Bayer requirements
9. Absence of off-color
10. Absence of acetic odor
11. Freedom from wicking or wadding
12. Frequency and severity of tablet miscount
13. Rate of tablet breakage
14. Clarity of package size
15. Clarity of aspirin concentration
16. Legibility of label copy
17. Presence and adequacy of indications for use
18. Adequacy and accuracy of dosage instructions
19. Presence of required caution
20. Presence of required warnings
21. Use of package insert

22. Provision of sealed package
23. Provision of safe, undamaged container
24. Presence of security-closed caps
25. Control number presence and legibility - carton
26. Control number presence and legibility - container
27. Chipping
28. Miscellaneous contents imperfections
29. Wadding presence and adequacy
30. Deficiencies in container appearance

Most of the attributes were either required under regulatory or compendial requirements or were regarded as desirable by Sterling at that time. Many of the latter were later [278] incorporated into regulatory or compendial requirements (Marcelli, Tr. 17411–12; CX 448I-J).

1129. RX 181A-E, a letter from Marcelli to Winig dated April 3, 1968, included a handwritten draft report form, including examples of the types of entries to be made on the basis of physical observation of the samples in New York and of the testing to be conducted at the quality control laboratory of the Trenton plant. RX 181 was written after consultation with Winig and Mannix. It was understood at that time that the laboratory testing was to follow the standard testing procedures used in the quality control laboratories (Winig, Tr. 13743; Mannix, Tr. 14609; Marcelli, Tr. 17409–10, 17585–88). Subsequent correspondence and discussion between Dr. Marcelli and Mr. Mannix concerned the format for presentation of results and the use of pass-fail standards (Mannix, Tr. 14621–22; Marcelli, Tr. 17409, 17585–86; RX 181F-G, K).

1130. In the 1960's and early 1970's, at the time the 223 Test was conducted (CX 448), the emphasis was on tablet disintegration rate. Today the emphasis is on dissolution rate and absorption-bioavailability (Danhof, Tr. 17067). At that time, there was no standard test for dissolution or bioavailability of aspirin tablets (Winig, Tr. 13756).

1131. CX 448P-U describes the various tests and observations carried out at the Trenton laboratory (Winig, Tr. 13739; Mannix, Tr. 14609–10; Marcelli, Tr. 17464–67; CX 448P-U).

1132. This work was done under the supervision of Mr. Edward Mannix, Director of Quality Control, at the plant. Mr. Jerome Winig, the plant manager, asked that the quality control laboratories undertake the work, at the request of the Sterling Medical Director, Dr. Tainter. As Director of Quality Control, Mr. Mannix was autonomous of the plant administration, reporting to quality control officials in the company. In the 223 Test, he was in charge of the testing. He set up the program, and participated in establishing the report format.

He assigned persons directly under his supervision to undertake the testing (Mannix, Tr. 14605–07; Winig, Tr. 14255; Marcelli, Tr. 17420–22).

1133. The test procedures used in the study were routine standard testing procedures, with which the laboratory staff were familiar. The personnel, equipment and procedures regularly used for testing chemical and pharmaceutical characteristics in the quality control laboratories were used in connection with this survey (Winig, Tr. 13738, 13743, 14261–62; Mannix, Tr. 14624–25, 14609; Marcelli, Tr. 17464–67; CX 429C).

1134. The tests conducted at the Trenton Laboratories were tablet count, color, odor, general appearance and disintegration. The tablet disintegration test was done by two different methods—the USP method and the Bayer method (Mannix, Tr. 14609–10; CX 448P-U). **[279]**

1135. The USP disintegration procedure used what is commonly termed the Vandercamp apparatus, described in the USP, using discs which move and hit the tablets as the apparatus is raised and lowered in the water medium. The Bayer basket technique involves a simple screen and stirring device and is used in the normal course of Bayer quality control. The Trenton plant's quality control staff was competent in both methods (Winig, Tr. 13740–42; Mannix, Tr. 14612–13).

1136. Analytic testing procedures in effect in the Trenton plant at the time are described in CX 429D-G, "Quality Control Specifications for Bayer Aspirin Tablets." These were the standard testing procedures used by the quality control group, and were taken from the plant monograph then in effect. They cover all the items dealt with in CX 448, except for color, odor, tablet count, and the USP disintegration method referred to above (Winig, Tr. 13742; Mannix, Tr. 14613–14; Marcelli, Tr. 17464–67).

1137. Testing began at the Trenton Laboratories for the CX 448 study in July 1968 and continued over a period of two years to August 1970 (Winig, Tr. 13737–38; Mannix, Tr. 14619; RX 181H, N-O).

1138. Two methods were used to transport the samples from Dr. Marcelli's custody in the New York office to the Trenton plant. Under one method, the samples were transported by the company's regular courier service to Secaucus, New Jersey, where there was a distribution center. There was routine transport between Trenton and Secaucus. The second method was through the mails. The mails were used when necessary (Mannix, Tr. 14628; Marcelli, Tr. 17444–46).

1139. After Mr. Mannix received the samples that were sent from Dr. Marcelli to the Trenton plant, he made assignments to various technicians to do the various tests according to their expertise and availability. The technicians were told that the work was in connec-

tion with a survey, and that they were to conduct it the same as they would an everyday procedure. No one told the technicians that the results of the study were to be used for advertising purposes. Testing of competitive products had been done previously on a routine basis (Winig, Tr. 13745; Mannix, Tr. 14622–24).

1140. At the time the testing was done, neither Mr. Winig nor Mr. Mannix had any understanding or information that the results were intended to be used for advertising purposes. It was their understanding that this was another survey of competitive products, like others that had routinely been done in the laboratories. The first information that Mr. Winig had of possible use or advertising was in the summer or fall of **[280]** 1971, at meetings to consider a new advertising campaign. Mr. Mannix was later informed of such consideration, and the resulting decision (Winig, Tr. 13736, 13752; Mannix, Tr. 14623–24).

1141. In conducting these laboratory procedures, following the usual procedures, the samples were not blinded. The purpose of the survey was to study commercial aspirin tablets in the form in which they were available to consumers, and this placed sharp limitations on blinding. Respondent's witnesses testified that since the survey employed routine, standardized tests, it was unnecessary to blind the samples and that blinding would have altered the physical composition of the tablets which is part of the evaluation (Rhodes, Tr. 11434–36, 11440–44; Banker, Tr. 12906; Fields, Tr. 16600–01).

1142. Mr. Mannix and his two supervisors were responsible for taking the data which constituted the test results from the notebooks and putting it into the reporting format to be sent to Dr. Marcelli. RX 181N-O is an example of the reporting format or reporting sheets sent from the laboratories in Trenton to the New York office (Mannix, Tr. 14625).

1143. In the collection of samples from the field, 14 Bayer samples were obtained and delivered to Dr. Marcelli. Seven of the Bayer field samples did not undergo laboratory examinations at Trenton because they were lost or destroyed in the course of transportation (Mannix, 14628–31; Marcelli, Tr. 17455–56; RX 181J; CX 429H).

1144. Efforts were made to find replacement samples for the lost Bayer samples of approximately the same age (or plant control number). These could not be found in retail establishments. Five replacements were found in Sterling facilities—four in "Free Goods 90 Park N.Y." which was a reference to aspirin which constituted part of an overshipment to a consumer or returned by a customer for credit, and one at the Sterling Winthrop Research Institute, Rensselaer. A total of 19 Bayer samples underwent physical observations in New York (Mannix, Tr. 14630; Marcelli, Tr. 17449–51).

1145. Laboratory tests were performed on the Bayer samples in August 1970. The samples tested at the laboratory were seven field samples and the five replacement samples (Mannix, Tr. 14631; Marcelli, Tr. 17454–55).

1146. In late January 1971, Dr. Marcelli undertook the analysis of the data resulting from the tests and performed the tabulation, analysis and writing of the report (CX 448). This work was done in January-March 1971. Dr. Marcelli had the advice and assistance of the Biometrics Section at the Sterling-Winthrop Research Institute, which is the expert biostatistical body at Sterling. The two handwritten tables in CX 430 were [281] prepared by the Biometrics Division. Dr. Marcelli had staff assistance in tabulating the data (Marcelli, Tr. 17478–81).

1147. Dr. Marcelli adopted a statistical cutoff for aspirin brands that were to be individually examined in the report, as the Biometrics Section indicated that comparisons should be confined to those brands where there was a minimum of six samples and at least four control lots in order to have a reasonable estimation of the distribution within the brand. All of the remaining brands were placed together in a single group labeled "Miscellaneous." At the time that this recommendation was made by the Biometrics Section, that section did not have access to any of the test results (Horner, Tr. 10762–63; Marcelli, Tr. 17482–85).

1148. It is Sterling's position that the overall conclusion arising from CX 448 is that Bayer Aspirin tablets were superior to all other plain 5–grain aspirin tablets represented in the study in terms of overall pharmaceutical quality. This was based upon analysis and evaluation of the data with respect to the 30 characteristics or categories reported on in the study. The 30 categories were divided into 26 primary categories and the 4 secondary categories. Bayer had 8 failures in 5 of the 30 performance categories, but no other brand with a reasonable representation equalled this record. The brand closest to Bayer in performance had twice as many failures. Only 3 of the Bayer failures fell into the 26 primary categories whereas with the other brands, 3 to 5 times as many failures fell into the primary categories (Marcelli, Tr. 17488; CX 448D-H; RPF 7.528).

1149. Dr. Marcelli reported that based on this 30 criteria employed in CX 448, Bayer was superior to 220 aspirin brands (CX 448D). Specifically, Bayer routinely yielded 324 mg aspirin per tablet with more lot-to-lot consistency than the other brands. One hundred fifty-seven competitors yielded at least one failure (CX 448D). For disintegration, Dr. Marcelli reported that Bayer consistently met a standard of beginning disintegration with 2 seconds and completing disintegration within 30 seconds, while 70 others failed to do so (CX 448D). For

FSA level, only one Bayer sample yielded FSA in excess of .035%, while 90% of competitive samples yielded such FSA values (CX 448E). Bayer showed a uniformly, pure white color while 9 major brands and approximately 20% of minor brands showed at least one off color sample (CX 448E).

1150. Dr. Marcelli also reported that Bayer showed perfect tablet count, but only four other major brands did so. One hundred seventy-eight minor brands showed tablet counts varying from the label claim (CX 448E, F). Bayer and 6 major brands manifested rare instances of broken tablets, while 50% of the minor brands registered broken tablets (CX 448F). Bayer and eight major brands manifested uniformly good label legibility. Most minor brands showed poor label legibility (CX 448F). On [282] packages of Bayer and five major brands, indications were clear. All minor brands registered deficiencies in the presentation of indications (CX 448F). She reported that only Bayer reliably included dosage recommendations, cautions and warnings on every sample (CX 448G). Only Bayer and St. Joseph provided package inserts (CX 448G). Among the major brands, only Bayer registered sealed units for every sample. Only a "handful" of minor brands provided such protection (CX 448G). Bayer routinely showed undamaged and safe containers, while other brands did not (CX 448G). Of the major brands, only Bayer manifested uniformly legible control numbers on cartons and bottles. Minor brands showed numerous failures, including omissions of these numbers (CX 448G). Bayer was free of extraneous dust and stray fragments, while 75% of the other brands manifested some deficiency detracting from the general appearance of the product of package (CX 448H).

1151. Dr. Marcelli concluded that Bayer alone showed failures in only 5 of the 30 categories, that those 26 parameters relating to efficacy, tolerance, stability, and safety, Bayer showed minor failures in 3 categories, and that no other brand with "reasonable representation" matched Bayer's record, and that competitive brands' failure rates ran three to five times Bayer's rate (CX 448H).

1152. Dr. Marcelli completed the report in March 1971, and distributed it to those on a list provided by Dr. Tainter. Her involvement ended with the submission of the report. Dr. Marcelli did not participate in any meetings later in 1971 that considered the report in the context of a proposed advertising campaign. She testified that she first heard of the possible use of the study for advertising purposes in late October or early November 1971, after she had left Sterling (Marcelli, Tr. 17532–33).

1153. In 1971, the 223 Test was considered by a group which included such company scientific experts as Dr. Blackmore, Director of Clinical Research at Sterling's Research Laboratory; Dr. Rosenberg,

Head of the Pharmacology Section of the Research Institute; Dr. Swarbrick, Head of the New Product Development Group and a former professor of Pharmaceutical Science (now Dean of the University of Southern California School of Pharmacy); Mr. Winig, Plant Manager at the Bayer Trenton plant and an expert in pharmaceutical and manufacturing standards in aspirin; Dr. Trout, Sterling Medical Director; and other medical personnel. A consensus was reached at those meetings that the 223 Test was valid and reliable and that it provided a basis for making a claim of superiority in pharmaceutical quality for Bayer Aspirin (Alberts, Tr. 9002–03; Winig, Tr. 13752; Trout, Tr. 16094).

1154. At the trial, respondent's expert witnesses, without exception, testified that because pharmaceutical quality is [283] related to therapeutic efficacy, physicochemical pharmaceutical tests, such as CX 448, provide a reasonable basis for conclusions regarding comparative therapeutic performance of aspirin products.

1155. Dr. Horner, respondent's expert biostatistician, testified that in assessing the clinical significance of differences in various parameters between aspirin brands, it is necessary to consider the net effect of all differences, rather than to isolate a single parameter. In this opinion, a critical issue is the assessment of overall pharmaceutical and therapeutic superiority, as opposed to making a series of independent decisions based on individual parameters (Horner, Tr. 10835).

1156. Dr. Feinstein testified that he reviewed the material in the "223 Test" (Feinstein, Tr. 16374). He described the study as containing the kind of evidence that he would resort to in making a decision as to which of the products would be better therapeutically (Feinstein, Tr. 16374, 16379).

1157. Dr. Rhodes, an expert in pharmaceutics, testified that, in his opinion, CX 448 is a valid pharmaceutical study, which demonstrates that when Bayer Aspirin was compared with a large number of other aspirin products available on the United States market, Bayer Aspirin was of better quality than those produced by its competitors. In his view, the 223 Test provides a reasonable scientific basis for the conclusion that Bayer Aspirin and Bayer Children's Aspirin have been tested against other brands of aspirin and found to be qualitatively superior to other brands. It would also be reasonable to draw a therapeutic conclusion based upon the therapeutic importance of the parameters measured in this study. In his opinion, a reasonable drug company in the late 1960's and early 1970's would have acted reasonably in selecting the parameters measured in the study in attempting to determine the pharmaceutical and therapeutic superiority of its aspirin over competitive brands of aspirin (Rhodes, Tr. 11425–26, 11434–43).

1158. It was the judgment of Dr. Banker, an expert in pharmaceutics, that the 223 Test was a comprehensive test which evaluated meaningful parameters of different brands of aspirin tablets. In Dr. Banker's view, the methodology is valid, and clearly established that Bayer was the most nearly optimized brand of aspirin tablets, and was pharmaceutically and therapeutically superior to the other brands evaluated. In his opinion, a reasonable drug company would have a right to rely on a study such as the 223 Test in making superiority claims for its aspirin and a reasonable hospital pharmacist or clinician would be justified in relying on such a study in selecting Bayer Aspirin over other brands of aspirin tablets for treatment of patients (Banker, Tr. 12779–81; Danhof, Tr. 16946–67). **[284]**

1159. Dr. Fields testified that the 223 Test was of value in selecting an aspirin brand to be used in the NIH Stroke Study because of the relationship between such characteristics and the therapeutic performance or side effects of an aspirin tablet. In the absence of controlled clinical studies, the 223 Test was considered to have a bearing upon determining which aspirin brand would produce the least variability, the most likelihood of bioavailability, and the least side effects. Among the physical and chemical characteristics considered were disintegration, amount of impurities, including free salicylic acid, and the stability of the tablet. According to Dr. Fields, the expert pharmacologists relied upon the 223 Test and other information in selecting Bayer Aspirin for use in the NIH Stroke Study (Fields, Tr. 16585–86, 16598–600, 16566, 16744–45).

1160. Dr. Scoville, a former FDA official, also testified that the 30 physical and chemical characteristics studied in the 223 Test were included in the type of material that is reviewed by the FDA, together with appropriate clinical data, in reaching judgments as to the safety and efficacy of a drug product and in determining whether to approve a drug for marketing or to seize or recall a drug product from the marketplace (Scoville, Tr. 14448–49).

1161. Dr. Falliers, an expert in allergy, testified that to the extent that the 223 Test demonstrated that Bayer Aspirin is pharmaceutically superior in the characteristics tested, it provided a reasonable basis to conclude that Bayer Aspirin is therapeutically superior (Falliers, Tr. 13320–21, 13326).

1162. The record shows that Sterling knew in August 1971, about five months after the completion of the 223 aspirin survey, that there were some 328 plain aspirin brands in the United States. Sterling knew in November 1971 that the number of competing aspirin brands was possibly as high as 442 (CX 363A; Alberts, Tr. 9045–47).

1163. The record shows that the reliability of the 223 Test and the validity of its findings are subject to serious doubts because of perva-

sive methodological deficiencies throughout the entire survey. Despite the elaborate and considerable research trappings which adorn the 223 Test, it is fair to conclude that its overall quality falls short of that generally required to substantiate unqualified claims of pharmaceutical or therapeutic superiority with respect to plain 5-grain aspirin brands. However, this does not detract anything from CX 448's utility as an ongoing internal quality monitoring tool, as its more modest predecessors had been (F. 1124, 1127, 1139, *supra*).

1164. It is well recognized that for a properly designed and well-controlled scientific study, a protocol must define in sufficient detail all of the important aspects of the study, including any plan for statistical evaluation (Moertel, Tr. [285] 6275, 6287; DeKornfeld, Tr. 8393, 8400; Horner, Tr. 10818–19, 10890–91, 10897). CX 448 does not contain any protocol (Rhodes, Tr. 11803). The draft report form contained in Dr. Marcelli's April 3, 1968 letter to Mr. Winig (RX 181A-E), although informative, cannot be characterized as a "protocol," and it does not include any description of contemplated statistical evaluation. In fact, no "protocol," in the conventional sense, was established for CX 448 (Marcelli, Tr. 17585–87). Mr. Mannix, who was responsible for overseeing the Trenton testing phase, testified that he had begun reporting test results to Dr. Marcelli before she had committed to paper what it was he was to be testing (Mannix, Tr. 14663).

1165. Test samples were collected by Sterling's field salesmen in 1968 (Mattimore, Tr. 5338–41, 15343). The record indicates that the sales representatives who collected the test samples received no written instructions concerning the manner of collection (Mattimore, Tr. 15370–79). The sales representatives simply received requests to pick up certain brands of aspirin (Mattimore, Tr. 15339). The witness called by respondent to provide evidence on the method and reliability of the collection state of this study did not remember whether or not written instructions ever existed concerning the collection of samples for major brands (Mattimore, Tr. 15374–79). No means exist to determine whether that portion of respondent's sales force involved in the collection effort had had any earlier training or experience in selecting samples (Mattimore, Tr. 15374–79) or how the sales representatives chose the retail outlets for the collected samples (Mattimore, Tr. 15378). The record does not show that any attempt was ever made to randomize any phase of the selection process.

1166. The sales representatives were not asked to report on the condition of the samples (Mattimore, Tr. 15372). They received no instructions regarding the manner of shipping the samples (Mattimore, Tr. 15372–74). Thus, no means exist to determine how much disparity there was among the samples in terms of physical appearance and storage conditions at the retail level. No way exists to deter-

mine how much disparity, if any, there was among the samples in terms of their handling after purchase and before examination in respondent's New York City office. Thus, from the very beginning, no controls were incorporated in the test for disparate condition and treatment of the samples.

1167. Mr. Mannix, who was responsible for quality control at the Trenton plant and oversaw the Trenton testing phase of the 223 Test, testified that differences in handling must be avoided to prevent biasing the results of tests, such as powdering, chipping and aspirin content, all of which were tested in the 223 Test (Mannix, Tr. 14660–61; CX 448I). Nevertheless, Sterling exercised no control over the manner of handling samples. For example, some samples were transported by special **[286]** company courier, others were transported by mail (Mannix, Tr. 14628; CX 678, admission 982). Of those transported by mail, possibly different classes were used (Winig, Tr. 14257). Some samples were handled roughly in transit and others not (Mannix, Tr. 14671–72). It made no sense to conduct tests on samples that were not collected and handled in the same way (Mannix, Tr. 14664).

1168. The 223 report itself indicates that the study failed to control for age or circumstances of handling of the collected samples (CX 448V; Rhodes, Tr. 11665–66). Without these controls, the study cannot be expected to show whether the test results (*e.g.,* tablet breakage and FSA levels) reflected differences in manufacturing practices, or age, or retail storage practices or other circumstances, of the samples (*see e.g.,* Rhodes, Tr. 1168; Banker, Tr. 13008). The former Vice President of the Glenbrook Laboratories Division and a control chemist of 42 years' experience at Sterling agreed that variables such as age, as well as conditions of storage of an aspirin product "without doubt" must be controlled as a matter both of scientific interest and with respect to the conclusions one can draw from a pharmaceutical study such as the "223 study" (Winig, Tr. 14215–56). Failure to control for such factors also can lead to biased results with respect to disintegration time and FSA levels (Winig, Tr. 14213–14). Failure to control for differences in handling can affect physical properties such as powdering, chipping or tablet weight.

1169. Dr. Marcelli selected the 30 parameters after reviewing the results of CX 445, an earlier in-house comparative study of Bayer and 152 other plain 5-grain aspirin brands (CX 678, admission 1011; CX 445A). Thus, Dr. Marcelli knew on which parameters Bayer had fared poorly or well, in comparison with other brands. She decided which parameters belonged in the primary and secondary categories (CX 448I, J; Marcelli, Tr. 17592). Dr. Marcelli placed in the primary category 10 factors which J. Winig—co-author of CX 445 (CX 445B)—characterized as matters of "pharmaceutical elegance" (CX 445H-L),

including: label legibility, adequacy of indications, adequacy of dosage instructions, adequacy of caution, adequacy of warnings, provision of sealed package, presence of securely closed caps, and presence of control number on carton and container, tablet breakage (CX 445H, I; CX 448I, J). Thus, 10 of the 16 parameters applied by Dr. Marcelli, and 10 of the 30 parameters applied in this test, comprised "pharmaceutical elegance," and not pharmaceutical quality, to a Sterling employee who for over 40 years was responsible for the manufacturing or quality control of Bayer Aspirin.

1170. The first stage of recording observations on the collected samples was conducted by Dr. Marcelli herself in respondent's New York City offices (Marcelli, Tr. 17433–35). She assessed the test samples on a "first come-first served" [287] basis (Marcelli, Tr. 17596–97). Of the 30 parameters employed in the CX 448 (CX 448I, J), Dr. Marcelli administered 16: freedom from wicking or wadding (11); clarity of package size (14); clarity of aspirin concentration (15); legibility of label copy (16); presence and adequacy of indications for use (17); adequacy and accuracy of dosage instructions (18); presence of required caution (19); presence of required warnings (20); use of package insert (21); provision of sealed package (22); provision of safe, undamaged container (23); presence of securely closed caps (24); control number presence and legibility-carton (25); control number presence and legibility-container (26); wadding presence and adequacy (29); deficiencies in container appearance (30) (CX 448I, J). All 16 criteria involved assessments by visual examination, not laboratory tests (Rhodes, Tr. 11800–01; Winig, Tr. 14255; Marcelli, Tr. 17597–99). Thus, over half of the test data resulted from sensory tests conducted by an unblinded employee of respondent, in a nonrandomized manner. None of her observations were checked or replicated by anyone else (Marcelli, Tr. 17604). Thus, this study failed to control human error for over half the test.

1171. CX 448's remaining 14 parameters were tested by members of Trenton plant's quality control staff (Mannix, Tr. 14610; 14644–48; 14651–52; 14655–57; 14659). Of the 14, 7 involved 5 separate laboratory analyses(*i.e.,* aspirin content, disintegration by the official test method, disintegration by the Bayer test method, FSA, and tablet weight) (Rhodes, Tr. 11800; Mannix, Tr. 14655–57; CX 429E-G). The remaining seven(*i.e.,* absence of capping, off color, acetic odor, tablet miscounts, tablet breakage, chipping, and miscellaneous contents imperfections) involved sensory tests (Rhodes, Tr. 11800–01; Mannix, Tr. 14646–68; CX 429D). Thus, for 23 of the 30 criteria, or over two-thirds of the test, the data were obtained from sensory tests. Furthermore, these tests were conducted by unblinded employees of respondent (Wining, Tr. 14260; CX 678, admission 988), even though blinding was

possible for at least some of the laboratory analyses (Winig, Tr. 14260–61). Thus, all personnel conducting the examinations and analyzing the test results knew the identity of the samples. Additionally, these tests were not performed in a randomized order (Winig, Tr. 13738; Mannix, Tr. 14622). Therefore, this stage of the study failed to control for human error for the remaining half of the test.

1172. Reports of the test results generated by the Trenton staff were submitted by Mr. Mannix to Dr. Marcelli on an ongoing basis (Mannix, Tr. 14662; Marcelli, Tr. 17606). The first such report, dated October 15, 1968, shows that the Trenton staff had begun conducting and reporting test results on competitive 5-grain aspirin samples before final agreement was reached on the totality of test parameters (RX 181A, F, G; K; Mannix, Tr. 14663; Marcelli, Tr. 17585–87). This, in addition to the lack of a formal protocol, increases the likelihood that the Trenton [288] tests were not conducted uniformly or in a standardized manner on all the tested samples. Without such standardization, the study failed to control for unequal treatment and inadvertent error by the testers.

1173. Mr. Mannix submitted 17 reports of Trenton test results to Dr. Marcelli (Marcelli, Tr. 17613). Dr. Marcelli had received 14 of these reports before she sent the 12 Bayer samples for testing at Trenton (Marcelli, Tr. 17605–13; RX 181J). After receiving some of Trenton's test results on competitive 5–grain aspirin samples, Dr. Marcelli stated: "[A]t least there is some encouragement in these early results to suggest that it will be possible to show differences between brands—especially between ours and others." (Mannix, Tr. 14688; RX 181K). Thus, the employee who later analyzed the data and prepared the report actually knew some test results during the course of the test. Also, she expressed an expectation that the test would show differences between Bayer and other aspirin brands. Such advance notice of partial test results and such an early expectation regarding the desired outcome of the study introduce a distinct likelihood of bias influencing the test results (Moertel, Tr. 6346; Banker, Tr. 12918–19).

1174. The 12 Bayer samples analyzed at Trenton (Mannix, Tr. 14628–30; Marcelli, Tr. 17605) were specially packed by Dr. Marcelli and sent by company courier, unlike the competing aspirin samples (Mannix, Tr. 14672–73). All or most competitive samples apparently were mailed (Mannix, Tr. 14673; Winig, Tr. 14257). Some samples sent by mail were handled roughly (Mannix, Tr. 14671). Thus, the test samples were not treated in a standardized manner and were subjected to different handling, with Bayer samples handled most carefully. Such differences in handling strengthens the possibility that the test results were, with respect to physical tablet characteristics such as

powdering, chipping and aspirin content, as likely due to chance, or a systematic bias, as to actual differences in the tested samples.

1175. Five of the 12 Bayer samples analyzed at Trenton (Mannix, Tr. 14628–30; Marcelli, Tr. 17605) came from "free goods," *i.e.*, samples returned to respondent, and thus they did not represent Bayer samples as found on the store shelves (Mannix, Tr. 14668–70; Marcelli, Tr. 17450, 17628, 17634–36). These five were not retained samples which respondent's witness Dr. Rhodes, stated would have been appropriate replacements (Rhodes, Tr. 11438–40). Thus, unlike the other brands, about 50% of the Bayer samples, which were analyzed on all parameters applied by the Trenton laboratory tests and subsequently compared to other brands, were not commercially available samples in the usual sense (Rhodes, Tr. 11803–04; Marcelli, Tr. 17450). Furthermore, despite Sterling's recognition that a range of lots and samples are necessary to draw conclusions about manufacturers (Marcelli, Tr. 17443), the range of lots **[289]** and samples represented in the 223 study were unequal. Of the major brands, Bayer was represented by 12 samples, 12 lots; Lilly by 6 samples, 4 lots; McKesson by 35 samples, 21 lots; Norwich by 36 samples, 29 lots; Rexall by 12 samples, 3 lots; St. Joseph by 43 samples, 35 lots; Squibb by 40 samples, 30 lots; Upjohn by 4 samples, 3 lots; and Walgreen's by 12 samples, 6 lots (CX 448XZ013–25; Marcelli, Tr. 17631–33). This may explain the fact that the eight other brands generated information about intra-lot variability while Bayer did not. Thus, the study provides relatively more information about at least four other major brands, *i.e.*, McKesson, Norwich, St. Joseph, and Squibb, than about Bayer.

1176. The Trenton plant staff measured FSA and disintegration, each against two limits (CX 448I). For each of the two factors, one was Bayer's internal standard employed in the production of Bayer (CX 448I). These two in-house specifications have remained constant since at least the early 60's (Winig, Tr. 14227, 14232; Mannix, Tr. 14603; 14675). The Trenton staff routinely applied these two limits in its ongoing quality control work (Mannix, Tr. 14609–10, 14675). However, in 14 of the 17 Trenton reports, the staff nevertheless employed two limits which had never been used as internal limits (Mannix, Tr. 14678; Marcelli, Tr. 17605–13). Such an error about respondent's own standards calls the reliability of this test into question.

1177. Despite the importance of setting forth the plan for statistical analysis in advance (DeKornfeld, Tr. 8400), the plans for statistical evaluation of the 223 Test data were not described in advance (Marcelli, Tr. 17637–79). Decisions to exclude certain brands from statistical analyses were made after the test results became available (Marcelli, Tr. 17482, 17637–79). One such rule was to exclude, from brand-to-

brand comparisons, several minor brands which appeared to Dr. Marcelli simply to be different bottles of the same brand, *i.e.,* they showed similar looking bottles, caps, and labels (Marcelli, Tr. 17440, 17518–25). The rationale for this decision was that the number of physical packaging similarities indicated that the *manufacturer* of the aspirin tablets was the same, and that these different labels represented only one brand (CX 48Z041; Marcelli, Tr. 17642–43). However, another equally plausible explanation, based on common industry practice, is that one repackager packaged aspirin tablets from different suppliers into similar bottles with similar caps (*see,e.g.,* Marcelli, Tr. 17644–65; Rhodes, Tr. 11371). The investigators did not, and could not, determine which explanation was correct (Marcelli, Tr. 17645), yet chose to exclude the performance ratings of several aspirin brands (CX 448Z041–Z045). The effect of grouping these brands into one "score" was to nullify the individual, good ratings of the different brands along individual parameters (Marcelli, Tr. 17719). **[290]**

1178. Another rule adopted after the tests was to exclude, from a brand-to-brand comparison, several minor brands which were represented by fewer than six samples from four lots (CX 448Z077; Marcelli, Tr. 17483, 17637–41). The rationale was that such representation was insufficient for permitting a reasonable estimate of brand performance (CX 448Z027; Marcelli, Tr. 17443, 17483). Dr. Horner, respondent's expert witness, testified that this decision was justified under the circumstances and reflected accepted statistical procedures (Horner, Tr. 10807–09, 10894–97). However, this rule (statistical cut-off point) was not applied to major brands. Both Rexall and Upjohn were represented by three lots, and yet subjected to statistical evaluation, including tests for statistical significance in comparison with Bayer (CX 448Z001–Z002, Z004; Marcelli, Tr. 17642).

1179. Apart from the considerable and extensive methodological problems discussed above, the 223 Test results do not show Bayer is significantly superior to all other tested brands in the tested respects. Six brands (McKesson, Norwich, Rexall, St. Joseph, Upjohn, and Squibb) began and completed disintegration at rates which were statistically insignificantly different from that for Bayer (F. 633, *supra*). Since the investigators failed to test for statistical significance of the brands' aspirin content averages, this survey does not show that Bayer was statistically significantly superior to all other brands in terms of aspirin content (F. 677, *supra*). It also does not show that Bayer more consistently yielded 100% of label claim than all other tested ASA brands (F. 677, *supra*). Since the investigators similarly failed to test FSA level data for statistical significance, this survey does not show that Bayer yielded a statistically significantly lower

FSA level than all other tested brands (F. 733, *supra*). The 223 Test did not include dissolution or blood level tests.

1180. Thus, several significant methodological deficiencies occurred throughout the stages of the "223 Test." These deficiencies cannot be brushed aside. They raise a distinct possibility that the test results were as likely due to chance, or a systematic bias, as to actual differences in the tested brands. This possibility leaves the validity and utility of the test data in question. Even if these problems were disregarded, the test data discussed in F. 1179, *supra,* alone show that this test did not demonstrate that Bayer was pharmaceutically or qualitatively superior to all the other 220 tested brands in the designated respects.

1181. Since the "223 Test" is not a clinical test, conforming to accepted scientific guidelines for a well-controlled clinical trial, it cannot offer any reliable conclusion regarding the therapeutic performance of the tested brands. It purported to assess various brands of plain 5-grain aspirin only in terms of pharmaceutical characteristics. **[291]** Although respondent's expert witnesses testified that the 223 Test (CX 448) is sufficient basis for reaching a clinical judgment that Bayer is superior overall to other brands, the record as a whole is clear that none of these pharmaceutical characteristics has been shown to relate, directly and reliably, to comparative therapeutic performance of aspirin.

1182. In addition, the medical director for Glenbrook Laboratories from 1971–1974 believed it very unlikely that therapeutic differences could be shown between two brands, both passing parameters (aspirin content-USP), 3 (tablet content-USP), and 7 (FSA limit-USP) (CX 448I) (John, Tr. 5564). Furthermore, the record as a whole shows that the scientific community has not recognized tests of such physicochemical properties, as were tested in CX 448, as providing anything more than a hypothesis concerning possible clinical effects (Grossman, Tr. 7499–7500). The hypothesis must be subjected to clinical testing, and the hypothesis alone provides no evidence of clinical or therapeutic superiority (Grossman, Tr. 7499–7500; DeKornfeld, Tr. 8414–17; Moertel, Tr. 6308–09).

1183. Dr. Marcelli stated that a brand with superior ratings for eight factors (for FSA, color, odor, wicking or wadding, aspirin content, capping, breakage, and disintegration), would likely be therapeutically superior to other brands (Marcelli, Tr. 17670–82). Applying this standard to the reported results of the "223 Test," Parke-Davis, Acme, and Tripple AAA would be likely to be therapeutically superior to Bayer (Marcelli, Tr. 17670–82; CX 430A-B).

1184. Therefore, the 223 Aspirin Test is not sufficient to substantiate the representation, as alleged in Complaint Paragraph Twenty,

that Bayer Aspirin is qualitatively superior in the designated respects, including speed of disintegration, to all other aspirins tested. Moreover, the 223 Test is not sufficient to substantiate the representation, as alleged in Complaint Paragraph Twenty, that Bayer is therapeutically superior to all other brands because at the time of this representation a substantial question existed, as recognized by experts qualified by scientific training and experience to evaluate the safety and efficacy of OTC analgesic agents, concerning the validity, significance and application of such tests to the question of therapeutic superiority.

1185. Respondent has also offered, as Bayer advertising substantiation, certain reports of recalls of brands of plain 5–grain aspirin published by the FDA (RX 152 for identification; *see, e.g.,* Banker, Tr. 12581–82). Such recalls have been undertaken voluntarily by manufacturers (Miller, Tr. 6941–42). Many of the recalls occurred because of violations of the GMP (Miller, Tr. 6945–48, 6950, 6956). These violations typically involved deficiencies in a manufacturing or distribution facility, rather than with the recalled products (Miller, Tr. [**292**] 6941–42, 7104). The record indicates that these recalls neither constituted a widespread problem (*see, e.g.,* Banker, Tr. 19582), nor involved major aspirin manufacturers, *e.g.,* St. Joseph, Squibb, Rexall, McKesson, and Norwich (Miller, Tr. 7105; Banker, Tr. 12930). Therefore, the history of recalls reported in this record does not provide a reliable basis for predicting the comparative therapeutic performance of plain 5–grain aspirin brands.

1186. Respondent has also offered, as Bayer advertising substantiation, certain reports of complaints by commercial institutional customers about various pharmaceutical attributes of some competitors' aspirin. The first, RX 215, consists of complaints to Monsanto, a major manufacturer and supplier of aspirin powder (Rhodes, Tr. 11336–43). The second, RX 217, consists of complaints to Norwich-Eaton, a manufacturer of aspirin tablets (Rhodes, Tr. 11347–70). Respondent acquired this material after 1974 (Tr. 8821–32). This record does not indicate that these complaints resulted in therapeutic problems to consumers. This record does not cite complaints to other major aspirin manufacturers or distributors, *e.g.,* St. Joseph, McKesson, Squibb, Dow (*see, e.g.,* RX 215). No witness pointed to information concerning the duration of the complaints (*see, e.g.,* Banker, Tr. 12987). In addition, this material contains notes and several pages of handwriting (Banker, Tr. 12987, 12992). Dr. Banker, respondent's witness, stated he was unable to vouch for the accuracy or precision of the information presented in RX 217 (Banker, Tr. 12992). Therefore, these reports do not provide a reliable basis for predicting the comparative therapeutic performance of plain 5–grain aspirin brands.

1187. Respondent has also offered, as Bayer advertising substantiation, materials concerning manufacturing specifications used by various aspirin manufacturers and tabletters, including respondent. They include RX 205, Dow's specifications for aspirin powder (Rhodes, Tr. 11327); RX 214, Monsanto's standards for supplying aspirin powder to various tabletters, *e.g.,* Bristol-Myers, Upjohn (Rhodes, Tr. 11332); and RX 169–170, respondent's specifications (Rhodes, Tr. 11305–08). These specifications typically present internal standards for pharmaceutical attributes, *e.g.,* FSA level, aspirin content, and for various stages in the formulation and tabletting of aspirin (*see, e.g.,* RX 214A, RX 169, respectively). Although aspirin manufacturing process and quality control clearly related to pharmaceutical quality of the finished product, the record does not show a direct and reliable correlation between these manufacturing standards and aspirin's clinical performance.

1188. No witness pointed to information concerning the effective period of the specifications or their completeness. This material does not include specifications for St. Joseph, Rexall, and McKesson (Banker, Tr. 12889, 13167). Moreover, no [**293**] witness for respondent had reviewed specifications for all of the approximately 100 aspirin manufacturers (*see, e.g.,* Banker, Tr. 12615, 12887–90; Rhodes, Tr. 11716–18). Therefore, these manufacturing specifications fail to provide a reliable basis for predicting comparative therapeutic performance of plain 5-grain aspirin brands.

1189. The combining of results from multiple tests into one composite "score" or "pooling" is scientifically appropriate only under certain conditions: (1) when plans to pool tests' results have been set forth in advance in the respective tests' protocols (Rickels, Tr. 8062–64); (2) when each test has been conducted according to the same protocol (Rickels, Tr. 8062–64); and (3) when each test has been sufficiently controlled to yield reliable results (*see, e.g.,* Banker, Tr. 12904).

1190. The various tests reviewed and discussed in this record do not meet these criteria. Very few of these tests included protocols. Those with protocols did not contemplate pooling, and those with protocols did not follow the same protocol.

1191. The record shows that none of the pharmaceutical studies respondent relies on were sufficiently controlled to yield reliable results. Many involved inadequate sampling or inadequate information to determine the adequacy of sampling. Respondent's expert witnesses have stated that representation by less than 10 samples is inadequate (Rhodes, Tr. 11478), representation by one lot is inadequate (Banker, Tr. 13145), and that variables which can affect the property under investigation, *e.g.,* age as it affects FSA level and aspirin content, must be ruled out or controlled (Banker, Tr. 12904).

1192. More importantly, before differences in any test data can be attributed to real differences in test samples (rather than chance), appropriate statistical evaluation must show that the differences were statistically significant. Most of the tests failed to include statistical evaluation and those which did generally lacked sufficient controls. Therefore, the tests discussed in this record do not meet the criteria for pooling and no composite score based on pooling can provide any scientific basis for a comparative pharameutical or therapeutic conclusion regarding different brands of plain 5–grain aspirin tablets.

1193. Hundreds of brands of adult plain 5–grain aspirin were commercially available during most, if not all, the time period of 1969–1974. In a September 22, 1971 memorandum, J. C. Marshall reported the identification of 422 different brands as of November 1971 (CX 363A).

1194. Therefore, at the time of the representation alleged in Complaint Paragraph Ten A, Sterling did not have a [**294**] reasonable basis for making the representation that Bayer is superior in terms of significant therapeutic effect to any other aspirin, because respondent lacked competent and reliable scientific evidence sufficient to support this representation.

VII. ARTICLES INCLUDED IN AND TESTIMONY BASED ON RX 250 FOR IDENTIFICATION DO NOT PROVIDE A REASONABLE BASIS FOR CHALLENGED ADVERTISING CLAIMS

A. *Legal Standard*

1195. During the trial, as part of its attempt to show a reasonable basis for certain challenged advertising claims, Sterling cited a large number of scientific articles and other textual material, collectively marked RX 250 for identification, as substantiation of these claims, and relied on testimony from its witness with respect to these articles. In order for these articles and the testimony elicited about them to constitute substantiation, there must be some showing that Sterling both possessed and relied upon these articles at the time it disseminated the challenged advertising claims. *Pfizer, Inc.*, 81 F.T.C. 23, 64–67 (1972).

B. *Possession*

1196. Dr. George Goldstein, medical director of Sterling from 1975 to the present, was respondent's sole witness with respect to possession of the RX 250 for identification materials. Because he was not employed by Sterling prior to 1975 (RX 274A-B), he could not testify regarding whether Sterling possessed any of these materials when

these claims were being made by drawing inferences from certain other facts either from his personal knowledge or by drawing appropriate inferences from other facts known to him during that period.

1197. Dr. Goldstein attempted to show possession of 19 of the RX 250 articles listed in Table I (CPF 452–453) solely by inference from the fact that they were turned over to the FTC pursuant to a 1971 subpoena. This may show that Sterling physically possessed these articles at the time they were produced to the FTC in 1971, but there is no evidence in the record to indicate that they possessed them *prior to* that time. Therefore, respondent has failed to carry the burden of proof of showing that these articles were possessed by it at the time it made the challenged advertising claims disseminated prior to 1971.

1198. Dr. Goldstein attempted to show Sterling's possession of 71 of the RX 250 articles listed in Table II (CPF 454–458), by inference derived solely from the fact that for each article another article from a different edition of the same journal in which it appeared had been turned over to the FTC pursuant to 1973, 1971 or 1966 subpoenas. From this Sterling would infer **[295]** that it was a regular subscriber to these journals and therefore possessed these articles at the time of the dissemination of the challenged advertisements. There is no evidence in the record, however, that Sterling subscribed to or regularly received these journals or that any corporate official or others responsible for preparation, review or approval of the advertising claims in question read them or relied on what he read with respect to any advertising claim. Indeed, Sterling has maintained that it has not made any of the advertising claims alleged in the Complaint. In these circumstances, Sterling has failed to carry the burden of proof of showing that these articles were possessed by and relied on it at the time it made the challenged advertising claims.

1199. Dr. Goldstein attempted to show possession of nine RX 250 articles listed in Table III (CPF 459) by inference derived solely from the fact that these articles had been turned over to the FTC pursuant to a subpoena or demand not otherwise identified in the record. There is no evidence in the record to indicate the year of the subpoena or subpoenas in response to which any of the nine articles was produced. Accordingly, there is no evidence in the record to indicate *when* Sterling possessed these articles. Thus, respondent has failed to carry the burden of proof of showing that it possessed these articles at the time it made the challenged claims to which these articles may relate.

C. *Reliance*

1200. Dr. Goldstein attempted to show Sterling's reliance on 19 of the RX 250 articles listed in Table IV (CPF at 460–461), solely by inference from the fact that these articles had been turned over to the

FTC pursuant to a 1971 subpoena. Dr. Goldstein accordingly left to inference that the articles were supplied for the same purpose in 1971 as at trial, an inference for which there is no evidence in the record. There is no evidence in the record to indicate for what purpose they were produced to the FTC in 1971. Sterling has failed to carry the burden of showing that these articles were relied on by it for substantiation of the challenged claims at the time of their dissemination.

1201. Dr. Goldstein attempted to show Sterling's reliance on 12 of the RX 250 articles listed in Table V (CPF at 462), solely by inference from the fact that these articles had been turned over to the FTC pursuant to a 1966 subpoena. Dr. Goldstein accordingly left to inference that the articles were supplied for the same purpose in 1971 as at trial, for which there is no evidence in the record. There is no evidence in the record to indicate for what purpose they were produced to the FTC in 1966. Sterling has then failed to carry the burden of showing that these articles were relied on by it for substantiation of the challenged advertising claims at the time of their dissemination. **[296]**

1202. Dr. Goldstein attempted to provide reliance on 71 of the RX 250 articles listed in Table VI (CPF at 463–467), solely by inference from the fact that for each article another article from a different edition of the same journal in which it appeared had been turned over to the FTC pursuant to 1973, 1971, or 1966 subpoenas. Dr. Goldstein testified that since the literature of interest to the medical staff at Sterling is too voluminous to keep abreast of, the Sterling library staff digests pertinent articles (Goldstein, Tr. 14778–80). His testimony left to inference the following: that Sterling subscribed to or regularly received these journals; that these journals were regularly received by Sterling's library staff; that these journals were regularly digested by Sterling's library staff; that the criteria which Sterling's library staff used in determining which articles should be digested encompassed each of these articles; that these articles were digested accurately, correctly, and adequately; that these digests came to the attention of the appropriate person on Sterling's medical staff; that this person read and understood these digests; that (s)he was sufficiently interested in the articles after reading the digests to request copies of these articles from the library; that this person received, read and understood these articles; and that this person relied on the information in these articles to substantiate the challenged claims. There is no evidence in the record to support any of these inferences. Respondent has thus failed to carry the burden of proof of showing that they relied on these articles at the time it made the challenged claims.

1203. Dr. Goldstein attempted to show that 16 of the RX 250 arti-

cles, listed in Table VII (CPF at 468–469), substantiate challenged claims. All of these articles were published after 1964, postdating the period when the challenged claims were made. There is no evidence in the record to indicate that respondent possessed or had knowledge of these articles prior to their publication date. Therefore, Sterling has failed to carry the burden of showing that it possessed and relied on this material for substantiation of the challenged claims at or prior to the time of their dissemination.

1204. Dr. Goldstein attempted to show Sterling possessed and relied upon 49 of the RX 250 articles, all published from 1970 to 1974, and list set forth in Table VIII (CPF at 470–473). There is no evidence in the record, nor any inference suggested, to indicate that Sterling possessed or had knowledge of any of these articles prior to their publication. Therefore, Sterling could not have possessed and relied on these articles to substantiate challenged claims that were disseminated prior to their publication date (CX 630–634).

1205. Dr. Goldstein attempted to show Sterling possessed and relief on RX 250-Wood solely by inference from the fact that [**297**] this paper was presented prior to its publication at a meeting of the American Pharmaceutical Association in New York in 1964 (G. Goldstein, Tr. 14940–41). This evidence relies on the following inferences: that a Sterling representative was present at this presentation; that complete and accurate copies of the unpublished RX 250-Wood were available; that a Sterling employee obtained a copy of RX 250-Wood at the presentation; that this employee brought the paper to the attention of the appropriate person on Sterling's medical staff; that this medical staff person read and understood the paper; that this person would have relied on this paper to substantiate challenged claims at the time of their dissemination; and that the paper presented at the APHA meeting was identical to RX 250-Wood in all material respects. There is no evidence inthe record to support any of these inferences. Respondent has thus failed to carry the burden of proof of showing it possessed and relied upon this article at the time it made the challenged claims.

1206. Dr. Goldstein attempted to show that Sterling possessed and relied upon three of the RX 250 articles listed in Table IX (CPF at 474), by inference from the fact that they are currently in Sterling's library. This evidence relies on the following inferences: that this article was in Sterling's possession at the time the challenged claims were disseminated; that the journals they appeared in were regularly reviewed by Sterling's library staff; that these journals were regularly digested by the library staff; that the criteria which Sterling's library staff used in determining which articles should be digested included each of these articles; that these articles were digested correctly and

adequately; that these digests would have come to the attention of the appropriate person on Sterling's medical staff; that this person would have read and understood these digests; that this person would have been sufficiently interested in the articles after reading the digests to request a copy of these articles from the library; that this person would have received these articles; that this person would have read and understood these articles; and that this person would have relied on the information in these articles to substantiate the challenged claims. There is no evidence in the record to support any of these inferences. Respondent has thus failed to carry the burden of proof of showing that it possessed and relied on these articles at the time it made the challenged claims.

1207. Dr. Goldstein attempted to show reliance on nine RX 250 articles, listed in Table X (CPF at 475), by inference from the fact that the articles were turned over to the FTC pursuant to a subpoena request. There is no indication in the record as to the year of the subpoena request pursuant to which each was produced. Dr. Goldstein left to inference that the articles were supplied for the same purpose in the subpoenas as at trial, [298] for which there is no evidence in the record. There is no evidence in the record to indicate for what purpose they were produced to the FTC. Sterling has thus failed to carry the burden of showing that these articles were relied on by it for substantiation of challenged claims at the time of their dissemination.

D. *Testimony of Witnesses Contacted by Sterling After Dissemination of Challenged Claims Do Not Provide a Reasonable Basis for Such Claims*

1208. In order for the expert opinion testimony adduced at trial to be relied upon by Sterling as providing a reasonable basis for advertising claims, such opinion must have been possessed and relied upon by Sterling at the time of dissemination of the challenged claims. *Pfizer, Inc.*, 81 F.T.C. 23, 67, 71 (1972).

1209. Sterling offered no evidence that Drs. Horner, Rhodes, Banker, Stander, Scoville, Fields, Feinstein, and G. Goldstein, were contacted by it prior to the commencement of this proceeding, and the opinion testimony cannot be relied upon by Sterling now to show reasonable basis for any of the challenged advertising claims involved in this proceeding.

VIII. SECTION 5 LIABILITY FOR DISSEMINATION OF "INCONSISTENT" ADVERTISING CLAIMS REGARDING DIFFERENT DRUG PRODUCTS

1210. As found hereinabove (F. 398–402, *supra*), Sterling made

contemporaneous and inconsistent advertising claims with respect to Bayer Aspirin, Vanquish and Cope.

1211. Complaint Counsel urge that contemporaneous dissemination by an advertiser of "inconsistent" performance claims for different OTC internal analgesic drug products is unreasonable and unfair and constitutes an unfair trade practice within the meaning of Section 5 of the FTC Act (CB at 110–116).

1212. Complaint counsel's theory of Section 5 liability would introduce a new legal requirement of "consistency" in advertising regulation. It would go beyond determining whether an advertising claim regarding a product is false or deceptive or otherwise unfair to consumers and place upon an advertiser the requirement that its advertising claims for different products it markets at a given time be consistent with each other. Complaint counsel have not spelled out the content of this new requirement, except to argue generally that, by definition, mutually inconsistent claims cannot be both true at the same time. Complaint counsel have not pointed to any authority for their view. It is the administrative law judge's view that such a requirement will have a chilling effect on [**299**] protected commercial speech by effectively placing every multiproduct advertiser at peril to take care that its advertising claims for all its products at any given time not be "inconsistent" although each claim may be based on a reasonable basis.

IX. CORRECTIVE ADVERTISING WITH RESPECT TO BAYER IS NOT WARRANTED

A. *There is Substantial Evidence Showing That a Significant Portion of Consumers Held Superior Efficacy/Safety Image for Bayer Aspirin From 1967 Through 1975*

1. Image and Advertising Penetration Studies

1213. An advertising penetration study measures the level and content of what consumers recall from past advertisements for a brand (Ross, Tr. 5802–03). Unlike a copy test which tests the consumer's recall of a specific advertisement immediately after or within a day of his or her exposure to it, a penetration study tests the consumer's recollection of advertisements for some period of time past, in terms of either the recent past or a specified number of months past (Ross, Tr. 5804). A penetration study accordingly reflects advertising themes from a variety of advertisements that may have been disseminated over some period of time prior to the measurement (Ross. Tr. 5804). On the other hand, an image study is aimed at measuring the level and content of consumer images of a particular brand or brands apart from advertising recall as such. To this extent, image

studies provide more direct and meaningful information regarding a brand's image than do penetration studies, which focus on consumer recall of advertising themes and can indirectly shed light on brand images as related to past advertising campaigns.

1214. There are three so-called image studies in the record: (1) CX 395, the Assets and Liabilities Study of Adult Analgesics ("Assets and Liabilities") by Dancer-Fitzgerald-Sample, Inc., dated December 1, 1967; (2) CX 404, A Study of Vanquish's Market Opportunities ("Vanquish Study"), by Benton & Bowles, Inc., November 1979; and (3) CX 521, the Consumer's view of the Relative Effectiveness of Various Brands of Aspirin ("The Zeisel Image Study") by Hans Zeisel, 1975 (Joint Hearings Tr. 426, 2055; Zeisel, Tr. 4742). The Assets and Liabilities Study and the Vanquish Study were done for commercial purposes by advertising agencies (CX 395, CX 404; Miller, Joint Hearings Tr. 209–10; Pernica, Joint Hearings Tr. 1891). The Zeisel Image Study was conducted at the request of complaint counsel for this proceeding (Zeisel, Tr. 4649–50).

1215. A brand image is a group of general impressions that people hold about a particular brand. It is the personality or [**300**] character of a brand (Miles, Tr. 9355; Haley, Tr. 10567, 10605–10; Lipstein, Tr. 12228).

1216. There are multiple factors that are involved in brand imagery: favorable or unfavorable experience with the brand word-of-mouth; nonverbal communication, such as the package, graphics, name, or price; favorable or unfavorable publicity; sheer longevity and visibility in the market; the amount of advertising regardless of content; advertising content; product innovation; store displays; promotional activities, and other sources. A brand image is not a direct reflection of advertising (Miles, Tr. 9355–59; Haley, Tr. 10567, 10605–10; Lipstein, Tr. 12037).

1217. Brand images are formed in different ways, both within and across product categories. For example, the image of a product that has recently been introduced into the market would stem more from advertising and positioning. The longer a brand is in existence the less its image stems from the content of advertising and, particularly, from the content of a specific advertising campaign (Miles, Tr. 9355–56, 9366; Haley, Tr. 15069).

1218. There are a number of cues that relate to the overall evaluation of a product as an excellent product, some of which come from the product itself, from price, from the label, from the length of time it has been on the market (Haley, Tr. 10670).

1219. The role of advertising in the formation of brand images varies according to the age of the brand. Newer products have images that are more malleable; for longer term brands, the images are

formed early and subsequent to that the role of advertising is primarily to remind consumers of the brand's existence (Haley, Tr. 10569, 10651).

1220. Brand imagery comes substantially from advertising for those brands where the advertising theme is the brand's reason for being and that theme is consistently advertised year after year. For example, Excedrin's reason for being was that it was an "extra strength" pain reliever, different from plain aspirin. This advertising claim was used consistently over time and as a consequence, advertising contributed heavily to its extra strength image (Miles, Tr. 9356).

1221. For an old established brand, such as Bayer, brand familiarity is the primary influence: predominance in stores, word-of-mouth, and length of time on the market are all very important. Advertising maintains the brand's salience but has little influence on the image (Haley, Tr. 10651).

1222. It is recognized that there are several basic principles in interpreting consumer research, including image research, which must be taken into account. **[301]**

1223. One of the fundamental principles that has been confirmed by studies over decades is that consumers will generally rate an advertised brand higher than an unadvertised brand. The fact of television advertising visibility indicates to the consumer that the brand is a well-known brand. The absence of advertising indicates the opposite. Thus, advertising produces a substantial effect on ratings in an image study whether or not people can remember any specific sales points from the nationally advertised brand (Miles, Tr. 9396–97).

1224. Complaint counsel's expert, Dr. Ross, acknowledged the principle that consumers will rate advertised brands better than nonadvertised brands and noted that there are studies that confirm it (Ross, Tr. 5863–64).

1225. Consumers generally tend to rate advertised brands higher than unadvertised brands on desirable product attributes, whether or not the attribute had been advertised or not advertised by the particular brand (Miles, Tr. 9389). This propensity is particularly pronounced with respect to generic attributes, such as being an effective or generally good pain reliever in case of an analgesic product.

1226. National brands are more highly regarded than non-national brands or store's own brands. Consequently, consumers will tend to rate national brands higher than store's own brands or unadvertised brands on most attributes (Ross, Tr. 5863–64; Miles, Tr. 9389; Amstutz, Tr. 10100–07; Haley, Tr. 10577–80, 10585; Lipstein, Tr. 12032).

1227. In a chapter written by Dr. Ross, "Applications of Consumer Information to Public Policy Decisions," contained in Scheth and Wright, *Marketing Analysis for Societal Problems,* Dr. Ross describes

"surrogate indicators of quality" and the first one listed is that nationally advertised brands are considered to be of higher quality than store brands (Ross, Tr. 5863–64).

1228. Users of a brand generally tend to rate their brand higher than nonusers of a brand (Ross, Tr. 5862–63; Miles, Tr. 9389, 9405; Amstutz, Tr. 10125; Haley, Tr. 10585; Lipstein, Tr. 12031).

1229. There is a high relationship between familiarity and awareness of a brand and high ratings of that brand (Haley, Tr. 10585, 10604; Lipstein, Tr. 12030).

1230. In viewing the image data in evidence, it is important to take into account the influence of Bayer's long-term position in the market. **[302]**

1231. Bayer Aspirin's brand image is in the entire analgesic market and not just in relation to all other 5–grain aspirin (Miles, Tr. 9360; Haley, Tr. 15076–78; Lipstein, Tr. 12028–29).

1232. The history of the analgesic market must be taken into account because it is an important factor in brand imagery. Bayer Aspirin was invented around the turn of the century and for years had the exclusive right to use the word "aspirin" (Alberts, Tr. 8988; Miles, Tr. 9361).

1233. The most important factor in Bayer's image is its heritage, its longevity, the length of time it has been in the market. It is the original aspirin (Miles, Tr. 9361; Haley, Tr. 10651).

1234. Advertisers often refer to the original brand, the founding brand of the category, as having the "grandfather rights" because such brands have certain characteristics in common: greater longevity, long-term visibility; they stand for integrity, purity, reliability, honesty, trustworthiness and lack of risk. Such brands generally have overall favorable images among users and nonusers alike (Miles, Tr. 9362).

1235. These grandfather rights do not come from the content of advertising, although the visibility and continued advertising generally helps to maintain them (Miles, Tr. 9364, 9366).

1236. The grandfather rights provide to the brand image a high regard, an overall generalized favorability irrespective of use. This generalized feeling may be reinforced for both users and nonusers by advertising weight regardless of specific content, by visibility in stores and people's homes; sometimes by nostalgia (Miles, Tr. 9364, 9366).

1237. Complaint counsel's witness Dr. Ross acknowledged that Bayer is a product with a long-standing, reliable position or "franchise" in the consumer's mind, analagous to Gerber's Baby Food and others in their respective fields. Dr. Ross discussed this point with reference to products which have long-established market positions, illustrating with Gerber's Baby Food the fact that the mere mention

of a brand name may cause a consumer to bring an understanding of the product in mind. In this situation, "the advertiser need only mention the brand name and a substantial percentage of the audience may understand the message as 'Here we are again, old reliable Gerber's, the safest, most nutritious, most reliable baby food your baby can eat.' " (Ross, Tr. 5866–68, 5880–81).

1238. Dr. Ross also agreed that Bayer Aspirin had "lineage," described as familiarity with the product, "my mother **[303]** used it," the good old reliable standby used at home (Ross, Tr. 6059–62).

1239. The image of Bayer Aspirin is of a mild, middle-of-the-road, overall high quality, no risk product. It is regarded with a generalized, overall favorability, a good brand that has been around for a long time (Miles, Tr. 9364, 9373).

(a) *The Assets and Liabilities Study (CX 395)*

1240. The 1967 "Assets and Liabilities Study of Adult Analgesics" (CX 395) was designed by Dancer-Fitzgerald-Sample, Inc. (DFS), and executed by Crossley Surveys for Sterling Drug, Inc. Its stated purpose was to "provide assets and liability profiles for Bayer Aspirin and other leading brands of analgesics products" and to "serve as a 'benchmark' against which data for future assets and liabilities studies may be measured" (CX 395F; Miller, Tr. 209–10). It is a replication of an earlier study that Crossley Surveys had done for DFS (Leonard, Tr. 88–89).

1241. Lloyd Miller, who designed CX 395, was and is Vice-President and Associate Director for Research of DFS. Mr. Miller testified concerning the design and analysis of the study.

1242. The sample for this study was a "multi-stage stratified area sample." The sample design provides for the selection of individual respondents by dividing the country as a whole into smaller and smaller units, from major markets to minor civil divisions to blocks, and from blocks to households. "Stratification" refers to that control on the sample which insures that it accurately represented particular demographic attributes of the population as a whole. Such stratification related to sex (50% male and 50% female), and to geography. The sample was designed to be representative of the U.S. population in terms of the proportional representation of the four geographic regions, three sizes of standard metropolitan statistical areas and one size of nonmetropolitan counties in the U.S. (Leonard, Tr. 95–96). Thirty-five primary units, or markets, were selected from a national probability sampling frame of 80 primary sampling units to be representative of the whole United States. Within those 35 markets, Crossley Surveys selected minor civil divisions in proportion to their relative population (Leonard, Tr. 97–98). Within individual divisions,

urban block clusters were selected systematically from census block statistics whenever that was possible. Once a particular block was selected, a random technique was used to designate a starting point on the block for interviewers to commence their interviewing. From that starting point, interviewers were given explicit instructions on which houses to contact (CX 1007). These instructions left no discretion whatever in the hands of the interviewer (Leonard, Tr. 100). **[304]**

1243. The sampling procedure outline above is consistently used by Crossley Surveys. It yields results upon which marketing decisions are made (Leonard, Tr. 102–05). The procedure was discussed with, and explicitly approved by, DFS (Leonard, Tr. 102).

1244. The questionnaire for this study consisted of a notebook with 31 pages. Each page was a self-contained rating scale on a separate attribute, positive ratings at the top and negative ratings at the bottom. The rating of the products was to be made by the interviewees themselves inserting cards bearing the names of products into one of six pockets, corresponding to the intensity of their feeling about those products on each attribute (CX 395D, Z158–Z160).

1245. The design of CX 395 was similar to that of other image studies commissioned by DFS (Leonard, Tr. 86–88), and the "Assets and Liabilities" type of notebook-questionnaire used in this survey had been used by DFS since 1953 or 1954 for major clients such as General Mills and Falstaff Brewing Company (Miller, Tr. 214). This study design is comparable in quality to others for measuring images of products (Leonard, Tr. 94).

1246. CX 395 was executed according to Crossley Surveys' normal survey procedures. Most of the field work supervisors and interviewers on the project were people with whom Crossley Surveys had had substantial favorable experience (Leonard, Tr. 107). All interviewers were personally briefed by their supervisors and provided with detailed written instructions for administering the questionnaire (Leonard, Tr. 87, 107–10; CX 1000, 1002).

1247. Validation of interviews at Crossley Surveys was a two-step procedure, conducted both by interviewer supervisors and then by Crossley's headquarters (Leonard, Tr. 110, 115, 138–39; CX 1001). This process provided a total of 15% of total interviews validated. As a third check on the interviewers' work, DFS itself validated an additional 10% of the interviews (Miller, Tr. 229–30).

1248. Coding of the results of the survey was performed by Crossley's editing and coding department. A trained, experienced editor was normally responsible for that task. Given the absence of open-ended questions on the questionnaire necessitating interviewers' recording verbatim responses, coding for this project was a ministerial task. After the coding and editing tasks were accomplished by Cross-

ley, the results were delivered to DFS who analyzed them and prepared the report (Leonard, Tr. 115–16; Miller, Tr. 235).

1249. CX 395 was not conducted in anticipation of litigation. Sterling was DFS's client and requested the study **[305]** in the regular course of business. Sterling was satisfied with the quality of the work and its presentation (Miller, Tr. 209–10, 235–36). Crossley Surveys itself had no direct contact with Sterling nor any interest in any particular outcome of the study (Leonard, Tr. 87). As a result, there was no reason for the survey to be biased.

(b) *Study of Vanquish's Market Opportunities (CX 404)*

1250. The 1970 "Study of Vanquish's Market Opportunities" (CX 404) was designed by Benton and Bowles, Inc., an advertising agency for its client, Sterling, as part of the development of an advertising campaign for Vanquish. CX 404 was designed to measure consumers' attitudes toward analgesics in general, their opinion of some leading analgesic brands, including Vanquish, and to determine what sort of consumer Vanquish was most likely to attract (CX 404E). The backup data to CX 404 was received in evidence as CX 440.

1251. Joseph Pernica, the Associate Research Director and Vice-President of Benton and Bowles, Inc. at the time of the Vanquish Study had full responsibility for developing the design, methodology, and questionnaire for the survey, and for overseeing its execution (Pernica, Tr. 1893). He testified for complaint counsel concerning those substantive areas.

1252. Lieberman Research Corporation of New York was responsible for executing CX 404. Arnold Fishman, the Vice-President of Lieberman Research, testified for complaint counsel concerning the procedures used for conducting the study, including sampling procedures, interviewing, and coding and tabulating.

1253. The sampling procedure for the 1970 Vanquish Study, CX 404, was developed by Lieberman Research according to specifications set by Joseph Pernica of Benton and Bowles. These specifications included the sample size, the number and type of markets in which the survey would be conducted and the desired 50/50 sex distribution of the respondents. Benton and Bowles instructed Lieberman to investigate the Mid-Atlantic and Pacific region and also wanted to concentrate some interviews in three high-share Vanquish markets, Atlanta, New Orleans and Oklahoma City. Lieberman Research was given a list of cities in the Mid-Atlantic and Pacific regions to choose from and it chose the ones in which it had the best interviewers (Fishman, Tr. 1292–93; Pernica, Tr. 1918–19).

1254. Within each market chosen, the sample was randomly selected from addresses listed in telephone books. A random number was

picked as the one which to enter each phone book, and to get to successive pages, a skip interval equal to the number **[306]** of remaining pages divided by the number of desired interviewing clusters was determined. In order to minimize the sampling error due to use of telephone listings, interviewers were instructed to interview a resident of the house adjacent to the one picked from the phone book (Fishman, Tr. 1299–1301). This procedure left no discretion to the interviewer in selecting respondents.

1255. This sampling procedure was standard at Lieberman Research, and the sampling instructions given to interviewers were the company's standard written instructions (Fishman, Tr. 1339–40, 1300). It was not designed to produce a national probability sample. However, the degree of deviation from strict adherence to all probability standards in this sampling pattern was small enough that Lieberman typically recommended that marketing decisions could be made based upon the data generated (Fishman, Tr. 1367–68).

1256. The Vanquish Study (CX 404) was based on personal interviews. The questionnaire was carefully reviewed and revised by Arnold Fishman at Lieberman Research in order to eliminate ambiguities and to ensure correct question order. After it was put into final form, it was pretested in the field to ensure that it could be easily administered. The pretesting indicated that there were no significant problems with the interview (Fishman, Tr. 1295–97). Lieberman Research chose its interviewers and supervisors carefully, using only supervisors who were known to have done timely work of high quality in the past, and encouraging the supervisors to use only their best interviewers. The supervisors were responsible for training interviewers, for passing on Lieberman Research's standard written instructions, for acting as intermediaries between them and the central office, and for validation of the interviewer's work. Lieberman did not rely solely upon the supervisor's validation, but validated an additional fifteen percent (15%) of all questionnaires in the central office. If validation of an interview uncovered a problem, all of the work of the interviewer who conducted it would be validated. In addition to these two validations, a third validation check was run by an outside service to ensure objectivity (Fishman, Tr. 11317–18).

1257. Coding, keypunching and tabulations were performed by Lieberman Research according to its normal procedures for studies of this type. The codes for open-ended answers were developed by Lieberman Research's coding staff under Arnold Fishman's supervision. Joseph Pernica, of Benton and Bowles, approved the final codes (Pernica, Tr. 1929). A portion of every coder's work was checked by the coding staff supervisors to verify that coders were correctly interpreting verbatim responses (Fishman, Tr. 1319–21). Keypunching and

tabulations were performed by Data Probe, a research computer company selected by Lieberman Research with the approval of Benton and Bowles, Inc. All of the coded questionnaires were [307] "machine-cleaned" (checked for the logic of responses) and all the keypunching was verified as accurate by machine at Data Probe. Data Probe produced the tabulations of the results, CX 404, according to specifications set by Benton and Bowles, and Lieberman Research checked the tables for conformity with those specifications. Mr. Pernica received the tabulations from Lieberman Research and used them as the basis for his analysis presented in CX 404 (Fishman, Tr. 1321–25; Pernica, Tr. 1929–30).

(c) *1971 Advertising Penetration Study (CX 565)*

1258. CX 565 was designed and analyzed by Ted Bates & Company, Inc. ("Bates"), and was conducted by Valley Forge Information Service (hereinafter "Valley Forge"), a wholly owned subsidiary of Burlington Industries, at the request of and for the benefit of Bristol-Myers Corporation (CX 1019–1020). Its purpose was to measure the advertising penetration of Bufferin and other OTC analgesics (CX 565E-K, CX 1009). The questionnaire design is typical of earlier Bates penetration studies, many of which are also performed for Bristol-Myers Corporation. Employees of both Bates and Valley Forge testified that the questionnaire was typical of those used in assessing advertising penetration (Weitz, Tr. 731; Fratto, Tr. 810).

1259. Bates is the advertising agency for the Bristol-Myers Company for the Bufferin account. Ms. Anne Jack,[6] a Vice President of Bates, testified for complaint counsel regarding the design and analysis of CX 565. Bates' research department performs a wide range of research on all types of products for its clients (Weitz, Tr. 809).

1260. Valley Forge designed the sampling plan for this survey. The first step was the construction of a "master probability sample." This was obtained by dividing up the entire country, according to published photostats from the Census Bureau, first into a census region, and then into four city-size classifications within the census regions. The "sampling points" within the four city-size classifications are randomly selected from within the counties listed in each classification. While one could obtain any number of sampling points, the 100 points used in this survey were found more than adequate (Fratto, Tr. 737–38). [308]

1261. The telephone numbers of individual survey respondents were selected randomly from within these sampling companies for each county in the master sampling plan, a standing order being

[6] When Ms. Jack testified about CX 565, her name was Anne Weitz. Therefore all citations to her testimony refer to "Weitz."

placed with each company to ensure that the directories were recent. If, for example, 1,000 complete interviews were required, 2,500 numbers would be selected, 25 from each of the 100 sampling points in the master sample. A randomized "skip pattern" within each phone book, starting from a random starting point, would also be established (Fratto, Tr. 738–40).

1262. All interviewers were instructed orally about the correct way to select a particular column on a page and a particular number down in that column. In other words, the smallest detail was attended to as carefully as the drawing of the original master sample (Fratto, Tr. 739–41). In order to minimize a nonresponse bias, each number at which there was no response received two call backs (Fratto, Tr. 744).

1263. The questionnaire was quite easily administered, because it required no skips and only very simple probes (CX 1009). Nevertheless, all interviewers received both written and oral instructions in conducting the interviews (CX 1021; Fratto, Tr. 740).

1264. In addition, training of the interviewers involved actual testing of their ability by supervisors who had at least one year's experience in interviewing and who were experienced in dealing with people (Fratto, Tr. 724). This degree of care in conducting interviews was a standard procedure at Valley Forge (Fratto, Tr. 720).

1265. The interviewers' WATS lines were connected to a monitoring facility so that each interview could be listened to as it was conducted without the interviewers being aware of the monitoring process (Fratto, Tr. 742). In addition, all completed questionnaires were checked by Valley Forge's supervisors for thoroughness and accuracy. Finally, there would be a third check by a group of editors who would review the questionnaires before they were sent to the client (Fratto, Tr. 745).

1266. Coding, keypunching and tabulation were performed by Bates after it received the completed questionnaires (Fratto, Tr. 745). Because the questionnaires contained open-ended verbatim responses, Bates employees expended a large degree of time and effort in developing appropriate codes for the verbatims despite the fact that the basic framework for coding had been developed during earlier Bates market penetration studies (Weitz, Tr. 823–24; CX 1016).

1267. The mechanics of coding and tabulating were performed by hand by Ms. Jack herself and a trainee under her close supervision (Weitz, Tr. 826). **[309]**

(d) *The 1973 Headache Remedy/Pain Reliever Usage and Advertising Penetration Study (CX 553)*

1268. CX 553 was designed to determine current advertising penetration and usage levels of selected analgesics (CX 553C). The study

was designed, executed and analyzed by Sobel-Chaikin Research Associates at the request of and in cooperation with American Home Products Corporation (American Home) (Sobel, Tr. 461–64). Sobel-Chaikin Research Associates is the research division of Market Probe International (hereinafter, "M.P.I."), an organization formed in approximately 1974 to perform market research, computer analysis and data processing for manufacturers and advertising agencies (Sobel, Tr. 451–53).

1269. Charles Sobel testified for complaint counsel regarding both the design and the execution of CX 553 for which he had ultimate responsibility. Mr. Sobel is Senior Vice-President and Director of the research group at M.P.I., and the founder of Sobel-Chaikin Research Associates.

1270. The study design called for a telephone sample to be randomly selected from telephone books in 10 major urban markets (CX 553C, CX 1007; Sobel, Tr. 467–68). Interviewers in each market were assigned a random starting page in the telephone book for the market and were instructed to skip a random interval number in order to obtain each succeeding page (CX 1007). They were instructed to start at the top of the second column on each page and proceed down the column until they had completed a series of five interviews. These instructions clearly left no discretion to the interviewer in the selection of respondents (Sobel, Tr. 467–68).

1271. The questionnaire for this survey was short, and it was easy to administer because it contained few skip patterns for interviewers to follow (CX 553Z101–Z104). The questions were unambiguous and were directed both to advertising recall and usage of analgesics. The questionnaire was developed in consultation and with the approval of American Home, and it was typical of questionnaires used previously by Sobel-Chaikin for advertising penetration studies (Sobel, Tr. 461–62, 484).

1272. The survey was conducted according to standardized procedures followed by Sobel-Chaikin Associates in all their research work. All interviewers received extensive instructions regarding the administration of the questionnaire and were personally trained by supervisors who were known either to the principals of the firm or to one of their field supervisors, on the basis of prior favorable experience (Sobel, Tr. 471–72). **[310]**

1273. Completed interviews were validated in a two-step procedure. Supervisors were instructed to validate work received from all their interviewers. In addition, 15% of the completed interviews submitted by supervisors were validated by an outside validation service hired by Sobel-Chaikin (Sobel, Tr. 477–81).

1274. M.P.I.'s in-house coding department coded the responses on

the completed questionnaires. The task involved building codes for verbatim responses to open-ended questions on the questionnaire asking about advertising recall. The final codes were prepared by Mr. Sobel and were approved by American Home. Checks on the quality of coding were supplied by M.P.I.'s coding supervisor and by having individual coders re-do each other's work for comparison purposes (Sobel, Tr. 483–85; CX 1005, 1006).

1275. M.P.I.'s own data processing group keypunched the completed questionnaires. The keypunching was performed by experienced operators and was checked both by verification and by automatic controls placed into the computer programing that produced the tabulation runs. The tabulation plan was developed in accordance with specifications approved by American Home. The report of CX 553 was prepared by Mr. Sobel and as submitted to American Home (Sobel, Tr. 484–87).

(e) *Zeisel Image Study (CX 521)*

1276. The purpose of CX 521, "The Consumer's View of the Relative Effectiveness of Various Brands of Aspirin" (Zeisel Image Study) was to identify consumers' images of the relative effectiveness of various brands of 5-grain aspirin. More specifically, CX 521 measures the comparative image of Bayer compared to other brands of aspirin with respect to effectiveness and speed of relief (Zeisel, Tr. 4649; Crespi, Tr. 4341; CX 521B). This study was conducted for use in this litigation by Dr. Hans Zeisel and the Gallup Organization under contract with the FTC.

1277. The principal author of CX 521 is Dr. Hans Zeisel. Dr. Zeisel was primarily responsible for the design of the study, the design of the questionnaire, the designation of samples, and the drafting of the final report (Zeisel, Tr. 4650–01). The Gallup Organization and Dr. Irving Crespi, then of Gallup, participated in the design questionnaire and executed the fieldwork for the study (Crespi, Tr. 4341; Zeisel, Tr. 4722–23). The fieldwork for CX 521 was conducted in substantially the same fashion as the fieldwork for CX 520, the Zeisel Copy Tests (Crespi, Tr. 4345–52).

1278. Dr. Crespi reviewed the draft questionnaire to CX 521 and pretested it to see that it conformed to good professional practice. The pretest indicated that, without exception, the [311] questionnaire was professionally acceptable. The pretest did show that some respondents failed to rate a brand of aspirin because they had never used that brand. Because the objective of the study was to measure images of brands among consumers who were familiar with a particular brand regardless of whether or not they used it. Question 4 was revised, explicitly instructing respondent to rate all the brands he or she was

familiar with regardless of whether or not the respondent had used the brand. With this change, the final questionnaire in CX 521 (CX 521Z002–Z005) fully conformed to good professional practice and was considered to be a "state of the art" questionnaire for consumer image research (Crespi, Tr. 4341–45).

1279. All surveys based upon probability samples are subject to sample error. "Sample error" is the measure of the extent to which the survey results may differ from the "true value" that would be obtained if the whole population was interviewed. Appendix III of CX 521 sets forth the estimated "sample error" at the 95% confidence level for the percentages reported in CX 521. Appendix III indicates the range (plus or minus the figure shown) within which the results of repeated sampling will occur 95 times out of 100 (CX 521Z014; Crespi, Tr. 4324–25, 4347).

1280. The response rate to CX 521 was about 60%. This response rate meets generally accepted standards for research of this nature (Crespi, Tr. 4351).

1281. The interviewers' work was validated by Gallup in the same manner as set forth in F. 219 (Crespi, Tr. 4350). The final and completed questionnaires were put through a standard quality control procedure to verify that the interviews had been conducted in accordance with instructions. The answers to CX 521 were statistically weighted according to demographic characteristics. This procedure is described in F. 221, *supra* (Crespi, Tr. 4351–52).

1282. The tabulated results of responses to the questions asked in CX 521 are set forth in Tables One through Nine (CX 521I-Y).

1283. After the Gallup interviewer selected the designated person within the household to be interviewed, the respondent was first asked whether he or she uses pain relievers. If the answer was "no," then the interview was terminated (Crespi, Tr. 4346). For those who indicated that they used pain relievers, they were then asked what brands of aspirins they have ever heard of or bought. The Gallup interviewer recorded all such responses. Next, the respondent was given a set of cards, each of which contained the name of a brand of aspirin, and then was asked if he or she had ever bought or heard of any of those brands of aspirin. The brands on the cards were A&P, Bayer, McKesson, Norwich, Rexall, Safeway, Squibb, St. Joseph's Aspirin [312] for Adults, and Upjohn. The Gallup interviewer then recorded all such responses. If the respondent mentioned two or more brands of aspirin, the respondent was then asked to rate on a scale of one to ten how effectively and how quickly those brands work. Finally, standard brand usage and demographic questions were asked of the respondents (CX 521G, Z002–Z005; Zeisel, Tr. 4726–33).

1284. There were 501 respondents interviewed in CX 521 (CX 521G).

Respondents were selected on a random, probability basis for Gallup's master probability sample as set forth in F. 217, *supra* (Crespi, Tr. 4345; CX 521Z009–Z012).

1285. The purpose of CX 521 was to determine how consumers of pain relievers viewed the various brands of aspirin which they had bought or heard of (CX 521B). The brand image of Bayer vis-a-vis various other plain 5–grain aspirin products were studied. A brand image is a group of general impressions that people hold about a particular brand. It is the personality or character of a brand (Miles, Tr. 9355; Haley, Tr. 10567, 10605–10; Lipstein, Tr 12228).

1286. Respondents argue that there are literally dozens of factors that are involved in brand imagery: favorable or unfavorable experience with the brand word-of-mouth; nonverbal communication, such as the package, graphics, name, or price; favorable or unfavorable publicity; sheer longevity and visibility in the market; the amount of advertising regardless of content; advertising content; product innovation; store displays; promotional activities, and other sources. A brand image is not a direct reflection of advertising (Miles, Tr. 9355–59; Haley, Tr. 10567, 10605–10; Lipstein, Tr. 12037).

1287. However, the crucial determination is not whether there is a multitude of sources for a particular brand image, but rather, whether advertising "in part" created a false impression or "played a substantial role in creating or reinforcing" a false and material belief. *Warner-Lambert Co.*, 86 F.T.C. 1398, 1499, 1503 (1975).

1288. The role of advertising in the formation of brand images varies according to the age of the brand. Newer products have images that are more malleable (Haley, Tr. 10569, 10651).

1289. The results of CX 521 are projectable with acceptable confidence to the total noninstitutionalized U.S. population of persons 18 years or older who use over-the-counter pain relievers (CPF 148–159; Crespi Tr. 4345).

1290. Subjects were asked to identify all brands of plain 5–grain aspirin products they had ever bought or heard of. To the extent that subjects did not recall on an unaided basis any of **[313]** nine nationally available aspirin brands, they were asked if they had ever bought or heard of those brands on an aided basis by being shown cards on which were printed the brand names (CPF 152; Zeisel, Tr. 4726–29). Subjects were then asked in questions 4 and 5 of the survey to rate each of the brands they recalled, whether they had used the brand or not, first on the basis of "how effectively it relieves pain," and second "how quickly it relieves pain," on a ten-point scale with a verbal anchor at the tope, "outstanding" (10), and at the bottom, "very poor" (1) (CPF 152; Zeisel, Tr. 4731–33). The use of ten points and verbal anchors were appropriate in the context of this survey (Crespi, Tr. 4342–44).

Data from these two questions were generated by considering simply whether each subject rated each brand higher, lower, or the same on each scale vis-a-vis Bayer, regardless of what points (s)he chose on the scales for the brands rated. Data for this question is reported as percent of ratings of Bayer as higher, the same, or lower than other brands (Zeisel, Tr. 4742–43; CX 521M, Q). The percent of consumers who hold comparative images of two products can be assessed by determining their beliefs about *both* products (CX 521; Ross, Tr. 5828–29). The sixth question asked which brands of pain reliever each subject used "most often." If the subject failed to identify a plain 5–grain aspirin product, (s)he was asked if (s)he ever takes aspirin and if so, what aspirin. With respect to the percentage of respondents who indicated that they used Bayer most often, the study separately reported their images on effectiveness and speed of relief (Zeisel, Tr. 4738–39; CX 521V, X).

1291. As an initial step in the analysis of the data generated in CX 521, the Zeisel Copy Tests, responses were assigned sample weights. This is done routinely by large survey organizations in national probability samples to adjust for discrepancies between the population in fact sampled and what the actual U.S. population is like. These estimates of what the U.S. population is like are based on annual census reports on population demographics such as age, sex, education or region of the country. Weights are assigned to responses in order to bring the characteristics of the survey population as close as possible to the U.S. population (Crespi, Tr. 4339–40, 4352). By inadvertance a few respondents were assigned "0" weights, which effectively eliminated their responses from the study. This error was insignificant and affected none of the results by any more than 1% (Crespi, Tr. 4362–63; Zeisel, Tr. 4819–20).

1292. Several criticisms of the Zeisel image study, CX 521, were offered by respondent's witnesses. The first criticism was that Dr. Zeisel had failed to take into account the possible relationship between the subject's brand awareness **[314]** and his rating (Amstutz, Tr. 10095). A second, related criticism was that CX 521 made no adjustment for the "halo" phenomenon (Haley, Tr. 10599). A third criticism was that the ten-point rating scale was inappropriately long (Haley, Tr. 10600). A fourth was that the results of CX 521 are explainable by the phenomenon of user loyalty.

1293. Dr. Amstutz was the major proponent of the argument that the Zeisel image study failed to consider brand awareness vis-a-vis rating. He undertook to remedy this in RX 142, a reanalysis of data from CX 521 prepared under his supervision (Amstutz, Tr. 10096). Dr. Amstutz's approach was to analyze ratings depending on whether subjects' recollection of brands was unaided, probed or aided, and

depending on whether subjects placed the brands they rated on the 1–5 end of the scales as opposed to the 6–10 end. His assumption in making the latter cut in the data was that a 1–5 rating is a low rating for any aspirin and a 6–10 rating is a high rating (Amstutz, Tr. 10397–99).

1294. This assumption was rejected by another of respondent's witnesses, Professor Russel Haley, who characterized as inaccurate a procedure whereby all the 6–10 ratings are treated as high and all 1–5 ratings are treated as low (Haley, Tr. 10704). People use scales differently; one person's "5" could be a high rating to him or her whereas another person's "6" could be low. CX 521 accounts for these differences between subjects by considering each person's relative ratings separately, each subject in effect becoming his own control (Haley, Tr. 10696).

1295. Thus the process used by Dr. Zeisel reduces concern for error in analysis that arises from the fact that different people use different portions of the scale or don't use all points of the scale (Haley, Tr. 10704). In any event, regardless of any error in the reanalysis done in RX 142 caused by Dr. Amstutz's division of the data into 1–5 and 6–10 ratings, Dr. Amstutz conceded that the relationship of subjects ratings of brands and their brand familiarity in RX 142 does not exclude the possibility that there are other factors than awareness, such as advertising content, that are leading to the efficacy ratings (Amstutz, Tr. 10435).

1296. Dr. Haley testified that the ten-point scale used in CX 521 is one shown by research to be less reliable because it uses too many points. Respondents were given too many choices and experience shows that it is dangerous to give any significance to a one-point difference on this type of scale (Amstutz, Tr. 10118; Haley, Tr. 10600–01; CX 521Z006, Z007).

1297. Professor Haley's testimony with respect to his concern about the length of the rating scales used in CX 521 seems to conflict with his previously expressed views. Despite [315] the fact that he testified that the ten-point scale with verbal anchors used by Dr. Zeisel was "not usual," Professor Haley admitted that in his article, CX 709 for identification, "Testing Thirteen Attitude Scales for Agreement and Brand Discrimination," published in the Fall 1979 Journal of Marketing, he evaluated a ten-point scale with verbal anchors as "commonly used" (Haley, Tr. 10693). He also conceded that when a numerical scale is used subjects have to be given verbal instructions so that they know which end of the scale is good and which end is bad because otherwise some subjects might regard ten as highest and others regard one as highest (Haley, Tr. 10695). Moreover, Professor Haley's opinion concerning the ten-point scale was contradicted by other ex-

pert testimony. Dr. Irving Crespi, an expert in the design and execution of consumer research and an expert with considerable experience with the use of rating scales such as those used in CX 521, said that the scales here were appropriate under the circumstances (Crespi, Tr. 4311, 4342–44). Dr. Amstutz acknowledged a variation of opinion regarding ten-point vs. six-point rating scales (Amstutz, Tr. 10286).

1298. Respondent also argues that there are other significant problems with the rating scale itself: For example,

(a) The scale used in CX 521 is also less reliable because it combines both words and numbers. In this situation, some people will use the words; others would use the numbers; while still others would use a combination of the two. Put together, they are not additive in statistical terms, and if the scale were repeated the next day with the same person, he might change from words to numbers or vice versa, and give different reactions. It is much better to utilize either numbers or words (Haley, Tr. 10600–01; *see* CX 521Z006, Z007).

(b) When shown the scale, consumers will think they are rating overall quality, not effectiveness or speed (Haley, Tr. 10715; *see* CX 521Z006, Z007).

(c) A substantial problem with the scale is the verbal anchors themselves. On the effectiveness scale, CX 521Z006, the verbal anchors should have been a phrase related to effectiveness, such as "completely effective" and "completely ineffective," rather than the overall general terms, "outstanding" and "very poor," that were used. Similarly, on the speed scale, CX 521Z007, the verbal anchors should have been something like "extremely fast," "extremely slow," rather than the same overall general terms, **[316]** "outstanding" and "very poor" (Miles, Tr. 9676; Haley, Tr. 10601; Lipstein, Tr. 12207–08; *see* CX 521Z006, Z007).

(d) Based on Dr. Haley's experience in scaling, when a scale says "outstanding" to "poor," people tend to give an overall evaluation. This is borne out by the similarity of the results obtained in the two questions (Haley, Tr. 10676; RPF 6.111; *see* CX 521N, P).

(e) Respondents would have focused on the scale itself and were unlikely to notice the heading at the top of the page. The respondents would not look to the top because their attention would be focused on the scale to try and understand the task they were asked to perform, particularly if it was unfamiliar to them (Haley, Tr. 10716; *see* CX 521Z006, Z007).

(f) The format of the two scales is so similar that it is likely that people will not note the change in the one word in the heading from "how effectively does it relieve pain" to "how quickly does it relieve pain." The data bears this out. The results of the two scales questions

are almost identical (Haley, Tr. 10601, 10676; *see also* Lipstein, Tr. 12207–08; Miles, Tr. 9416; CX 521N, P).

(g) In addition, the method of administration compounds the problems caused by the use of the scale. The interviewer reads a long paragraph to each respondent which, although briefly mentioning "effectiveness" and "speed" at the beginning, explicitly referred to the verbal anchors "outstanding" and "poor" throughout the instructions, concluding with the statement: "Rate them just the way you feel about each brand." This will lead to an overall rating by the respondent in terms of whether the brand is considered to be "outstanding" or "poor," rather than ratings related to "effectiveness" and "speed" (Haley, Tr. 10602–03; Lipstein, Tr. 12207–08; *see* CX 521Z004).

(h) If meaningful verbal anchors such as "relieves pain effectively" and "does not relieve pain effectively" had been used, they would have tied respondents to the scale and focused them on the issue of concern, instead of inviting them to rate a brand according to a general reputation or [317] aura, *e.g.,* a high rating for a national brand (Amstutz, Tr. 10118–20).

1299. However, the evidence is persuasive that the design of CX 521 is basically sound for the purpose for which complaint counsel seek to rely on in this proceeding, although care must be taken in interpreting the responses to the various questions.

1300. Respondent's expert witness, Dr. Amstutz, testified that, in his experience, users of a product are more apt to give their brands higher positive ratings on generic attributes (Amstutz, Tr. 10125–26; RX 142M, Table H).

1301. CX 521 is useful in judging consumer images of Bayer's relative effectiveness and speed vis-a-vis other plain 5-grain aspirin in 1975. Results of CX 521, reported at page C, as corrected by Dr. Zeisel on November 13, 1979) were that 40% of consumers (weighted) rated Bayer higher than all other brands with regard to effectiveness; 39% rated Bayer higher with regard to speed of relief. Of the remainder 34% and 35% respectively rated all brands equally on effectiveness and speed and 15% and 13% respectively rated Bayer highest with other brands. These results are projectable to the noninstitutionalized U.S. adult population who use OTC pain relievers.

1302. In his evaluation of the Bayer image in 1975, Dr. Ross considered responses to questions 4, 5 and 6 of CX 521 (Ross, Tr. 5826). Dr. Ross prepared a chart, CX 541, in order to look at the image of Bayer compared to other aspirin in a way that removed or adjusted for user bias. This chart incorporates an analysis of the way in which Bayer was rated among its users and the way other aspirin brands were rated by their respective users (Ross, Tr. 5827).

1303. In order for a respondent to be included in CX 541 (s)he would have had to have rated another brand besides the one (s)he used most often now. In other words, the respondent must have had a belief about more than one brand in order to qualify as having a comparative image (Ross, Tr. 5828–29). Results reported at page CX 541A demonstrate that 80 out of 136 respondents who reported that they now used Bayer most often rated best in effectiveness, while only 3 of 35 who now used another brand most often rated that brand above all others on that attribute. Eighty out of 136 versus 3 out of 35 is "clearly" statistically significant, though Dr. Ross went on to confirm this by performing a chi square analysis (Ross, Tr. 5828–29). This analysis showed that the image of Bayer among its users is superior in terms of efficacy compared with the image of other 5–grain plain aspirin among its users (Ross, Tr. 5829; CX 541A).

1304. Dr. Ross performed the same analysis on the attribute "speed," reported at CX 541B. Here, 74 out of 135 rated their [318] product, Bayer, as best compared with other aspirin as opposed to the 5 out of 35 who rated their non-Bayer 5–grain aspirin as best. The lopsided nature of the difference in the "speed" image of Bayer among its users, compared with other aspirins' speed images among their users, led to a statistically significant chi square value (Ross, Tr. 5829; CX 541A-B). Both speed and effectiveness results led Dr. Ross to conclude that the image of Bayer is superior to the image of aspirin on these measurements of effectiveness (Ross, Tr. 5829).

1305. CX 541 was prepared to account for user bias (Ross, Tr. 5827, 5829). According to Dr. Ross, these tables remove whatever contribution user bias might have had to the ratings by consumers because the analysis looks only at the image of each brand among its respective users (Ross, Tr. 5829). From CX 541, Dr. Ross concluded that, in 1975, the image of Bayer versus other plain 5–grain aspirin, was that Bayer was a superior product with respect to effectiveness and speed (Ross, Tr. 5830). Because speed is an indicium of effectiveness, Bayer was viewed generally as superior in terms of effectiveness (Ross, Tr. 5830).

1306. Dr. Ross testified that his conclusion that consumer images of Bayer's superior therapeutic effectiveness are not a consequence of user bias is supported by Tables 8 and 9 of CX 521 appeared at pages X and Y. Table 8 shows that 24% of the respondents not using Bayer rated it "best" as against other aspirin in terms of effectiveness; 23% of respondents not using Bayer rated it best in terms of speed. These nonusers could not have had their images of Bayer contributed to by usage of the product. Therefore, the image must have come from some other source, not use (Ross, Tr. 5830–32; CX 521X and Y). This data led Dr. Ross to conclude that a substantial number of people held an image of Bayer in 1975 that it was therapeutically superior to other

aspirin and this image would not be explained as a result of use or user bias (Ross, Tr. 5832).

1307. Two advertising penetration studies in evidence, the 1971 Ted Bates Advertising Penetration Study (CX 565), and the 1973 Sobel-Chaikin Study (CX 553), included questions related to Bayer Aspirin advertising penetration. Both surveys asked subjects to identify what recent advertising said about Bayer Aspirin. The 1973 study (CX 553) differed slightly from CX 565 in that it asked first whether the subject had seen or heard any recent advertisement for any headache remedies or pain relievers, and if so, what. This question generated an "unaided" response, *i.e.,* a response that was elicited by a question that made no reference to any product. Respondents were then asked whether they had seen any advertisement for specific products by brand name, including Bayer. Neither "aided" question in either study (*i.e.,* aided in the sense that the questions referred to brand names such as "Bayer") added anything to suggest the content of the advertising (Ross, Tr. 5810–11). **[319]**

1308. Evidence from CX 565 (the Ted Bates Advertising Penetration Study) confirms that consumers remembered Bayer's competitive effectiveness claims. Page K of CX 565 shows that 48% of the sample surveyed reported recall of Bayer advertising, and 16% of the sample recalled claims of competitive superiority for Bayer ("Competitive Superiority, Net"). It is not known exactly what was tabulated within that 16%. However, Page U, showing breakdowns of "Net Competitive Superiority, 16%," indicates that the therapeutic superiority category with respect to Bayer could be as low as 3% ("stronger" and "safer") or as high as 5% ("stronger", "safer" and "more relief"). The 8% shown for "Better/best/more effective" on S may also indicate the upper limit of "therapeutic superiority" recall, although the probability that 8% may include "better/best" in "better/best for quality," sense as some pre-1970 Bayer advertisements expressly claimed, cannot be excluded.

1309. With respect to CX 553, the 1973 study, about 51% of the respondent recalled Bayer advertising (Ross, Tr. 5811). Other tabulations show that therapeutic superiority claim recall for Bayer could be as low as 8.6% ("faster acting" - CX 553Z46) or as high as 25.1% ("best, better than other aspirin" - CX 553Z46), although there is a distinct likelihood that the 25.1% figure would also include "best, better than other aspirin for quality" as most of the Bayer advertisements in the early '70's expressly claimed. However, Dr. Ross found no penetration of advertising themes relating to manufacturing quality in CX 553.

1310. Thus, the advertising penetration data in evidence generally

show that a significant portion of consumers in the 1970's remembered Bayer's claim of therapeutic superiority.

1311. In the context of analgesic advertising, when consumers believe that the attributes of a particular OTC product make it perform better (*e.g.*, faster, safer, or more effectively than another product), they also believe that the superiority of that product on those attributes has been supported by scientific evidence (Ross, Tr. 5756). Were this not the case, consumers assume that the advertiser would be prohibited from making that statement, most typically, by the government (Ross, Tr. 5756). Dr. Ross' opinion in this regard is confirmed by a Food and Drug Administration study entitled "A Study of Health Practices and Opinions" dated June 1972. At page 270 the survey reported that 38% of American adults agreed with the statement, "Most of the things that advertisements say about medicines and health aids must be true or they wouldn't be allowed to say them" (Ross, Tr. 5766).

1312. Consumers hold beliefs about products and services (Ross, Tr. 5815). Those beliefs are measured in terms of [**320**] attributes or dimensions or characteristics of the product. Consumer attitudes are measured in terms of both the nature of the content of their beliefs and the effect or desirability of their beliefs (Ross, Tr. 5815). The most typical way of measuring consumer images or beliefs is to conduct a consumer survey, an image study, to measure the nature of those beliefs through either open-ended or close-ended responses (Ross, Tr. 5816; Crespi, Tr. 4341).

1313. Dr. Ross selected from CX 395 18 attributes from among the 33 that were asked about, attributes that he felt were most pertinent to evaluating consumer beliefs about relative therapeutic benefits of aspirin products (Ross, Tr. 5817). CX 637, "Usage and Selected Image Characteristics of Bayer, Norwich and Store Aspirin," is a table of attributes from CX 395 that Dr. Ross considered (Ross, Tr. 5819). Dr. Ross considered attributes such as "never upsets your stomach," "relieves pain most quickly," "relieves pain for a long time" as pertinent to consumer beliefs concerning therapeutic superiority (CX 637), and attributes such as "hear about it all the time," "high priced brand," or "a company that cares about the consumer," as not pertinent. This was an appropriate selection of data for purposes of the complaint. CX 521 was already composed of a question specifically dealing with effectiveness and a second one dealing with speed which Dr. Ross felt were appropriate attributes to look at for purposes of the complaint (Ross, Tr. 5816–17).

1314. Dr. Ross prepared CX 637 for the purpose of removing biases that are inherent in studies such as CX 395. This bias must be accounted for in order to arrive at a conclusion about the comparative

image of brands in a survey in which there are different numbers of brand users of the brands in the survey. Dr. Ross regarded user bias to be unfair where, as in CX 395, there were more exclusive users of Bayer than there were users of other brands in the survey. Without adjusting for the fact that there are different brand shares, an analysis of CX 395 would "load the dice," as it were, for many attributes for the brand that was most popular (Ross, Tr. 5820). Since Dr. Ross lacked the underlying data to perform what he regarded as the preferred adjustments—either to calculate the user image of Bayer and contrast it with the user image of Norwich or store, or to calculate the respective brand image of Bayer, Norwich and store among nonusers of each respective brand—what he did was to first determine the number of exclusive users of each brand and subtract those percentages from the percentages who rated particular attributes of the respective brands as "top pocket" (Ross, Tr. 5820–21). The term "top pocket" refers to the method of eliciting data. Subjects in CX 395 were asked in each **[321]** question to express their preferences by placing cards with brand names printed thereon in one of six envelopes or "pockets" ranged vertically on a page, at the top of which was printed a verbal anchor such as "relieves pain for a long period," and at the bottom of which was printed "relieves pain for a short time" (CX 395D, Z158–Z160).

1315. In performing these calculations Dr. Ross realized that some results would be illogical (Ross, Tr. 5821). For example, only 29% of Bayer users regarded Bayer as top pocket for the attribute "relieves pain for a long period." If one subtracts the 36% exclusive Bayer users figure from the 29% top pocket figure, the result is −7%, which is inexplicable from the viewpoint of a real number in this context (Ross, Tr. 5821). Nevertheless, despite the inadequacies of this adjustment in some cases, it was the best procedure available to Dr. Ross to remove user bias given the absence of underlying data (Ross, Tr. 5820–21). Dr. Ross followed a similar procedure for the other brands, subtracting the 2% exclusive Norwich users from the Norwich top pocket data, and subtracting 7% exclusive store brand use from the store brand top pocket data (Ross, Tr. 5823).

1316. As a last step, Dr. Ross calculated the statistical significance of differences in the adjusted top pocket ratings between Bayer, Norwich and store for all but two of the 18 attributes listed on CX 637 (two attributes, "good for relieving nervous tension" and "good for premenstrual/tens/depression," he decided were not relevant to therapeutic superiority allegations of the complaint (CX 637; Ross, Tr. 5822). He used a common statistical test known as a chi square, at a 90% confidence interval (10% alpha level). The results were reported

at CX 637 in the far right column headed "differences" (Ross, Tr. 5823).

1317. Dr. Ross found that for five of the characteristics Bayer was superior in image to both Norwich and store brand as follows: often recommended by doctors, never upsets the stomach, good for occasional mild headaches, effective in reducing fever, and good for aches and pains of colds and flu (Ross, Tr. 5824). Bayer's image was superior to store but equal to Norwich on two attributes: relieves pain most quickly and good for all kinds of pain (Ross, Tr. 5824). Bayer was inferior to either Norwich or store or both on several attributes: strength or strong product, and specialized kinds of pain such as muscular and arthritic pain (Ross, Tr. 5825).

1318. Based on his analysis of top pocket attributes on which Bayer was superior and those where it was not, Dr. Ross concluded from CX 395 that Bayer was believed by consumers to be a superior general pain reliever, but it was not seen as comparatively strong aspirin nor as a unique, distinctive or specialized pain reliever compared to other aspirin (Ross, Tr. **[322]** 5825). Dr. Ross viewed the fact that Bayer was rated higher on the "often recommended by doctors" measure than competing aspirins as relevant to the establishment component of the case (Ross, Tr. 5825).

1319. Two image studies in evidence, CX 395, the 1967 "Assets and Liabilities" Study, and CX 521, the Zeisel Image Study (conducted in 1975 and reported in 1976), were relied on by Dr. Ross in reaching an opinion about consumer beliefs about the therapeutic superiority of Bayer Aspirin to other plain 5-grain aspirin. Dr. Ross concluded from his analysis of CX 395 and 521 that a significant number of consumers believed that Bayer is therapeutically superior to other aspirin (Ross, Tr. 5816). This conclusion is reasonable and is supported by a preponderance of credible evidence in the record.

B. *The Record Evidence Is Insufficient To Show That Respondent's Unlawful Advertising Claims Played a Significant Role in Creating and Reinforcing Consumers' Beliefs in the Therapeutic Superiority of Bayer Aspirin Over Other Plain Aspirin During the Relevant Period*

1320. A variety of factors contribute to the creation or reinforcement of an image about a product. The most frequently mentioned factors in the professional and trade literature are usage, advertising, word-of-mouth, the package, news stories in the media, and the store in which it is bought (Ross, Tr. 5832). Of these, the literature in the field of marketing regards usage and advertising as the most likely source of product image (Ross, Tr. 5832–33). Word-of-mouth generally

derives from usage or advertising, so that it is not regarded as a separate source of image (Ross, Tr. 5833).

1321. Advertising creates expectations for consumers about what benefits are to be achieved or realized through the use of the advertised product (Ross, Tr. 5833). If the consumer buys the product, advertising will guide or assist the consumer in coming to some impression or conclusion about what the product is like. It will guide their perceptions as to how the product is performing (Ross, Tr. 5833). Advertising creates expectations about product performance which then translate into either causing people to try the product, or, if they are already users of the product, reinforcing images or beliefs that consumers already have about the product (Ross, Tr. 5833).

1322. If a product permits the consumers to evaluate it correctly, in the sense of sensory or physical qualities of the product, that will be the primary basis of consumer image or attitudes towards the product (Ross, Tr. 5834). Most consumers will trust their own senses as a basis of evaluating a product more than they will advertising or someone else's word for it. For most products and services, direct user experience or **[323]** perception of that experience is the basis of the image for that product (Ross, Tr. 5834).

1323. It is more difficult for consumers to evaluate the comparative performance of a product than it is to evaluate the absolute performance of a product (Ross, Tr. 5836–37). This difficulty is compounded where, as here, the product is a pain reliever and the consumer's only bases for evaluation are his own possibly uncertain recollection of past pain experiences, which may vary from time to time (Brock, Tr. 5163–65), and the effects of different analgesics, which may be indeterminate because of the subjective nature of individual response to pain and analgesia and the placebo effect.

1324. When a consumer cannot evaluate the performance of a product by virtue of his or her own sensory abilities, especially in a comparative sense, then direct usage experience will play an increasingly lesser role as sensorially discernible differences among brands in a product category became more indistinct (Ross, Tr. 5837). When the differences in performance between products in the brand category tend to be small, then the role of advertising in forming comparative images about the products tends to increase. In other words, as usage diminishes in its ability to be a contributor to the image of a product, then so does advertising increase in its role as a contributor to that image (Ross, Tr. 3838). The inability of consumers to evaluate the performance of drugs or, as here, the comparative performance of drugs, is supported by classical psychological research which shows that user "perceptions" of the drugs are influenced not by actual

product performance but by external information, such as advertising (Brock, Tr. 5054).

1325. In circumstances such as use of OTC analgesics, where usage cannot be appropriately evaluated by consumers, usage experience cannot explain away images that consumers hold about the product. Their expectations are, by definition, confirmed or supported when they use the product (Ross, Tr. 5838). Since usage and advertising are the most important sources of product image, the unimportance in fact of usage and comparative usage experience to the consumer in his or her formation of images of over-the-counter internal analgesic brands leaves the advertising induced expectations of performance as the major source for the creation and reinforcement of consumer beliefs about OTC analgesics.

1326. Dr. Ross, complaint counsel's witness, concluded that Bayer advertising disseminated between 1969 and 1974 served to either cause or to reinforce or contribute to Bayer's superior therapeutic image (Ross, Tr. 5841). The basis for this conclusion included his view that virtually all Bayer advertisements disseminated during this period represented that Bayer was therapeutically superior to other aspirin. This view is contrary to F. 293–294, *supra;* F. 1335, *infra.* Dr. Ross [324] also relied on the penetration studies (CX 553, 565) regarding what consumer recollections were present, and the "Assets & Liabilities" (CX 396) and "Zeisel image studies" (CX 521) which showed that the image of Bayer is that it is therapeutically superior to other aspirin (Ross, Tr. 5841). Moreover, in Dr. Ross' view, the image studies in evidence show that both users and nonusers hold essentially the same belief about the superior therapeutic performance attributes of Bayer.

1327. The evaluation of Dr. Ross' testimony as a whole makes clear that his conclusion that advertisements played a substantial role in creating or reinforcing the therapeutic superiority image of Bayer rests mainly upon his view that therapeutic superiority of Bayer was the dominant theme of all Bayer advertisements disseminated between 1969 and 1974. The administrative law judge, however, found that a relatively small number of the advertisements in evidence covering that period can be said to contain a therapeutic superiority claim.

1328. To the extent CX 521 (the Zeisel Image Study), conducted in 1975 and reported in 1976, noted a higher level of therapeutic superiority image for Bayer than the image data found in CX 395, reported in 1967 (and before the bulk of the advertisements in evidence were disseminated), it is arguable that the post-1967 Bayer advertisements played a significant role in raising the level of Bayer's therapeutic superiority image.

1329. The record shows, however, that CX 395 (the 1967 Study) is likely to have significantly understated Bayer's therapeutic image vis-a-vis other plain 5–grain aspirin brands (which is in issue in this proceeding). CX 395 essentially presented rankings by consumers of the then five major national OTC analgesic brands (Bayer, Alka Seltzer, Anacin, Bufferin and Excedrin) as well as Norwich aspirin and a catchall category called "Store's Own Brand." Although data for Norwich and Store's Own Brand were included for some thirty-three attributes, the in-depth analysis based on user/nonuser breakdowns was presented only for four major nationally advertized brands, Bayer, Anacin, Bufferin and Excedrin (CX 395, Z146–Z153). Thus, the study had as its thrust the image of Bayer as compared to the combination products, and the inclusion of only two possible aspirin brands undoubtedly shifted the focus of Bayer's comparative therapeutic image rankings away from aspirin brands. It is fair to say that CX 395 collected the Bayer-Norwich-Store's Own Brand data only incidentally, and as a result substantially understated Bayer's relative position vis-a-vis other plain 5–grain aspirins. In this light, the rise in Bayer's comparative image from 1967 to 1975 shown in CX 521 appears less significant.

1330. In this connection, what is noteworthy of CX 521 is that the proportion who rated all aspirin brands alike or other **[325]** aspirins better than Bayer in terms of effectiveness and speed of relief, were significantly greater than those who rated Bayer better. *See* CX 521C, Table.

1331. More importantly, the analyses of CX 521 data by complaint counsel's expert witnesses made no distinction whatsoever between Bayer, which is the only nationally advertised and distributed brand of plain 5–grain aspirin, and eight other aspirin brands, all regionally distributed brands and none of which does any advertising to speak of. Thus, complaint counsel's experts totally ignored the well recognized fact that consumers generally consider national brands or advertised brands to be superior products over local brands or unadvertised brands (Miles, Tr. 9397–98, 9598–99; Haley, Tr. 10577; Lipstein, Tr. 12028–29). It is especially difficult to compare Bayer to any of the other regional or unadvertised brands, and the user/nonuser analysis performed by Dr. Ross does not remove this difficulty.

1332. It is common knowledge that Bayer aspirin was introduced in this country during the early 1900's and has been the only plain 5–grain aspirin tablets marketed for some time. In fact, "aspirin" was a trademark identified with Bayer Aspirin for a long time. Since other aspirin brands were introduced, Bayer aspirin has remained the only nationally advertised and nationally distributed plain 5–grain aspirin. In these circumstances, one would expect consumers to be very

familiar with Bayer Aspirin independent of personal usage of the product or of personal exposure to any of the Bayer advertisements found to contain a therapeutic claim. It is fair to conclude that Bayer is *sui generis* in terms of brand longevity and familiarity. And there is no dispute that consumers generally rate familiar products or advertised products higher than others apart from use experience. And in 1975, users and nonusers alike rated Bayer higher in terms of effectiveness and speed, both of which are generic claims.

1333. Bayer Aspirin is another product with a long-standing, reliable position in the consumer's mind (perhaps even more so than Gerber's Baby Food) (Ross, Tr. 5880–81). Dr. Ross, however, totally ignored the familiarity or longevity factor in his evaluation of brand image data regarding plain 5–grain aspirins. *See also* F. 1237, *supra.*

1334. The record evidence is consistent with the view that, because of its unique longevity and brand familiarity, Bayer has always enjoyed a fairly high level of favorable product image (particularly with respect to such generic attributes as efficacy and safety) among users and nonusers alike and equally among those who have been exposed to or remembered any *unlawful* advertisements and those who have not. [**326**]

1335. In fact, of some fifty-two Bayer Aspirin advertisements found to be offensive, those containing the "faster acting" or "gentler" type of therapeutic claims numbered a scant dozen. The remainder contained claims of "best" or "world's best" aspirin, both bordering on puffery. *See* F. 293–294, 306–07, *supra.*

1336. There is also some evidence that the relatively small number of Bayer advertisements found to contain therapeutic claims did poorly in terms of related recall scores in copy tests, meaning that these advertisements were not effectively conveying the advertising messages to the audience. *See* RPF 5.204–5.212.

1337. From all of the foregoing, it is found that the record evidence is insufficient to show that respondent's unlawful advertisements played any significant role in creating or reinforcing a consumer image of therapeutic superiority for Bayer to the extent such image is found to exist in this record.

C. *The Record Does Not Show A Need For Corrective Advertising Regarding Bayer Aspirin*

1338. As discussed in B hereinabove, the record shows that in 1975 a substantial number of consumers had an image of Bayer Aspirin as being superior to other aspirins in terms of effectiveness and speed. However, the record evidence is insufficient to show that Bayer's superiority image, to the extent it existed in 1975, was attributable in any significant respect to respondent's advertising claims of

Bayer's therapeutic superiority. Furthermore, in view of the relatively small number of advertisements containing therapeutic superiority claims, disseminated during a relatively short period of time almost a decade ago, an inference that such image, to the extent, if any, attributable to the offending Bayer advertising, is likely to persist into the 1980's in the absence of such advertising since the middle 70's is not reasonable. *See* F. 1335.

1339. Recent research regarding corrective advertising has shown that the fashioning of a corrective message is a very difficult task. A message which appears acceptable can sometimes convey to consumers meaning beyond or outside what was intended by the author (*See* RPF 6.318–6.320). In this case, the problem is compounded by the need not to inhibit dissemination of advertising information containing true and unmisleading claims of product quality, including comparative claims where appropriate. Since the record evidence is insufficient to support an inference that there is an image [**327**] attributable to unlawful advertising claims with respect to Bayer, the justification for any corrective advertising involving Bayer is lacking in this record.

1340. Furthermore, the effect of a corrective advertising order regarding Bayer Aspirin may be to injure all 5–grain aspirins on the market. Bayer Aspirin is the only product which advertises for aspirin *per se* and which defends aspirin as a generic category against the antiaspirin advertising of the combination and acetaminophen products. Bayer is identified with the generic class of straight aspirin. A corrective statement or a disclaimer statement would weaken any Bayer advertising and would be likely to injure all straight aspirin products (Miles, Tr. 9464).

1341. Corrective advertising would be likely to be punitive as against Bayer Aspirin, in that its effect would not be limited to correcting an alleged incorrect belief but would injure the product's image and position generally, including product attributes the correctness of which is not questioned (Miles, Tr. 9436).

1342. Corrective advertising for Bayer Aspirin would be harmful because the basic image of Bayer Aspirin is that of old fashioned, reliable high-quality, which rests in substantial part on a stock of goodwill. As the "grandfather brand" it has been building this goodwill for more than 50 years. Once lost, Bayer's goodwill would be practically impossible to retrieve (Miles, Tr. 9463).

X. RESPONDENT LOIS HOLLAND CALLAWAY IS LIABLE FOR ITS CREATING AND DISSEMINATING CERTAIN CHALLENGED ADVERTISING CLAIMS

1343. Respondent Lois Holland Callaway, Inc. ("LHC") actively participated in the creation and dissemination of all challenged Van-

quish advertisements disseminated after April 1971. That participation included development of Vanquish advertising copy strategies, and development jointly with Sterling's Glenbrook Laboratories Division of Vanquish marketing plans, beginning with the 1972 marketing plan (CX 678, admission 225; CX 681, admission 268).

1344. LHC played a substantial role as Sterling's advertising agency for the development of the following advertisements for Vanquish between April 1971 and 1974: CX 252–256 and 258–264 (CX 632A, B). These advertisements were disseminated from May 1971 to December 1974 (CX 632A, B), and made all representations as alleged in Complaint Paragraph 8(B)(2), 8(C), 12(B)(1), 12(B)(2), 12(C) and failed to disclose the presence of aspirin as alleged in Complaint Paragraph 23. **[328]**

1345. Through CX 252–256 and 258–264 LHC represented that a recommended dose of Vanquish is more effective for the relief of pain than a recommended dose of aspirin or buffered aspirin, and that this comparative superiority is established. Through CX 252, 253, 255, 256, 258–264 LHC represented that a recommended dose of Vanquish is more effective for relief of pain than the largest "extra strength" tablet, and that this comparative superiority is established. Through CX 252–256 and 258–264 LHC represented that because Vanquish contains "gentle buffers" it will result in less gastric discomfort than any internal analgesic not containing buffers, and that this comparative superiority is established. Throughout CX 252–256 and 258–264 LHC failed to disclose that Vanquish contains aspirin.

1346. With respect to comparative efficacy and safety claims regarding Vanquish, the record indicates that there were some literature and research support for those propositions at the time these claims were disseminated, although they had not been convincingly demonstrated by well-controlled clinical studies. In these circumstances, it was reasonable for LHC to have relied on its client's scientific judgment in favor of these claims, based in part on Sterling's in-house research data. The record does not show that LHC in fact knew of any cogent scientific evidence to contradict its client's scientific judgment in this regard. The view that Section 5 requires an advertising agency to conduct its own study or to obtain independent scientific opinion in order to verify the scientific validity of proposed advertising claims is rejected.

DISCUSSION

A. *Introduction*

The instant proceeding (D. 8919) is one of the three related OTC internal analgesic advertising cases instituted by the FTC in February 1973 under Sections 5 and 12 of the FTC Act. An Initial Decision has been filed in each of the other two cases, *i.e., Bristol-Myers Co., et al.,* (D. 8917), September 28, 1979 [102 F.T.C. 21 (1983)]; *American Home Products Corp., et al.,* (D. 8918), September 1, 1978 [98 F.T.C. 136 (1981)]. D. 8917 (*Bristol-Myers*) involved certain advertising claims for Bufferin (a buffered aspirin product), Excedrin (a combination aspirin product also containing acetaminophen, salicylamide and caffeine) and Excedrin P.M. (a combination aspirin product also containing acetaminophen, salicylamide and methapyrilene fumarate). D. 8918 (*American Home Products*) involved certain advertising claims for Anacin (an aspirin-caffeine combination product) and Arthritis Pain Formula (a buffered aspirin product). D. 8917 (*Sterling Drug*) involves Bayer Aspirin (plain 5-grain aspirin tablets), Bayer Children's Aspirin (plain aspirin tablets for children), Cope (a buffered **[329]** combination aspirin product also containing caffeine and methapyrilene fumarate), Vanquish (a buffered aspirin product also containing caffeine and acetaminophen) and Midol (a combination aspirin product also containing caffeine and cinnamedrine HCL). Joint hearings in the three cases were held in June, July and August of 1977, followed by further separate hearings in each of the three cases.

Although the three advertising cases involved various OTC analgesic products of different formulations, they have certain common core issues of law and fact, namely, the appropriate legal standards governing the advertising claims of simple or comparative efficacy and/or safety found to have been made for the various OTC analgesic products and the adequacy of medical-scientific evidence the advertiser relied on as substantiation of its advertising claims.

In this Initial Decision, the ALJ has attempted to follow the same legal standards articulated in the earlier cases in light of the evidence contained in this record and endeavored to fashion self-explanatory findings. Therefore, the discussion which follows will be limited to certain key issues which are unique to this case. Among such issues are (1) comparative advertising claims of drug product quality (or pharmaceutical quality) as distinguished from implied efficacy or safety claims, (2) the reasonable basis required for a comparative pharmaceutical quality claim, (3) physicochemical evidence and blood level data as bases for superior therapeutic claims regarding plain

5–grain aspirins, (4) Section 5 liability for the so-called inconsistent claims, and (5) the propriety of corrective advertising requirement with respect to Bayer Aspirin.

B. *Consumers Recognize Pharmaceutical Quality As A Distinct Attribute Apart From Therapeutic Performance Of Drug Products Although The Two Are Ultimately Related*

Complaint counsel argue that an express claim that Bayer Aspirin has superior pharmaceutical quality over other USP aspirins in certain respects is an implied comparative therapeutic claim because consumers will perceive such a claim to mean that Bayer Aspirin's therapeutic performance is significantly superior to that of other aspirin brands. Thus, complaint counsel lump drug product quality claims (or pharmaceutical claims) and therapeutic claims (or medical claims) together and refuse to distinguish between an advertising claim that Bayer Aspirin is "faster" or "safer" than other aspirins and a claim that Bayer Aspirin is "a better **[330]** quality aspirin one can count on," or "a product that will do what aspirin tablets are supposed to do."[7]

Complaint counsel's argument is deficient in several important respects. *First,* it is clear from common sense and daily experience that drug product quality or pharmaceutical quality is a familiar concept which is readily recognized and understood as such by consumers apart from drug performance or efficacy. It is beyond dispute that consumers understand and desire drug product quality not necessarily because they believe that drug quality can affect the performance of a drug but primarily because they want quality, purity or freshness in a drug product for its own sake.[8] The fact that drug product quality can ultimately affect the therapeutic performance is not a good reason to ignore, in the guise of consumer protection, an important concept consumers recognize and on which they base their purchasing decisions in their daily lives.

Secondly, the inevitable and regretable consequence of complaint counsel's position will be to inhibit free and unfettered dissemination of true and nonmisleading information regarding significant drug quality improvements. Such a position would not only run counter to the Commission's established policy of encouraging free flow of important product information (*e.g., The Eyeglasses Industry Rulemaking Proceeding; The American Medical Association* (D. 9064)) but also may

[7] The ALJ recognizes that certain claims are comparative therapeutic claims even though they may be couched in terms of physicochemical or pharmaceutical terms. For example, a claim of superior dissolution speed is a comparative therapeutic claim because the speed of aspirin tablet dissolution can have no independent meaning apart from speed of pain relief action to consumers.

[8] *See, generally,* Professor Boulding's thoughtful presidential lecture at the American Association for Advancement of Science annual meeting in January 1980. Kenneth E. Boulding, "Science: Our Common Heritage," *Science,* 207:831–836 (1980).

have such a chilling effect as to constitute an unreasonable prior restraint on legitimate commercial speech protected by the First Amendment.[9] **[331]**

Finally, since there is no dispute that public policy should encourage improvement of the pharmaceutical quality of drug products, complaint counsel's technical interpretation of drug quality advertising claims would reduce incentives for drug quality improvements and thus be counterproductive while having no significant redeeming features.[10]

C. *An Affirmative Product Claim Must Be Based On A Reasonable Basis And With Respect To A Drug Product Quality Claim Based On Physicochemical Studies, Such Studies Must Show Statistically And Clinically Significant Difference*

It is now well-established that an affirmative product claim must be based on a reasonable basis at the time such a claim is made and that certain advertising claims must be adequately supported by appropriate scientific evidence. *Pfizer, Inc.,* 81 F.T.C. 23 (1972); *Firestone Tire & Rubber Co.,* 81 F.T.C. 398 (1972), *aff'd,* 481 F.2d 246 (6th Cir. 1973), *cert. denied,* 414 U.S. 1122 (1973); *National Dynamics,* 82 F.T.C. 398 (1972), *aff'd,* 492 F.2d 1333 (2d Cir. 1973), *cert. denied,* 419 U.S. 993 (1974). **[332]**

A claim that Bayer Aspirin is superior in quality to other aspirin brands implies that the claim is supported by appropriate scientific evidence. Sterling's express reliance on scientific tests in its Blue Book advertising campaign shows Sterling's acceptance of that substantiation requirement. Indeed, in this proceeding Sterling's first-line defense with respect to the challenged Bayer advertising claims is that there was adequate scientific substantiation for the pharmaceutical claims it made. Reason and common sense require that, to the extent aspirin quality claims are based on physicochemical comparisons, such studies be of a sound design consonant with recog-

[9] *Va. State Board of Pharmacy* v. *Va. Citizens Consumer Counsel, Inc.,* 425 U.S. 748 (1976); *Bates* v. *State Bar of Ariz.,* 433 U.S. 350 (1977); *Central Hudson Gas & Elec. Corp.* v. *Public Serv. Commission,* 100 S.Ct. 2343 (1980). *Also see* Tribe, *American Constitutional Law,* at 651–56, 712–14, 721–24, 728–30 (1978).

[10] The ALJ recognizes the close questions of policy involving certain competing considerations which demand a careful deliberation regarding Bayer's superior quality claims. *First,* Section 5 of the FTC Act mandates the Commission to prevent deceptive, or confusing or spurious therapeutic superiority claims based on insignificant physicochemical differences which are capable of misleading consumers. *Second,* it is an important, recognized public policy objective to promote the improvement of drug product quality, and a requirement that every pharmaceutical claim which may be said to convey a therapeutic message to some consumers be supported by well-controlled clinical trials may run counter to the policy of encouraging all improvements in drug product quality independent of their therapeutic importance. *Third,* a legal requirement that every comparative pharmaceutical quality claim be substantiated by well-controlled clinical trials may have a chilling effect upon true and honest claims of product quality, and be tantamount to unlawful prior restraint on commercial speech. Finally, it may run counter to the Commission's established policy of encouraging free flow of significant product information.

After due deliberation, the ALJ is of the opinion that the resolution reached herein is reasonable and realistic in light of the record as a whole.

nized scientific design principles and the results show statistically significant differences. This is essential in order to insure that the results of such studies do reflect a true difference and are not a result of chance. Furthermore, with respect to a drug quality claim, mere statistical significance in terms of some physicochemical characteristics could be meaningless unless it is also clinically significant, not in the sense that the superiority of one aspirin has been demonstrated through clinical studies but in the sense that the difference observed can reasonably be expected to (or is known to) have a significant clinical impact in the opinion of biomedical experts.

On the other hand, if advertisers were allowed to claim superior drug quality for their products without adequate medical-scientific substantiation outlined above, consumers will be hopelessly confused and misled by claims of differences which are, in reality, illusory or meaningless. The result would not only be unfair to consumers but also to other competitors who do not make drug quality claims unless they do have appropriate medical-scientific substantiation.

D. *The Record Evidence Regarding The Physicochemical Characteristics Of Bayer Is Insufficient To Substantiate Therapeutic Superiority Claims Of Bayer*

During the trial, Sterling advanced a position which would apply different standards of substantiation to comparative therapeutic claims for buffered or combination aspirin products, on the one hand, and to similar claims for plain 5-grain aspirin, on the other hand. With respect to the former, Sterling would require well-controlled clinical studies. As to the latter, Sterling argues that the record evidence regarding physicochemical characteristics of Bayer and other brands constitutes adequate substantiation of comparative therapeutic claims for Bayer. This argument is rejected for several reasons.

First, the record as a whole is persuasive that a comparative therapeutic claim of one brand of plain 5-grain aspirin over another USP 5-grain aspirin based solely on [333] physicochemical differences remains an hypothesis to be clinically tested although the hypothesis does appear rational and plausible in terms of recognized pharmaceutical and pharmacological principles. Until so tested and confirmed, the superiority claim stands unsubstantiated.

The record is clear that since the early 1960's there has been little dispute in the biomedical scientific community that the efficacy of drugs must be demonstrated by well-controlled clinical studies, including appropriate replication. This basic principle was applied equally to new therapeutic agents and to new formulations of drugs recognized as effective. The record shows that until this trial Sterling has subscribed to this view in various representations made to the

FTC and the FDA. In recent years, a vocal dissent from that position has been heard. There are respected scientists who sincerely believe that the strict FDA requirements regarding clinical trials have exacted excessive costs in terms of delayed introduction of important new drugs as well as in terms of research and economic resources. They urge that other less costly alternatives short of well-controlled clinical studies should be accepted. However, the record shows that this remains a minority view in this country.

In any event, reason and common sense argue that the need for clinical demonstration becomes more acute when the claim is not of simple efficacy but is that one brand of plain 5–grain aspirin is therapeutically superior to other USP aspirins. There appears to be a paucity of reports of clinical studies comparing different brands of plain 5–grain aspirin. A possible explanation of this fact may be that biomedical scientists believe, as a basic proposition, that generic equivalents (such as different brands of plain 5–grain aspirin) are therapeutic equivalents until the contrary is shown to be the case. It is reasonable and fair that those who claim therapeutic superiority of one brand of drug product over another generically equivalent product demonstrate the therapeutic superiority of their product through well-controlled clinical trials. Until this has been done, such therapeutic superiority claims remain unsubstantiated, and physicochemical data alone are insufficient to fill the fundamental gap.

In defense of Bayer's therapeutic superiority, Sterling placed a heavy reliance upon research literature related to biopharmaceutical and dissolution rate-time characteristics of drug products, and vigorously argued that pharmaceutical equivalents are not therapeutic equivalents. However, the record is clear that the only time bioavailability can significantly affect aspirin's therapeutic performance is when aspirin is administered in chronically high-level maintenance doses for the treatment of rheumatoid arthritis or rheumatic fever, an area found to be inappropriate for self-medication by the FDA-OTC Internal Analgesics Panel (CX 466). The research **[334]** and review articles Sterling relies on make clear that the "truth of the matter is that although drug formulations has been studied extensively *in vitro,* the clinical significance and the extent of generic inequivalence is unknown." "[D]rug activity is not necessarily related to drug concentrations in plasma. In addition, the plasma concentration of a drug depends not only on absorption, but also on individual characteristics and kinetics of drug distribution, methabolism and excretion" (RX 250-Prescott, at 287, 289). The same author concluded:

> The incidence and ultimate clinical significance of generic inequivalence is unknown .... [I]t is clear that complex *in vitro* studies of drug formulations cannot be relied on

to predict the performance of drugs in clinical practice. . . . There is an appalling lack of information on the equivalence of drugs in the very situation where it is most needed—in patients with diseases.

In 1972, in concluding a 40–page review article of tableting research and technology, a recognized pharmaceutical expert cautioned that "[i]n spite of increasing activity at the biological level [publication of research papers], the investigational gap in *in vitro - in vivo* correlations involving the tablet dosage form remains too wide" (RX 250-Cooper, at 1531, 1550). The record shows that these observations still hold true today. The record is convincing that evidence other than well-controlled clinical studies are insufficient to provide a reasonable basis for a comparative therapeutic proposition regarding plain 5–grain aspirin brands, which are generic equivalents.

Sterling's argument that conducting costly, large-scale clinical trials of different brands of plain 5–grain aspirin are fraught with many difficulties and that this is not an optimal utilization of biomedical research resources has considerable force. In the final analysis, however, there are no insurmountable ethical or logistic barriers to conducting well-controlled clinical studies of a relatively small number of leading brands of plain 5–grain aspirin. If Sterling insists on advertising therapeutic superiority claims for Bayer, Sterling should conduct the required clinical studies. On the other hand, it may well be that, because of the present state of art in analgesometry, whatever therapeutic differences that may exist among different brands of plain 5–grain aspirin may not be large enough to be observed or to reach statistical and clinical significance. Should this be the case, then Sterling would be put in no different position than its competitors with regard to their comparative therapeutic claims for OTC analgesic products. Furthermore, there is nothing to suggest that consumers make a distinction in terms of the kind and degree of scientific substantiation they expect with respect to therapeutic superiority claims for combination products, on the [335] one hand, and similar claims for plain 5–grain aspirin, on the other hand. For plain aspirins and combination products alike, a therapeutic superiority claim stands unsubstantiated until it is demonstrated by well-controlled clinical studies.

Sterling's argument that the FDA instructed the various OTC drug monograph panels, including the OTC Internal Analgesics Panel, that they may base conclusions regarding the effectiveness of OTC drugs under review solely upon data other than well-controlled clinical trials is not persuasive (RPF 7.879–7.891). In the administrative law judge's view, the most that can be inferred from the information Sterling relies on in this regard, is that (1) the FDA intended the OTC

drug monograph panels to consider all available scientific information, including data other than well-controlled clinical studies, and (2) the FDA intended to permit the panels to reach conclusions regarding the effectiveness of OTC drug ingredients under review, even in the absence of well-controlled clinical studies, on the basis of the available data. This is a far cry from saying that the FDA no longer requires well-controlled clinical trials in support of drug efficacy or that the FDA is willing to settle for something less than controlled clinical demonstration in all cases. Suffice it to say the record is devoid of any evidence to show that the FDA has approved an NDA involving any analgesic agent without the required clinical demonstration.

Second, in any event, the inference that can be drawn from physicochemical data regarding aspirin tablets is often a matter of degree and cannot provide a clearcut or definitive conclusion regarding the relative therapeutic performance of different brands being compared. There is no dispute that some physicochemical characteristics of aspirin tablets (such as dissolution profile) are expected to have a greater bearing on the tablet's therapeutic performance than other characteristics, or that some of the desirable physicochemical factors are mutually antagonistic in the sense that one may be enhanced at the expense of some others. Therefore, even in cases where statistically significant differences in some physicochemical characteristics are shown, the final, all important question of whether such differences in themselves are sufficient to make a significant therapeutic impact in actual use can only be resolved by well-controlled clinical tests. The assertion that, other things being equal, an aspirin brand which is better than another brand in terms of one or more physicochemical factors is preferable and that this "clinical" judgment requires no well-controlled clinical trials merely begs the question.

Finally, the various physicochemical evidence and non-clinical *in vivo* data (such as blood level data) Sterling relies on are equivocal or inconclusive. The various physicochemical studies suffer from significant deficiencies in design, execution and/or analysis, or fail to show statistical significance. Also, blood level data do not provide a reliable **[336]** answer to the question of comparative efficacy of aspirins because, as Sterling agrees, a precise correlation between the blood salicylate level and either the onset, duration or intensity of analgesia is yet to be established.

E. *Section 5 Liability Based On Inconsistent Advertising Claims Is Not Only Vague But May Constitute Unlawful Prior Restraint Upon Commercial Speech*

Complaint counsel argue that Sterling's contemporaneous dissemination of mutually inconsistent advertising claims regarding differ-

ent analgesic products it marketed is a violation of Section 5 not only because mutually inconsistent or conflicting claims regarding these products cannot be true at the same time, but also because they are unfair to consumers. This is a novel theory. Although this theory has some surface plausibility, it raises serious constitutional problems as it has been presented in this case.

First, the standard of consistent claims is not clearly articulated, nor is its content defined with sufficient clarity. Vague legal requirements accompanied by sanction are not consistent with due process.[11] *Secondly,* what is an advertiser to do when advertising claims for different products have equally reasonable basis although some of the claims may arguably not be consistent with some others? Should the advertiser forego some advertising claims having reasonable basis or make the claims at its peril? One thing is clear in these circumstances. The "consistency" requirement would have a chilling effect and may amount to an unreasonable restraint upon legitimate commercial speech in contravention of the First Amendment. For all of these reasons, complaint counsel's "inconsistent claims" theory is rejected.

F. *The Evidence Is Insufficient To Support A Corrective Advertising Requirement With Respect To Bayer Aspirin*

The basic rationale of the corrective advertising requirement in Section 5 cases is that because of the intensity and duration of the dissemination of unlawful advertising claims it may be reasonably inferred (1) that the offending advertising claims played a significant role in creating or reinforcing a mistaken product image and (2) that in the absence of a corrective advertising the mistaken product image will endure for a significant period of time. In my view, the record [337] evidence is insufficient to support either of the two necessary elements outlined above.

First, as detailed in F. 293–94, 314, 1335, *supra,* I have found that the number of offending Bayer Aspirin advertising in evidence is relatively small.[12] According to CX 630, these offending TV ads were run intermittently during a relatively short period of time, at some period between January 1969 and March 1973. Furthermore, although these advertisements were found to have implied a therapeutic claim, they were not like the more blatant comparative efficacy claims that Sterling made for Vanquish, for example, or those comparative claims made by some of Sterling's competitors.

Secondly, the record evidence is consistent with the view that Bay-

[11] *See* Tribe, n. 9, *supra,* at 718–720.

[12] Of some fifty odd Bayer Aspirin advertisements found to be offensive, those containing the more familiar "faster acting" or "gentler" claim numbered a scant dozen during the 1967–1973 period. The remaining ads were found offensive because they contained claims of "best" or "world's best" aspirin, both bordering on puffery. F. 293–294, *supra.*

er's relatively high therapeutic image is due to Bayer's unique longevity and brand familiarity among consumers. F. 1320–34, *supra.* In these circumstances, it is not reasonable to infer that Sterling's offending advertising played a significant role in creating or reinforcing Bayer's superiority image and to require Bayer to include a corrective advertising message in all future Bayer advertisements.

In sum, the administrative law judge is persuaded that the record evidence does not support a corrective advertising remedy with respect to Bayer. Corrective advertising is an equitable remedy and should be required only when there is convincing evidence showing its need. In the administrative law judge's opinion, this is not such a case.[13]

G. *Liability Of Lois Holland Callaway*

The law is well-settled that an advertising agency may be held liable for false advertising if it "actually participated in the deception ... In order to be held a participant in such [**338**] deception, the agency must know or have reason to know of the falsity of the advertising." *Doherty, Clifford, Steers and Shenfield, Inc.* v. *FTC,* 392 F.2d 921, 918 (6th Cir. 1968); also *Carter Products, Inc.* v. *FTC,* 323 F.2d 523, 534 (5th Cir. 1963); *ITT Continental Baking Co., Inc.,* 83 F.T.C. 865 (1973).

In determining liability, the agency will be strictly held to know what claims are made in advertisements. *In re Merck & Co.,* 69 F.T.C. 526, 559 (1966), *aff'd,* 392 F.2d 921 (6th Cir. 1968). *ITT Continental, supra.* Since LHC actively participated in the creation and dissemination of the challenged advertisements for Vanquish, the remaining issues regarding its liability is whether it knew or should have known that the advertisements were false due to failure to disclose material facts of the presence of aspirin and the existence of a substantial question in the medical scientific community concerning the validity of the "establishment" claims regarding Vanquish.

Complaint counsel argue that respondent's absolute and comparative efficacy (and related) claims for Vanquish were false because, having represented these claims as being "established" by scientific evidence, LHC knew or should have known that the data supporting the claims were subject to "substantial question" among experts and that the existence of such substantial question was a material fact which should have been disclosed to consumers. Complaint counsel also argue that the failure to disclose the presence of aspirin in Vanquish was false because LHC knew, or should have known, that since

[13] In *Warner-Lambert,* the cold-preventive image of Listerine was shown to be about three times as high as that of competitive products. 86 F.T.C. at 1503. Also in this connection, Professor Emerson's reminder bears repeating today in the context of this case. We should be ever mindful of the danger of imposing restrictive rules which may be valid in principle but may tend in actual operation to circumscribe freedom of expression. Emerson, "Toward A General Theory of the First Amendment," 72 Yale L.J. 877, 901–902 (1963).

aspirin may cause undesirable side effects in certain users, implicit promotion of these analgesics as containing ingredients other than aspirin and failure to disclose the presence of aspirin was false advertising by virtue of the fact that the presence of aspirin is material fact knowledge of which may cause some consumers to change their purchase decisions.

It is my determination that the record as a whole supports the conclusion that LHC's good faith reliance on Sterling's substantiation information with respect to the comparative efficacy claims for Vanquish was reasonable under the circumstances.

With respect to advertising agency's liability under the establishment/substantial question theory, it is my determination that the same standards applicable to drug manufacturing firms are not appropriate for advertising agencies. Here, as in my Initial Decision in *American Home Products,* Docket No. 8918, dated September 1, 1978 (p. 225) [98 F.T.C. at 340 (1981)], LHC is found to have acted reasonably in relying in good faith on the substantiation data provided by Sterling. As the record in this case amply demonstrates, scientific analysis or verification of the accuracy of clinical data is a highly complex, technical process, one for which LHC is not, and may **[339]** not reasonably be expected to be, equipped. Even where complaint counsel have shown the advertising agency to have been aware of some questions concerning the validity of its unqualified representations, LHC was not obligated to perform statistical or clinical analyses of the representations to determine the "substantiality" of the question or its "materiality." I reiterate my conclusions in *American Home Products:*

> This is not a case where the disparity between the advertising representations and the substantiation information is so great as to preclude a conclusion that the advertisements were conceived through reasonable reliance on the assurances of the manufacturer that the claim is true or has a reasonable basis. *Cf. Standard Oil Co. of California,* 84 F.T.C. 1401, 1474–75 (1974). Clyne [advertising agency] cannot be reasonably charged with the duty to conduct an independent investigation that the claim is scientifically established in the sense that there existed two or more well-controlled clinical demonstrations in support of the claim. In these circumstances. Clyne's good faith reliance on American Home's assurances, as embodied in CX 304, was reasonable.

H. *Relief*

It is well-established that in Section 5 cases the Commission has the power and duty to fashion appropriate remedies which are reasonably calculated to prohibit the unlawful practices found to exist. *E.g., Jacob Siegel Co.* v. *FTC,* 327 U.S. 608, 611–13 (1946); *FTC* v. *Ruberoid Co.,* 343 U.S. 470, 473 (1952); *FTC* v. *National Lead Co.,* 352 U.S. 419, 428–30 (1957). The remedy must have a reasonable relationship to the

unlawful practice and be no broader than is reasonably necessary to remedy the violation. *Jacob Siegel Co.* v. *FTC, supra,* at 613; *Beneficial Corp.* v. *FTC,* 542 F.2d 611, 619–20 (3d Cir. 1976). *See also Warner-Lambert Co.* v. *FTC,* 562 F.2d 749, 757–58 (D.C. Cir. 1977); *National Commission on Egg Nutrition* v. *FTC,* 570 F.2d 157, 164 (7th Cir. 1977).

1. Part I Of The Order

Part I of the Order would prohibit simple and noncomparative efficacy or safety claims that are not supported by a reasonable basis. The provision is justified by Sterling's failure to have a reasonable basis for various tension and depression relief claims for Bayer, Cope and Midol and by Sterling's failure to possess and rely on a reasonable basis for its superiority claims for Bayer and BCA, Vanquish and Cope. [340] Reasonable basis is defined to be "competent and reliable scientific evidence" for both simple and comparative effectiveness claims; however, as to comparative (therapeutic superiority) claims regarding OTC analgesic products, reasonable basis is further specified to be the well-controlled clinical evidence described in Part II(E). This further explication of reasonable basis standards for comparative claims involving a particular product class is based upon the extensive record evidence on the kind of data necessary to provide reasonable scientific support for such claims.

Inclusion of all OTC drug products in the reasonable basis requirement provision is appropriate in this case. Sterling appears to have been involved in a number of Section 5 proceedings which resulted in cease and desist orders or consent orders involving misrepresentation of a number of OTC drug products.[14] It is now time to place Sterling under a broad proscription with respect to all OTC drug products marketed by it. Furthermore, the proscription here is narrow and related to the particular type of claims involved in this case.

2. Part II Of The Order

Part II(A) prohibits any claim that an ingredient or combination of ingredients is unusual, special or exclusive when that ingredient or combination is available in other [341] analgesics. This is based upon

[14] Sterling has five outstanding advertising orders against it, four by consent and one after litigation. In a 1950 litigated order, Sterling was ordered to cease representing that Phillips Milk of Magnesia Skin Cream and Cleansing Cream was effective for keeping the skin free of enlarged pores and would control oiliness and from misrepresenting the benefits of these products. Sterling agreed not to represent Camphophenique, an antiseptic, as an effective cure for pimples, acne, skin rashes or as effective treatment for insect bites. 49 F.T.C. 1635 (1953). In a 1962 consent agreement, Sterling agreed to cease representing that Isuprel or similar drugs had no adverse side effects and could be taken without risk of toxic side effects. 61 F.T.C. 1008 (1962). In a 1968 consent agreement, Sterling agreed not to misrepresent the benefits of Ironized Yeast Tablets as a remedy for weakness, tiredness, frequent headaches, nervousness, loss of appetite, loss of energy or restlessness. 73 F.T.C. 979 (1968). Most recently, in a 1974 consent agreement, Sterling agreed not to represent that Lysol or any other household disinfectant will be of benefit in reducing the incidence or spread of influenza or throat infections and not to overstate the value of such disinfectants against strep and staph infections. 84 F.T.C. 547 (1974).

respondent's unfair and deceptive claims of the uniqueness of Cope's formula (Complaint Paragraph 22).

Part II(B) prohibits respondent from misrepresenting the identity of commonly known ingredients in its advertising, as the record shows they have done here, by falsely representing that these ingredients were something other than commonly known aspirin and caffeine (Complaint Paragraph 26).

Part II(C) requires respondent to disclose the presence of aspirin in every OTC drug product it advertises. The provision is based upon the record evidence which demonstrates that the presence of aspirin is a material fact which, if known to consumers, might influence their decision to purchase the drug. The provision is also justified by respondents' uniform, continuous advertising representations that have the tendency and capacity to lead consumers to believe that aspirin is not an ingredient in Midol (Complaint Paragraphs 24–25).

Part II(D) of the Order prohibits respondent from misrepresenting the results or analysis of any test, study or survey. The provision is based upon the misrepresentation of the results of the tests for Bayer and Cope challenged in Paragraphs 18 through 21 of the Complaint. The additional coverage extending to studies and surveys is justified because the technique abused in the representations challenged here is equally applicable to any study or survey.

Part II(E) of the Order prohibits representations that comparative effectiveness or comparative freedom from side effects of any OTC internal analgesic product has been established unless such is, in fact, the case. The requirements which must be met before an "established" claim can be made are based primarily on FDA's regulations, which set forth the criteria for "adequate and well-controlled" clinical investigations necessary to provide "substantial evidence" of effectiveness for new drugs (21 C.F.R. 311.111(a)(5)(ii)), and which have also been applied to OTC drugs (21 C.F.R. 330.10(a)(4)(ii)). *See* F. 449, *supra;* CB at 127–34. The FDA regulations have been modified for purposes of this Order in certain limited respects in light of the fact (1) comparative efficacy and comparative freedom from side effects are involved in this section of the Order and (2) only OTC internal analgesic drugs are involved.

Among the modifications are the following:

Part II(E) of the Order requires "*two or more* adequate and well-controlled clinical investigations conducted by *independent* experts . . ." The underlined portions have been added to the requirements of the FDA regulation to make it explicit that at least two studies must be conducted, and that the studies should [342] be done by different researchers. The language is virtually identical to the FDA Analgesic

Panel's conclusion that a Category III compound can achieve Category I status only on the basis of "at least two studies by independent investigators which conform to the guidelines [for well-controlled studies]" (CX 466 at 35445).

Since both comparative efficacy and comparative freedom from side effects are addressed in Part II(E), the FDA requirements have also been modified to reflect that fact. For example, the order provision requires that experts must conclude on the basis of the clinical studies that "the drug will have the comparative effectiveness or comparative freedom from side effects it is represented to have . . ." In addition, the Order contains a requirement, not found in the FDA regulation, that the "comparative effectiveness or comparative freedom from side effects [be] demonstrated by methods of statistical significance, and with levels of confidence, that are generally required by . . . experts." Such a requirement is necessary in light of the expert testimony in this record which demonstrated the need for statistically significant differences between drugs before any firm conclusions could be reached concerning comparative efficacy or side effects.

Part II(E), unlike the FDA regulation, requires that:

> [a]t least one of the . . . investigations to evaluate the comparative effectiveness of the drug shall be conducted on any disease or condition referred to, directly or by implication [in advertising]; or, if no specific disease is referred to, then the . . . investigations shall be conducted on at least two conditions or diseases for which the drug is effective.[15]

In other words, if a claim is made that one of respondent's OTC drugs is superior for a certain condition (*e.g.,* headache) to another product, at least one of the two studies *must* be on that particular condition. On the other hand, if a "general" superior efficacy claim is made, *e.g.,* "Vanquish is a more effective pain reliever than Drug X," the studies must show superiority in at least two conditions "for which [Vanquish] is **[343]** effective," such as headache pain and post-partum pain. Part II(E) is designed to insure that the covered superiority claims are not made with respect to a type of pain or condition for which superiority has not been demonstrated. Likewise, it is designed to preclude general claims of superiority based on studies conducted on conditions for which OTC drugs are not generally used.

Finally, Part II(E) of the Proposed Order requires that the studies be double-blind and placebo-controlled, even though the FDA regulation does not contain an explicit requirement for such controls. The regulation does indicate, however, that a placebo should be used ex-

[15] This portion of Part II does not apply to claims relating to comparative freedom from side effects. In other words, respondent could test its drug on healthy volunteers, who have no "condition or disease," and establish that its drug causes less gastric upset.

cept in circumstances where (1) "objective measurements of effectiveness are available and placebo effect is negligible," (2) the condition treated is such that administration of a placebo would be contrary to the interest of the patient," or (3) a drug is studied on "diseases with high and predictable mortality . . ." (21 C.F.R. 314.111(a)(5)(ii)(a)(4)(i) through (iv)). None of those situations is applicable to studies involving the comparative performance of OTC drugs, and for this reason a placebo control should be required in the Order. Likewise, even though the FDA regulation does not specify that all efficacy studies must be conducted under double-blind conditions—presumably because such a condition would be impossible or unethical when certain types of drugs were studied (*e.g.,* chemotherapeutic drugs)—it did require that the study must be designed to "minimize bias on the part of the subject and observer." It is my view that comparative analgesic studies should be double-blind to the extent possible.

Part II(F) of the Order prohibits respondent from making comparative effectiveness or comparative freedom from side effects claims in the face of a substantial question unless the existence of that substantial question is disclosed. It, thus, is directly related to two of the most basic allegations of lawfulness in this case, the unfairness of making comparative claims for drugs in the face of a substantial question and the misleading nature of advertisements which fail to disclose the material fact that there exists a substantial question concerning the validity of a comparative claim. The requirements of Part II(F) do not apply unless they are "triggered" by respondent's choice to make a comparative therapeutic claim.

3. Part III Of The Order

Part III of the Order is designed to implement the requirement that any comparative drug product quality claim or pharmaceutical claim be appropriately supported by a sound scientific study conducted by experts or other qualified personnel which show statistical and clinical significance of the physicochemical differences observed. This section is necessary because Sterling made superior quality claims without [**344**] a reasonable basis as alleged in Complaint Paragraphs Twenty and Twenty-One. This section would also limit the requirements as to statistical analysis to quality factors which can be quantified by generally accepted or appropriate procedures.

4. Part IV Of The Order

Part IV of the Order is directed to LHC, the advertising agency for Vanquish, and requires LHC to refrain from making certain advertising claims or failing to disclose a material fact with respect to nonprescription internal analgesic products containing aspirin.

Conclusions of Law

1. The Federal Trade Commission has jurisdiction over the advertising of Bayer Aspirin, Bayer Children Aspirin, Midol, Cope and Vanquish under Section 5 of the Federal Trade Commission Act.

2. Each of the various charges of the Complaint has been sustained by a preponderance of credible evidence, except with respect to Complaint Paragraph Twenty-Nine insofar as it relates to Complaint Paragraph Seventeen. Respondents' use of false, misleading and deceptive representations as herein found has had and now has the capacity and tendency to mislead members of the purchasing public into the erroneous and mistaken belief that said statements and representations were and are true and into the purchase of substantial quantities of Bayer Aspirin, Bayer Children Aspirin, Midol, Cope and Vanquish by reason of this erroneous and mistaken belief. In the absence of an appropriate cease and desist order, including appropriate affirmative disclosure requirements, consumers will continue to be misled by respondents' advertising representations regarding efficacy or safety or quality of said products that such representations are supported by scientific evidence generally accepted by the scientific community as establishing such propositions or have adequate substantiation.

3. The acts and practices of respondents as herein found were and are prejudicial and injurious to the public and to respondents' competitors and constituted and now constitute unfair methods of competition and unfair and deceptive acts and practices in commerce in violation of Sections 5 and 12 of the Federal Trade Commission Act.

4. The accompanying order is necessary and appropriate for the purpose of prohibiting the continuation of the proscribed acts and remedying the injury and unfairness to the consuming public. **[345]**

Order

I.

It is ordered, That respondent Sterling Drug, Inc., a corporation, its successors and assigns, and respondent's officers, agents, representatives and employees directly or through any corporation, subsidiary, division or other device, in connection with the labeling, advertising, offering for sale, sale or distribution of any nonprescription drug in or affecting commerce, as "commerce" is defined in the Federal Trade Commission Act, do forthwith cease and desist from representing, directly or by implication, that, such product is effective, or therapeutically superior to any other drug, for any disease, symptom or condi-

tion, unless at the time such representation is made respondent possesses and relies upon a reasonable basis for such representation, which shall consist of competent and reliable scientific evidence. In the case of comparative representations regarding a nonprescription internal analgesic drug product, other than representations of pharmaceutical quality, "competent and reliable scientific evidence" shall be defined as the evidence described in Part II(E) of this Order. In case of representations of pharmaceutical quality, the provisions of Part III of this Order shall apply. For the purposes of this Order, "a representation concerning the pharmaceutical quality" shall mean any representation concerning the manufacturing processes or pharmaceutical quality (such as quality, purity, stability or product formulation) of a [346] nonprescription internal analgesic product which does not refer, directly or by implication, to the speed, onset, duration or intensity of action or to adverse effects.

II.

It is further ordered, That respondent Sterling Drug, Inc., a corporation, its successors and assigns and respondent's officers, agents, representatives and employees directly or through any corporation, subsidiary, division or other device, in connection with the labeling, advertising, offering for sale, sale or distribution of any nonprescription drug in or affecting commerce, as "commerce" is defined in the Federal Trade Commission Act, do forthwith cease and desist from:

A. Representing that such product contains any ingredient, or combination of ingredients which is unusual, special or exclusive when such ingredient, or combination of ingredients, is available in other nonprescription analgesic products.

B. Referring, directly or by implication, to aspirin, caffeine or any commonly known ingredient by any word or words without disclosing the common, or usual, name of such ingredient. [347]

C. Failing to disclose in the advertising of such nonprescription drug product the presence of aspirin when such product contains such ingredient.

D. Misrepresenting, in any manner, any test, study or survey or any or all of the results thereof.

E. Representing, directly or by implication, that a claim concerning the comparative effectiveness or comparative freedom from side effects of any internal analgesic product has been established unless such representation has been established by two or more adequate and well-controlled clinical investigations, conducted by experts qualified by training and experience to evaluate the effectiveness and

comparative effectiveness or comparative freedom from side effects of the drugs involved, on the basis of which it could fairly and responsibly be concluded by such experts (1) that the drug will have the **[348]** comparative effectiveness or comparative freedom from side effects it is represented to have, and (2) that such comparative effectiveness or comparative freedom from side effects is demonstrated by methods of statistical analysis, and with levels of confidence, that are generally recognized by such experts. At least one of the adequate and well-controlled clinical investigations to evaluate the comparative effectiveness of the drug shall be conducted on any disease or condition referred to, directly or by implication; or, if no specific disease or condition is referred to, then the adequate and well-controlled clinical investigations shall be conducted on at least two conditions or diseases for which the drug is effective. To provide the basis for the determination whether any clinical investigation is "adequate and well-controlled," the plan or protocol for the investigation and the report of the results must include the following:

1. A clear statement of the objective of the investigation. **[349]**
2. A method of selection of the subjects that:

a. Provides adequate assurance that they are suitable for the purposes of the investigation, and diagnostic criteria of the condition to be treated (if any);

b. Assigns the subjects to the test groups in such a way as to minimize bias;

c. Assures comparability in test and control groups of pertinent variables, such as age, sex, severity, or duration of disease or condition (if any), and use of drugs other than the test drugs.

3. An explanation of the methods of observation and recording of results, **[350]** including the variables measured, quantitation, assessment of any subject's response, and steps taken to minimize bias on the part of the subject and observer.

4. A comparison of the results of treatments or diagnosis with a control in such a fashion as to permit quantitative evaluation. The precise nature of the control must be stated and an explanation given of the methods used to minimize bias on the part of the observers and the analysts of the data. The investigation must be conducted double-blind, and methods of double-blinding must be documented. In addition, the investigation must contain a placebo control to permit comparison of the results of use of the test drugs with an inactive preparation designed to resemble the test drugs as far as possible.

5. A summary of the methods of analysis and an evaluation of data derived **[351]** from the study, including any appropriate statistical methods.

F. Making any representation, directly or by implication, concerning the comparative effectiveness or comparative freedom from side effects of any internal analgesic product, when there exists a substantial question, recognized by experts qualified by scientific training and experience to evaluate the efficacy and safety of such drug product, as to the validity of any such representation, unless respondent discloses the existence of such substantial question by including in the same advertisement a clear and conspicuous disclosure statement conforming to the following:

1. The disclosure statement regarding Bayer Aspirin shall state "Bayer Aspirin has not been proven to be therapeutically superior to other plain aspirins," or comprise such other statement approved by the Federal Trade Commission in advance or as respondent can demonstrate [**352**] (based on consumer surveys whose design is adequate and previously approved by the Federal Trade Commission) will convey the same message to consumers.

2. The disclosure statement regarding Bayer Children's Aspirin shall state "Bayer Children's Aspirin has not been proven to be therapeutically superior to other children's aspirins," or comprise such other statement determined and approved as set forth in 1 hereinabove.

3. The disclosure statement regarding Vanquish or Cope shall state "Vanquish [or Cope] has not been proven to be more effective [or faster or gentler] than aspirin," or comprise such other statement determined and approved as set forth in 1 hereinabove.

4. In print advertisements, the disclosure shall be displayed in type size which is at least the same size as that in which the principal portion of the text of the advertisement appears and shall be separated from the text so that it can be readily noticed. [**353**]

5. In television advertisements, the disclosure shall be presented simultaneously in both the audio and video portions. During the audio portion of the disclosure in television and radio advertisements, no other sounds, including music, shall occur. Each such disclosure shall be presented in the language principally employed in the advertisement.

III.

It is further ordered, That respondent Sterling Drug, Inc., a corporation, its successors and assigns and respondent's officers, agents, representatives and employees directly or through any corporation, subsidiary, division or other device, in connection with the labeling, advertising, offering for sale, sale or distribution of any nonprescrip-

tion drug in or affecting commerce, as "commerce" is defined in the Federal Trade Commission Act, do forthwith cease and desist from making any representation, directly or by implication, concerning the pharmaceutical quality of any nonprescription drug product manufactured or distributed by it, unless at the time such representation is made respondent possesses and relies upon a reasonable basis for such representation, which shall consist of competent and reliable scientific evidence. In the case of [**354**] such comparative representations, "competent and reliable scientific evidence" shall mean a pharmaceutical or physicochemical study or survey designed and conducted according to sound scientific procedures, by experts qualified by training and experience to evaluate the pharmaceutical quality or comparative pharmaceutical quality of the drug product class, on the basis of which it could be fairly and responsibly be concluded by such experts (1) that the drug product has the comparative pharmaceutical quality it is represented to have, (2) that such comparative pharmaceutical quality, to the extent suceptible of quantitation, is demonstrated by methods of statistical analysis, and with levels of confidence, that are generally recognized by such experts, and (3) that such comparative pharmaceutical quality has clinical significance. Such scientific procedure shall also include (4) appropriate controls of test samples for their age and condition of storage in a way generally accepted by such experts and (5) the nature of the control must be stated and an explanation given of the methods used to minimize bias on the part of the observers and the analysts of the results. The study must be conducted double-blind to the extent appropriate.

IV.

It is further ordered, That respondent Lois Holland Callaway, Inc., a corporation, its successors and assigns, and [**355**] respondent's officers, agents, representatives and employees directly or through any corporation, subsidiary, division or other device, in connection with the labeling, advertising, offering for sale, sale or distribution of Vanquish or any other nonprescription internal analgesic product in or affecting commerce, as "commerce" is defined in the Federal Trade Commission Act, do forthwith cease and desist from:

A. Referring, directly or by implication, to aspirin, caffeine or any commonly known ingredient by any word or words without disclosing the common, or usual, name of such ingredient, or

B. Failing to disclose in the advertising of such nonprescription drug product the presence of aspirin when such product contains such ingredient.

So much of Complaint Paragraph Twenty-Nine as it relates to Complaint Paragraph Seventeen is hereby dismissed.

OPINION OF THE COMMISSION

BY CLANTON, *Commissioner.*

I. INTRODUCTION

The American Indians knew that you could make a medicine for treating pain from the leaves and bark of the willow tree. However, it was not until the mid–1800's that acetylsalicylic acid (a substance similar to the salicin contained in willows) was synthesized and at the end of the 19th century it was first marketed commercially under the trade name "Aspirin" by the German concern, Farbenfabriken Bayer AG. During World War I the United States government seized Bayer's American assets and these (including the "Bayer" name) were sold in 1918 to Sterling Drug, Inc. ("Sterling"). At about the same time, the patent on the manufacturing process expired and shortly thereafter Sterling lost the "Aspirin" trademark in private litigation.[1] [2]

Nonetheless, for many years, Bayer aspirin ("Bayer") remained the nation's leading over-the-counter ("OTC" or nonprescription) analgesic (pain reliever). In recent years, competition from other aspirin and acetaminophen-based analgesics has eroded Bayer's market share to the point that it is no longer the market leader. In order to stave off this decline, Sterling has devoted a substantial amount of money to advertisements for its aspirin. In fact, from 1967 through 1973, Sterling spent $118.5 million on television advertising for Bayer.

In 1973 the Commission issued a complaint against Sterling (and its two advertising agencies, Dancer-Fitzgerald-Sample, Inc., and Lois Holland Callaway, Inc.) charging that advertising for Bayer and for four other analgesic products manufactured by Sterling (Bayer Children's Aspirin, Cope, Vanquish, and Midol) violated Sections 5 and 12 of the Federal Trade Commission Act (15 U.S.C. 45, 52).[2] Specifically, the complaint charged that respondents made the following false, deceptive, or unfair claims:

1) Bayer is therapeutically and qualitatively superior to any other aspirin and this superiority has been shown by tests (Comp. ¶¶ 10(A), 20), and the therapeutic superiority has been established (Comp. ¶ 8(A)(1);

2) Bayer Children's Aspirin is therapeutically superior to any other

[1] *Bayer Co. v. United Drug Co.*, 272 F. 505 (S.D. N.Y. 1921).

[2] On the same date, the Commission issued a complaint against American Home Products regarding its advertising of Anacin and Arthritis Pain Formula and a complaint against Bristol-Myers Company regarding its advertising for Bufferin, Excedrin, and Excedrin P.M.

children's aspirin (Comp. ¶ 10(B)), and this superiority has been established (Comp. ¶ 8(A)(2));

3) Vanquish is a more effective pain reliever than aspirin, buffered aspirin, or the largest selling extra strength tablet (Comp. ¶ 12(B)(1),(C)), and this superiority has been established (Comp. ¶¶ 8(B)(1), (C);[3] [3]

4) Cope is more effective for the relief of "nervous tension headache" than any other OTC internal analgesic (Comp. ¶ 12(A)), this superiority has been shown by tests (Comp. ¶ 18), and this superiority has been established (Comp. ¶ 8(A)(3));

5) Vanquish will result in less stomach upset than any other unbuffered OTC analgesic (Comp. 12(B)(2)), and this superior freedom from side effects has been established (Comp. ¶ 8(B)(2));

6) Bayer, Cope, and Midol can relieve nervous tension (Comp. ¶ 15);

7) Cope is the only OTC analgesic containing both a pain reliever and a sedative (Comp. ¶ 22);

8) The analgesic in Midol is other than ordinary aspirin and its stimulant is other than caffeine (Comp. ¶ 26).

The complaint also alleged that respondent's ads failed to disclose that Vanquish, Cope, and Midol contain aspirin and caffeine (Comp. ¶¶ 23, 24, 25) and that ads for Bayer, Vanquish, and Cope made mutually inconsistent claims regarding the superior effectiveness of Bayer and Cope for the relief of "nervous tension headache," and the freedom from side effects of Bayer and Vanquish (Comp. ¶ 17).

Dancer-Fitzgerald-Sample, Inc. was charged with responsibility for all ads relating to Bayer, Bayer Children's Aspirin, and Cope. Lois Holland Callaway, Inc. was charged with responsibility for some of the ads relating to Vanquish.

This case was assigned to Administrative Law Judge Montgomery K. Hyun,[4] who reached an initial decision on January 30, 1981, finding against Sterling on all charges except for the charge related to the making of inconsistent claims. The advertising agency, Dancer-Sample-Fitzgerald, Inc., was dismissed by Sterling in 1976 and entered into a consent order settling the charges against it. 96 F.T.C. 1 (1980). Lois Holland Callaway, Inc. became insolvent and its creditors' committee chose not to defend in this suit. The ALJ found that the ad agency had adequate substantiation for the comparative efficacy and

[3] Complaint paragraph 8(C) does not specifically allege that respondents represented that Vanquish's superior efficacy over the leading "extra-strength" tablet has been established. However, since it appears that it was complaint counsel's intent to allege establishment, and since it appears to have been the parties' understanding that establishment had been alleged (*see* Contested Issues of Fact ¶¶ 2(d), 2(e), 2(f)) and since the issue was tried by the parties (*see* p. 89 of the Initial Decision), we will treat this issue as though it had been appropriately pleaded (*see* F.T.C. Rules of Practice Section 3.15(2), 16 C.F.R. 3.15(2)).

[4] The companion cases against American Home Products Corp. and Bristol-Myers Company were also heard by Judge Hyun.

safety claims it made regarding Vanquish, but he found it liable for failing to disclose the presence of aspirin in Vanquish. [4]

This matter is now before us on the appeal of both respondent Sterling and complaint counsel. Sterling's principal contentions on appeal are:

1) the ALJ erred in finding that Sterling's ads made representations of therapeutic superiority;

2) the ALJ applied inappropriate substantiation standards to the ads;

3) contrary to the ALJ's decision, the nonclinical evidence in the record provides a reasonable basis for the conclusion that Bayer is pharmaceutically and therapeutically superior to other aspirin;

4) the ALJ erred by excluding scientific materials proffered by Sterling; and

5) the order entered by the ALJ is overbroad.

Complaint counsel support the ALJ's order but argue that it should be broader. Specifically, they argue:

1) the order should require the same amount of substantiation for superior quality claims as for therapeutic superiority claims;

2) the order should have broader product coverage;

3) the order should define the substantiation necessary for noncomparative tension claims;

4) corrective advertising should have been required;[5] and

5) Sterling should have been prohibited from making mutually inconsistent performance claims for its products.

Although many of the issues in this case are similar to those recently considered by the Commission in *American Home Products,* 98 F.T.C. 136 (1981), *aff'd,* 695 F.2d 681 (3rd Cir. 1982), and *Bristol-Myers,* Docket No. 8917 (1983) [102 F.T.C. 21], there are some notable differences. As the Commission noted in *American Home Products,* 98 F.T.C. at 362, because aspirin is so homey and commonplace, "a maker of one aspirin-based pain reliever seeking [5] to differentiate its product from the rest faces a formidable marketing task." In order to accomplish this task, both American Home Products and Bristol-Myers attempted to dissociate their products from aspirin and then represent them as special and more effective. Sterling took a different approach. The advertising for Sterling's principal product, Bayer, specifically emphasizes the aspirin content of that product and its superiority over other aspirin-based analgesics. Rather than trying to

[5] Although complaint counsel indicated they did not intend to press the corrective advertising issue (Transcript of Oral Argument p. 27), they did brief this issue and we have dealt with it in this opinion (*infra* pp. 60–61).

disguise the aspirin ingredient,[6] respondent's ads trumpet the fact that Bayer contains aspirin and that, based on tests of competing products, Bayer is the best aspirin on the market. Sterling emphasizes that its advertising for Bayer benefitted the public because it was the only defender of aspirin. It argues that its ads countered "the tremendous volume of advertising" for combination products which "disparag[ed] aspirin." (Transcript of Oral Argument p. 13) Among these "disparaging" ads placed by Bayer's competitors are ads challenged in *American Home Products* and *Bristol-Myers*. Sterling further contends that its advertising was designed to show that the combination products all contained aspirin and that Bayer was as effective as any of them.

The Commission, of course, does not dispute the value of providing consumers with specific product information, especially information which facilitates product comparison. In fact, we encourage that kind of advertising.[7] The issue here, however, is whether Sterling's advertising made certain claims and whether those claims were supportable. In particular, a principal, and unique, focus of this proceeding concerns the extent to which general therapeutic superiority claims can be inferred from representations expressly referring to particular product attributes such as purity, freshness and speed of disintegration. Respondent argues that these ads make only manufacturing quality claims and that manufacturing quality is distinct from therapeutic superiority.[8] As we discuss more fully below, we believe that some of respondent's ads do make therapeutic superiority claims and references to scientific tests in those ads imply that the superiority of Bayer (and Cope) has been established.[9]

In connection with these claims of established therapeutic superiority, we find no reason to depart from our conclusions in *Bristol-Myers* and *American Home Products* that these claims must be supported by well-controlled clinical studies, evidence which Sterling lacked in this case. We also find that some of the separate product attribute claims (or pharmaceutical quality [6] claims) made by Sterling were misleading in light of the evidence relied upon to support those representations. Finally, our decision addresses a variety of other charges concerning noncomparative tension relief claims, unusual ingredient claims and material omission claims that are similar to issues considered in our other analgesic cases. We do note, at this point, two additional differences between this case and those

[6] While there are allegations of failure to disclose aspirin content for Midol, etc., those issues are secondary here to the principal claims involving Bayer.

[7] Indeed, we find ads such as CX 31 unobjectionable.

[8] For a discussion of the difference between manufacturing quality and therapeutic superiority, *see infra* pp. 11–12.

[9] The meaning of "establishment" claims is discussed in *American Home Products*, 98 F.T.C. at 373–376 and in *Bristol-Myers*, slip op. at 18–19 [102 F.T.C. at 331–332].

involving Sterling's competitors, Bristol-Myers and American Home Products. Unlike those cases, the complaint in this matter includes allegations that respondent's therapeutic superiority claims lacked a reasonable basis. Another allegation, unique to this case, is that some of respondent's advertising claims were mutually inconsistent. Complaint counsel contend this practice should constitute a separate violation of the Section 5 of the F.T.C. Act.[10]

II. COMPARATIVE EFFICACY AND SIDE EFFECTS CLAIMS

A. *The Advertisements.*

Paragraphs 8–14 and 18–21 of the complaint allege that respondent Sterling made comparative performance and freedom from side effects claims for Bayer, Bayer Children's Aspirin, Vanquish, and Cope and that these claims were not properly substantiated. In discussing these allegations, we first review respondent's advertisements. It is well settled that the Commission can interpret the meaning of advertisements without necessarily referring to extrinsic evidence. *Bristol-Myers,* slip op. at 4 [102 F.T.C. at 319], *The Kroger Company,* 98 F.T.C. 639, 728 (1981). However, when extrinsic evidence is presented to assist in interpreting ads, that evidence must be considered. *Cinderella Career and Finishing Schools, Inc.* v. *F.T.C.,* 425 F.2d 583, 588 (D.C. Cir. 1970). Accordingly, we have examined all the evidence which has been presented including the ads themselves, expert testimony and copy test results. Additionally, when interpreting advertisements, we consider the net impression made by the ad. *American Home Products* v. *F.T.C.,* 695 F.2d at 687; *Beneficial Corp.* v. *F.T.C.,* 542 F.2d [7] 611, 617 (3rd Cir. 1976), *cert. denied,* 430 U.S. 983 (1977). Therefore, we analyze each challenged advertisement as a whole.[11]

Sterling argues that its ads must also be examined in light of the advertising of its competitors to which it was attempting to respond. (R.A.B. p. 5) Although such a comparison may be helpful in interpret-

[10] The following abbreviations are used in this opinion:

F.	–	Initial Decision, Finding No.
I.D.	–	Initial Decision
CX	–	Complaint Counsel's Exhibit No.
RX	–	Respondents' Exhibit No.
Tr.	–	Transcript of Testimony, Page No.
C.A.B.	–	Complaint Counsel's Appeal Brief
C.An.B.	–	Complaint Counsel's Answering Brief
R.A.B.	–	Sterling's Appeal Brief
R.An.B.	–	Sterling's Answering Brief

[11] In interpreting ads, the Commission is concerned not only with representations conveyed by literal statements, but also with representations reasonably implied by the ads. However, we may not inject novel meanings into ads and then condemn them as unsupported. If an ad conveys more than one reasonable meaning and any one of these meanings is false, that ad may be found in violation of the law. Bristol-Myers, slip op. at 4–5 [102 F.T.C. at 319–320].

Challenged claims must also be material, *i.e.,* likely to influence consumers' purchasing decisions. In this case respondent has not raised any argument regarding materiality. False superiority claims are material because they may discourage consumers from shopping for less expensive and potentially equally effective alternatives. (*See* F. 11) Unsubstantiated tension-relief claims are material because they may encourage excessive use of aspirin, a potentially harmful drug, or otherwise discourage consumers from purchasing more effective products.

ing advertising, it cannot excuse the failure to adequately substantiate the claims which are clearly made in Sterling's ads. Sterling is accountable for the advertising which it promulgated, *see Chrysler Corp.*, 87 F.T.C. 719, 752 n. 43 (1976), and it cannot justify its failings by pointing to the conduct of its competitors. (Indeed, we have already found two of its major competitors in violation of the law.)

The complaint against Sterling alleges that it made eight comparative superiority claims for Bayer, Bayer Children's Aspirin, Vanquish and Cope. The ALJ found that Sterling's ads made all eight of the claims. In addition, the complaint alleges Sterling represented that seven of these claims had been established and the ALJ also found all seven establishment representations had been made. We agree with the ALJ that Sterling's ads make some of the alleged representations and we further agree that three of the comparative claims are represented as having been established. However, we find that some of the ads cited by the ALJ as making certain representations do not make those representations. Further, as we indicated above, we disagree with the ALJ's conclusion that every ad which makes a comparative superiority claim represents that the superiority has been established.

As we described in *Bristol-Myers*, slip op. at 6 [102 F.T.C. at 321], the complaints in these cases require us to distinguish three distinct types of comparative efficacy and freedom from side effects claims. The first group consists of "establishment claims," or claims that superiority has been scientifically *established.* This kind of representation may be made through the use of specific language, such as "medically proven" or through the use of visual aids, such as scientific charts and white-coated technicians. *See American Home Products,* 98 F.T.C. at 374–375. The second type consists of claims of superiority without any indication superiority has been established. An advertiser must [8] possess a reasonable basis for making this type of claim. *See Pfizer, Inc.*, 81 F.T.C. 23 (1972) The third type of claim is puffing, for which no substantiation is required. Puffing claims are usually either vague or highly subjective and, therefore, incapable of being substantiated.[12] Each of respondent's claims can be placed in one of these three groups and the substantiation necessary is dependent upon that characterization.

1. Claim that Bayer is therapeutically superior to any other aspirin.[13]

Paragraph 10(A) of the complaint alleges that Sterling represented

[12] The claim "Bayer works wonders" in CX 27 is an example of puffing.

[13] Complaint paragraphs 8(A)(1) and 10(A).

in its advertisements that Bayer is therapeutically superior to any other aspirin. The ALJ found that Sterling's ads made this claim. We agree with respect to some of the advertisements, but disagree with respect to others.

In evaluating Sterling's advertisements, it is important to understand the relationship between therapeutic effectiveness and "manufacturing quality" (also referred to as "pharmaceutical quality" or "product quality"). Therapeutic effectiveness refers to the medical effects of a drug—its effectiveness as a pain reliever, its freedom from unwanted side effects, and so on. By contrast, manufacturing or pharmaceutical quality refers to the care with which a product containing the drug was manufactured—*e.g.,* its purity (freedom from contaminants), any tendency of the pills to crumble or deteriorate over time, or the ease with which the pills can be dissolved. A recurring issue in this case is whether Sterling's ads made claims of superior therapeutic effectiveness, or whether they claimed only superior manufacturing quality.

We agree with the ALJ that Sterling made representations of therapuetic superiority in some instances. For example, CX 161 states:

> . . . Bayer tested its aspirin for quality against the other leading brands . . . 220 brands in all. 30 separate tests were conducted in 14 different categories. . . . During the 4-year study, tests [9] were made for purity, freshness, speed of disintegration, aspirin content, tablet count, overall quality control.
>
> The results were clear. Bayer was consistently superior. . . .

For several reasons we believe this ad represents that Bayer is therapeutically superior. First, the ad implies Bayer disintegrates faster than the other tested aspirins. As we discussed in *Bristol-Myers,* slip op. at 7 [102 F.T.C. at 322], consumers could reasonably infer that an aspirin that disintegrates faster provides relief faster. Since consumers want relief from pain as rapidly as possible, a pain reliever that works faster would reasonably be considered by consumers to be more effective.

Second, the reference to speed of relief is supplemented by the comprehensive nature of the comparative study. The ad emphasizes that 30 separate product attributes were tested. Although only six of those attributes are specifically mentioned, consumers could reasonably assume that some of those tests related to product effectiveness not only because of the reference to speed of disintegration but also because of the clear statement that a wide variety of attributes was

tested. After all, in purchasing aspirin, consumers are primarily concerned that the product purchased be able to relieve pain.

Respondent argues that this ad and others like it (*e.g.*, CX 47, 48, 79, 109, 155–158) actually discuss manufacturing quality and that speed of disintegration is just another attribute of manufacturing quality. It is true that these ads do mention specific product attributes (such as shelf life) that are not necessarily synonymous with a product's comparative therapeutic efficacy. Nevertheless, the ads refer to product characteristics, such as speed of disintegration, that are closely related to efficacy. In addition, the ads speak in such sweeping terms about quality that a consumer could reasonably infer that the tests measured Bayer in all respects, including efficacy. Certainly nothing in the ad indicates that the tests were limited to attributes relating to manufacturing quality. It is hardly reasonable to expect consumers to guess, without any prompting, that Bayer's tests of aspirin "for quality" omitted the very attribute (efficacy) that consumers value the most.

Respondent also argues that testimony of its expert witness, Dr. Miles, indicates that these ads only make representations regarding quality. (R.A.B. p. 8) We have examined Dr. Miles' testimony and find that although she said she reviewed all of the [**10**] challenged ads (Tr. 9258), she only discussed one of the ads listed above, CX 157. (Tr. 9331–33) Dr. Miles did state that she believed the ad made an "unambiguous quality representation." (Tr. 9332) Nevertheless, it was her opinion that consumers reading (or viewing) an ad such as this one would not infer a message of superior efficacy because consumers do not devote much mental effort to ads: "As I said before, they don't make inferential leaps. They don't do a lot of processing of advertising claims. They don't rationally process advertising communication." (Miles, Tr. 9311, 9333)[14] However, we believe that this analysis actually leads in the opposite direction. Efficacy is the most important feature of an analgesic, and a test of numerous product attributes designed to measure product "quality" would normally be assumed to test efficacy. Only upon application of substantial mental effort (effort which Dr. Miles believed the viewer was unlikely to apply) would it occur to a consumer that none of the tests mentioned in the ad directly measured therapeutic efficacy, and that perhaps effectiveness was *not* tested. Thus, we find that these ads[15] do represent that Bayer is therapeutically superior to other aspirin.

The ALJ also found that a representation of therapeutic superiority was made by advertisements claiming that Bayer was the best pain

[14] Dr. Miles indicated that consumers did not devote mental effort to the Bayer ads because they were boring, poorly made, dull, and not memorable. (Tr. 9259, 9273, 9295, 9303, 9304, 9309)

[15] CX 47, 48, 79, 109, 155–158, 161.

reliever or the best aspirin. (F. 294(b)) Once again, we agree that the representation was made at least by some of the ads cited in the ALJ's decision, although we place less reliance than did the ALJ on the ads' closing tag line. For example, CX 52 is a television ad featuring golfer Lee Trevino. He first describes an AMA study which indicates that aspirin is preferred *over combination products* for relief of pain (*i.e.*, on grounds of therapeutic efficacy). After referring to a separate study on aspirin performed by Bayer, he states, "You see, Bayer tested its aspirin for quality, for purity and for freshness against 220 other brands. The tests showed that Bayer makes the superior aspirin." The ad closes with the tag line, "Aspirin is the best pain reliever. And Bayer is the best aspirin." Because of the emphasis in the ad on the AMA report and the comparative aspirin study, consumers could naturally assume that the comprehensive comparative testing performed by Sterling included tests of relative effectiveness. This inference is especially likely given the fact that Sterling's test is mentioned in the context of the AMA report, which implies that aspirin is more effective than other pain relievers. The tag line at the end of the ad does nothing to alter the impression of therapeutic superiority created by the entire ad.

Respondent argues that the "Bayer is the best aspirin" portion of the tag line is puffing which consumers would not take seriously. To support this, it cites Dr. Miles (RAB p. 9–12). However, what Dr. Miles says is that the phrase [11] "world's best aspirin" is puffery and the phrase, "best [aspirin], all by itself is not likely to lose its puffery characteristic and take on some kind of superior therapeutic meaning." (Miles, Tr. 9271) We agree with this. Indeed, we find that in ads such as CX 13 the phrases, "Bayer is 100% aspirin—the world's best aspirin," and "Bayer works wonders," are merely puffing because the ad does not discuss any comparison of Bayer's "quality" with other brands of aspirin.[16] CX 52 is different. That ad mentions the AMA report and the Sterling test comparing 220 brands of aspirin. The tag line does not appear "all by itself," but appears in a context which invites the viewer to conclude that Bayer is therapeutically superior to other aspirin.[17]

Respondent also argues that the two parts of the tag line should be analyzed separately. It argues that the "Bayer is the best aspirin" portion of the tag line should be considered only in light of the study comparing Bayer to other aspirin. Although respondent would apparently agree that the AMA study involved comparative efficacy, it argues that the Bayer study concerned pharmaceutical or manufac-

[16] *See also*, CX 15, 19, 37, 38, 39, 117, 122, 123, 126, 145, 146, 147, 150, 152.

[17] Other similar ads which imply therapeutic superiority and contain the tag line are CX 50, 54, 56–64, 67–70.

turing quality only and that, therefore, the "Bayer is the best aspirin" tag line implies only pharmaceutical superiority.

We recognize (along with the ALJ, F. 322) that pharmaceutical or manufacturing quality is an attribute of analgesics which may, in some circumstances, be distinct from therapeutic quality. But since consumers buy aspirin only to reduce fever, alleviate pain or lessen inflammation, quality, the ability of a product to do what it is supposed to do, is closely linked to efficacy. In ads such as CX 52 and others which mention Sterling's test of 220 brands of aspirin, no effort has been made to limit the claim to non-therapeutic quality characteristics.[18] Each of these ads implies that Sterling's comparative testing was comprehensive by virtue of both the number of brands and variety of attributes tested. As we mentioned before, the natural inference is that efficacy, the most important feature of any analgesic, was also tested. Although these ads do speak of "quality," that term has not been limited to non-therapeutic quality. Thus, we do not agree with respondent that the mention of Bayer's superiority in the 220 test refers only to quality attributes distinct from efficacy. [12]

However, we do not mean by our decision to prevent Sterling from conveying information regarding Bayer's pharmaceutical or manufacturing quality. Indeed, we recognize that this information may be valuable to consumers. And, even though pharmaceutical quality is closely linked to therapeutic quality, it is certainly possible to convey information limited to non-therapeutic quality attributes. For example, CX 72 states in part, "Sometimes you can even smell a difference in aspirin. If you sense a vinegary odor, that's a sign of possible deterioration. So to get the best quality aspirin, always get Bayer." This ad discusses a particular attribute of product quality, shelf life, that consumers would understand as distinct from therapeutic superiority.[19] Although representations of superior "quality" will usually imply therapeutic superiority, that is not so with respect to CX 72 because it carefully defines product quality in terms of shelf life and thereby avoids making any representation regarding therapeutic superiority.

Respondent finally argues that it was improper for the ALJ to rely upon copy test results (F. 300–302) because the Zeisel Study (CX 520) to which he referred was flawed and because it showed that only a small percentage of consumers (11% for one ad and 13% for another) received a superior efficacy message from two of the ads (CX 52, 157) in question. (RAB p. 13 n. 20) The ALJ did, in fact, find that the Zeisel Study was flawed and that the questions asked of test participants

[18] Other similar ads which imply therapeutic superiority are CX 73–78, 80–83, 105–108, 110–116, 162, 163.

[19] A product with a longer shelf life may well be no more efficacious than one with a shorter shelf life. However, it will maintain its quality for a longer time.

were leading. For this reason, he rejected the results of all but the first two questions asked of participants. (F. 201) Furthermore, the only two questions which the ALJ found unobjectionable direct participants to the ad's major point. For this reason, it appears unlikely that the study would capture inferences and additional meanings drawn by consumers from the ad. Thus, the Zeisel Study cannot confirm respondent's interpretation of the ads. Indeed, another copy test in the record, CX 568 (performed by Audience Studies, Inc.), evaluated CX 50, an ad similar to CX 52. It showed that more than 22% of test participants drew a message of superior efficacy from the ad. On balance, the copy test evidence is not especially helpful because the Zeisel Study was flawed and because its results were contradicted by CX 568. However, in this instance, the representation of Bayer's therapeutic superiority flows clearly and logically from the ads. Therefore, copy tests were not necessary to aid in our interpretations.

2. Claims that Bayer's therapeutically superiority to any other aspirin is established.

Paragraph 8(A) of the complaint charges Sterling with having represented that the truth of the therapeutic superiority claims for Bayer had been established. The ALJ concluded that every [13] ad which made a representation of superior therapeutic efficacy also made an establishment claim, because consumers believe scientific evidence supports every claim of superiority regarding drugs. While we reject the ALJ's reasoning, we reach the same conclusion regarding the advertisements at issue in this case.

Respondent argues there is no justification for the ALJ's conclusion that any ad which makes a therapeutic superiority claim necessarily represents the claim is scientifically established. Sterling notes that a substantial number of ads cited by the ALJ do not mention tests at all, and those ads which do mention tests of Bayer indicate that only quality attributes were tested. (RAB pp. 14–17)

In *Bristol-Myers,* the record did not contain sufficient evidence to sustain the argument that consumers believed that every claim of superiority for an analgesic drug has been established to the satisfaction of the scientific community. Slip op. at 40–41 [102 F.T.C. at 350–351]. For the same reason—since no additional evidence has been presented in this case—we reach the same conclusion here. Nevertheless, as we also stated in *Bristol-Myers,* consumers *would* reasonably infer that a proposition in an ad has been scientifically established if the ad uses language or visual aids which suggest such a foundation. *Id.* at 6 [102 F.T.C. at 321]. Indeed, in this case all of the challenged ads in which we find a representation that Bayer is therapeutically superior include language and pictures which suggest Bayer's thera-

peutic superiority has been proven. All of the ads rely heavily on the AMA report and/or Bayer's own tests.

In *Pfizer,* the Commission found that the challenged ads portrayed a frivolous aura and thus, even though the ads mentioned tests, they did not convey serious scientific overtones. 81 F.T.C. at 59. Here, the opposite is true. Not only are the challenged ads serious in tone but also the format of the ads listed above generally consists of objective evaluations (based on tests) of Bayer's superiority, thus contributing to the scientific aura of these ads. In the ads which mention speed of disintegration, the impression that tests support Bayer's superiority is enhanced by a picture of the AMA report (*e.g.,* CX 155, 156) or by a picture of a booklet containing the results of Sterling's aspirin comparison tests (*e.g.,* CX 47, 48, 157, 158, 161). Even in those instances in which the ads do not mention speed of disintegration, the ads do indicate that Sterling's testing was comprehensive, in terms of both the number of brands included and the number of attributes tested. As discussed above (*supra* pp. 11–12), the natural inference is that these tests measured efficacy and demonstrated Bayer's superiority.[20] [14]

3. Claim that Bayer has been tested and found pharmaceutically and therapeutically superior to all other aspirin tested.

Paragraph 20 of the complaint alleges that respondent's ads represent that Bayer has been tested against 220 other brands of aspirin for quality, purity, freshness, stability, and speed of disintegration and that the tests demonstrate that Bayer is superior to the other brands in all these categories. The ALJ found that advertising for Bayer represented it was superior in overall quality and superior with respect to each of the other four listed attributes. We agree.

Numerous ads discuss the results of tests comparing Bayer with other aspirin. All of those ads indicate that Bayer was compared for quality against other aspirin.[21] The clear message is that the tests showed Bayer to be superior. For example, CX 108 states, "Bayer tested its aspirin for quality against all major brands of aspirin. And Bayer came out way ahead for quality." Other ads use such language as: "For quality—Bayer was shown superior" (CX 76), "The aspirin that tested better for quality" (CX 74), "Bayer was consistently better" (CX 109). In addition, some of the ads that mention the comparison mention specific attributes that were tested: purity (*e.g.,* CX 61,

[20] We do not mean to suggest that every reference in an ad to a study or test necessarily implies that the underlying claim has been scientifically proven or established. (Indeed, *Pfizer* itself is sufficient to disprove this notion.) What we do suggest is that where scientific evidence is cited in support of a claim, absent some explicit qualification it is unlikely that consumers would interpret such evidence narrowly to provide proof for only a limited portion of the claim.

[21] The tests comparing Bayer with other aspirin are mentioned in CX 47, 48, 50, 52, 54, 56–64, 67–70, 72–83, 101–116, 155–158.

79, 109), freshness (CX 47, 73, 155), stability (CX 79, 156, 158), and speed of disintegration (CX 48, 109, 157). Although the ads do not directly state that the tests showed Bayer superior for any specific attribute, superiority with respect to these attributes is clearly implied by the ads. For example, CX 79 states, "Bayer tested its aspirin against every other leading brand. For purity, stability, speed of disintegration, Bayer was consistently better." This ad and others like it (CX 109, 155, 156, 158) focus on the attributes tested as well as overall quality.

Paragraph 20 also charges that the comparative tests demonstrate Bayer's *therapeutic* superiority. This allegation is essentially identical to paragraph 10(A) of the complaint. Since we have previously determined that numerous ads describing Sterling's comparative tests represent Bayer's therapeutic superiority (*supra* pp. 8–12), there is no need to repeat that analysis here. [15]

4. Claims that Bayer Children's Aspirin is therapeutically superior to other children's aspirin and that this superiority has been established.

Paragraph 10(B) of the complaint alleges that respondent represented in its advertising that Bayer Children's Aspirin is superior to other children's aspirin in terms of significant therapeutic effect. Paragraph 8(A)(2) contains the corresponding establishment charge. The ALJ found that both of these representations were made by the challenged advertisements. (F. 339–351) In this instance, we agree with respondent and can find no representation of therapeutic superiority in any of the challenged ads.

For example, CX 183 states:

> It's a fever. But you know what to do. Doctors recommend aspirin to reduce the fever of a cold and relieve the aches. And you choose Bayer Children's Aspirin because when your child is sick, it's good to know you have Bayer behind you. You and Bayer. You take extra care to keep the aspirin safely stored. Bayer takes extra care to keep the aspirin pure and fresh by making over 200 quality control tests on every group of tablets. Part of a special Bayer process. You take extra care to read the label and give the right dosage. Bayer takes extra care by blending two kinds of aspirin crystals instead of one so its aspirin disintegrates smoothly and gently. Part of a special Bayer process. So when your child is sick, its good to know you have Bayer behind you.

For several reasons, this ad does not represent that Bayer is therapeutically superior. First, although the ad does indicate the Bayer manufacturing process is special, the overall impression created by the ad is not one of uniqueness. To be sure one could reasonably infer from this ad that Bayer Children's Aspirin is a good product or even that it is one of the best brands of children's aspirin available. But that

characterization does not constitute a representation that Bayer Children's Aspirin is *the best* children's aspirin. In addition, a central theme of this ad is that Bayer Children's Aspirin is a well-made aspirin. Unlike ads for regular Bayer discussed above (*supra* pp. 8–12) CX 183 does not stress a comparison between Bayer Children's Aspirin and all other brands. Indeed, the tests that are mentioned are quality control tests, not comparative tests. The ad does not in any way imply that Sterling is the only manufacturer which performs such tests. Thus, without [16] additional evidence we are unwilling to conclude that consumers would make the inferential leap from the representations in CX 183 to a representation of therapeutic superiority.[22]

Another Bayer Children's Aspirin ad cited by the ALJ is CX 167. That ad states, in part, "No one makes aspirin like Bayer. No one purifies aspirin like Bayer. No one protects aspirin like Bayer." Unlike CX 183, this ad makes a superiority claim for Bayer Children's Aspirin. Nevertheless, the message is not that Bayer Children's Aspirin is therapeutically superior but that Sterling's manufacturing process is superior. As we indicated above in connection with ads for regular Bayer aspirin, a representation of superior manufacturing quality does not necessarily imply therapeutic superiority. (*supra* pp. 11–12) In ads such as CX 167, Bayer Children's Aspirin's superiority is narrowly defined in terms of specific quality attributes. Thus, there is no representation of therapeutic superiority.[23]

Finally, the ALJ discusses CX 176. That ad opens with a mother concerned about a sick child. She consults a doctor, "she keeps the patient quiet and she gives her children's aspirin. . . . She chooses orange-flavored Bayer Aspirin for children because she knows Bayer makes the best children's aspirin." Although this ad does state that Bayer makes the best children's aspirin, the reference to "best" constitutes puffing. The tone is homey, familiar and secure. In this context, "best" implies only that Bayer Children's Aspirin is a dependable product that a parent can feel secure in giving to a sick child.[24]

5. Claims that Vanquish is a superior pain reliever and that that superiority has been established.

Paragraphs 12(B)(1) and 12(C) of the complaint allege that respondent represented in its advertisements that Vanquish is a superior pain reliever to aspirin, buffered aspirin and the largest selling "extra strength" tablet. Paragraphs 8(B)(1) and 8(C) allege this superiority

[22] Other ads which we find do not claim superiority are CX 182, 184, 209.

[23] Other ads representing pharmaceutical superiority only are CX 175, 185, 188, 196, 197, 205.

[24] Other ads in which "best" constitutes puffing are CX 168–170, 176–181, 194, 195, 198, 201–203.

has been established. The ALJ found Sterling had represented that Vanquish was a superior pain reliever to aspirin, buffered aspirin and to the largest selling "extra strength" tablet. We agree with this finding, although we find that some of the ads cited by the ALJ did not make the challenged representations. We also disagree with the ALJ's finding that the ads make establishment claims. [17]

Vanquish is represented as a superior pain reliever to aspirin in CX 224 and 226. CX 224 states:

> Vanquish is different. It gives you the proven effectiveness of aspirin in this tablet, plus extra medications in these. . . . Vanquish is the only leading pain reliever you can buy that combines the extra strength of three medications with two gentle buffers.

The point of this ad is that Vanquish starts with aspirin and adds extra medication. Consumers could reasonably assume that the purpose of this extra medication is to provide "extra strength"—*i.e.,* extra pain relief. As the court noted in *American Home Products* v. *F.T.C.:*

> Not only credulous purchasers are apt to conflate the idea of more pain reliever with that of more pain relief, but as the Commission explains in its brief, even rational and careful consumers will be apt to place such an interpretation on the advertisements, "If the presence of more pain reliever in a product did not result in greater pain relief (as may well be true of Anacin), disclosure of the extra amount could be a clear liability since consumers would logically expect that it contributed to an increased price." 695 F.2d at 696.

Another ad, CX 226, represents that Vanquish is a superior pain reliever because it has more ingredients than aspirin. In this ad, a man states that he wants more than buffered aspirin for his headache and the announcer notes that Vanquish has added extra medication to buffered aspirin. This clearly implies that Vanquish is a superior pain reliever to aspirin (and to buffered aspirin). CX 245 and 251 also make this implication. Both of these ads state that Vanquish has extra strength which the "leading buffered product" lacks. Ten other ads represent that Vanquish is superior to the largest selling extra-strength tablet.[25] CX 261 is typical of these ads and it states, "Vanquish contains more pain relievers than the largest selling extra-strength tablet."

Respondent argues that Vanquish was introduced in 1966 as a defensive measure to protect Sterling's share of the OTC analgesics market which was being steadily eroded by other "extra-strength" products. It contends that the Vanquish ads were designed not to represent that Vanquish was therapeutically superior but to describe the product to consumers who would [18] otherwise purchase a com-

[25] The ten ads are CX 252, 255, 256, 258–264.

petitor's "extra-strength" product. Although we appreciate that the market for OTC analgesics is highly competitive, it is important that advertisers not compete with false or deceptive advertisements. An advertiser must still be able to substantiate any claims it makes in attempting to compete.

The complaint also alleges, and the ALJ found, that Sterling made establishment claims regarding Vanquish's superiority over aspirin, buffered aspirin, and the largest selling "extra-strength" tablet. He found that the representation of establishment was made by the use of the phrase "medically-proven ingredients" in CX 254 and the phrase, "the proven effectiveness of aspirin" in CX 224. Also, he found that the mortar and pestle used in several of the ads (CX 224, 226) were "chemist's instruments" which help convey the impression that superiority is predicated upon scientific fact.

We are unable to agree that any ad represents Vanquish's superiority has been established. Although the phrase "medically-proven ingredients" might imply scientific testing, the phrase is used only in one ad and that ad does not compare Vanquish with any other product. Similarly, CX 224 states that aspirin has been proven effective but does not indicate or imply that anything has been proven regarding Vanquish. Finally, the mortar and pestle in CX 224 and 226 do not, in the context of those ads, appear to be chemist's instruments. Instead, they are used to demonstrate that Vanquish is a combination of several ingredients. Although in other instances a mortar and pestle might conceivably convey establishment connotations, they do not do so here.

6. Claims that Vanquish will cause less stomach upset than other OTC analgesics and that this representation has been established.

Complaint paragraph 12(B)(2) alleges that respondent represented that because Vanquish contains "gentle buffers"[26] it would cause less stomach upset than any OTC analgesic not containing buffers. Paragraph 8(B)(2) alleges that respondent represented that Vanquish's superior freedom from side effects had been established. The ALJ found respondent had represented Vanquish causes less stomach upset because it is buffered and we agree.

For example, CX 245 discusses the plight of Tuesdee Testa, "a successful female jockey" who "can't afford a headache." The ad observes that "she wants more than just extra strength, she [19] wants gentle action." After several scenes of Ms. Testa in action, the ad comments, "Look—this leading extra-strength pain reliever has no buffers. . . . Vanquish—it gives you extra strength and gentle buffers.

[26] The "gentle buffers" mentioned in the Vanquish advertising are two antacids, aluminum hydroxide and magnesium hydroxide.

Vanquish—all the strength you need for your headache pain, yet gentle enough to your system." This ad implies that Vanquish is gentle because it has buffers and that a product lacking buffers will be less gentle to the system. Respondent argues that it was merely attempting to inform consumers that Vanquish is a multi-ingredient product. However, the ad clearly does more than that. It represents that the buffers in Vanquish make it a product that is less likely to cause stomach upset.[27]

Most of the challenged Vanquish ads promote it as "the only leading pain reliever you can buy that combines the extra-strength of three medications with two gentle buffers," (CX 224) or contain statements such as, "Vanquish is different. It gives you the well-known pain reliever in this tablet, plus extra medication in this tablet and this tablet, and buffers as in this one. Three headache relievers and two gentle buffers. . . ." (CX 241) Statements such as these imply that Vanquish will produce less stomach upset than an unbuffered analgesic. After all, why else would respondent advertise that Vanquish has buffers and some other products do not? Why else would it refer to the buffers as "gentle"? The natural inference from such advertisements is that Vanquish is "gentler" or causes less stomach upset.[28] Indeed, even respondent's expert, Dr. Miles, appears to concede that consumers may draw inferences from the mention of an ingredient in an ad for a product. (Tr. 9492–93.

However, we disagree with the ALJ's finding that the challenged ads represent that Vanquish's superior freedom from side effects has been established. As we stated above consumers would not infer that claims in an advertisement had been established merely because the ad shows a hand using a mortar and pestle to grind tablets. And once again, we disagree with the ALJ's conclusion that every claim of comparative superiority impliedly represents that the superiority has been established. Thus, we find that respondent represented that the presence of buffers in Vanquish causes it to produce less stomach upset than analgesics not containing buffers. However, the ads do not represent that this claim has been established. **[20]**

7. Claims that Cope is more effective than other analgesics and that this representation has been established.

Paragraph 12(A) of the complaint alleges that Sterling's ads represent that Cope is more effective for the relief of nervous tension headache than any other OTC analgesic. Paragraph 8(A)(3) alleges

[27] Other ads disclosing the presence of buffers in Vanquish and containing the statement that Vanquish is gentle to the system are CX 246, 247, 251.

[28] Ads which promote Vanquish as a product containing gentle buffers are CX 224, 226, 235, 236, 241–247, 250–256, 258–264.

that respondent represented that this claim has been established.[29] The ALJ found that all of these claims had been made, and we agree.

CX 272 is typical of the ads for Cope. It states:

> Important studies made at the world's leading headache clinic show that for relief of severe nervous tension headaches a combination of a pain reliever and a sedative provides greater relief than either medication alone. Of all leading remedies you can buy for ordinary nervous tension headaches, only Cope combines a gentle relaxer with a powerful pain reliever for really effective relief.

The clear and direct message of this ad is that Cope is superior for the relief of nervous tension headache to any other analgesic[30] because of the formulation which it alone has. The importance of this formulation is emphasized by the reference to tests in the ad.[31] Respondent again argues that the Cope ads were merely a description of Cope's ingredients designed to introduce the product to the market. (R.A.B. p. 88) However, these ads plainly go further and describe the purpose of Cope's formulation and inform consumers that because of this formulation, Cope gives more effective relief.

This same ad also represents that Cope's superiority has been shown by studies. Read literally, the ad claims studies have demonstrated that a combination of an analgesic and sedative provides greater relief for *severe* nervous tension headache and that Cope with a similar formula provides relief for *ordinary* [21] nervous tension headaches. Given this juxtaposition, it seems reasonable for consumers to infer that if tests show that a combination of sedative and analgesic provides more effective relief for severe nervous tension headache, these results would also support the conclusion that the Cope formula provides more effective relief for ordinary nervous tension headaches. Indeed, this inference is almost inescapable.[32]

There are numerous other indicia of establishment in the ad. It mentions "important studies made at the world's leading headache clinic," the announcer is holding what appears to be a copy of a report, and he is standing in a room lined with ponderous books. The words and the visual images of this ad imply that the basic message of the ad (Cope's superiority) has been established.[33]

---

[29] Paragraph 18 alleges that respondent's ads represent that tests or studies prove this superior efficacy. This is the same allegation contained in paragraph 8(A)(3). (*See supra* pp. 14–15)

[30] The Cope ads actually compare Cope to the "leading remedies." Consumers could reasonably assume that a product which is superior to the leading remedies is superior to all remedies.

[31] Other ads which imply that Cope is superior for the relief of nervous tension headache to any other OTC analgesic are CX 273–276, 283, 287, 292–294.

[32] Other ads which imply that Cope's superior efficacy for the relief of nervous tension has been proven by tests or studies are CX 283, 287.

[33] Other ads making an establishment claim are CX 283 and 287.

B. *Required Substantiation for Establishment Claims*

As our analysis of the challenged advertising has shown, Sterling has represented that it is established that Bayer is therapeutically superior to any other aspirin and that it is established that Cope is more effective for the relief of nervous tension headache than any other nonprescription internal analgesic. Paragraph 28 of the complaint alleges that the truth of these claims, in fact, has not been established, and that the establishment claims are, therefore, false. The ALJ agreed. It was his determination that well-controlled clinical studies are necessary to establish an analgesic's comparative superiority, and that Sterling did not possess that sort of evidence.

Sterling has appealed the ALJ's conclusion. First, it argues that the ALJ's approach conflicts with the reasonable basis theory enuciated in *Pfizer* in that *Pfizer* precludes finding a violation based simply on a conflict in scientific opinion. (R.A.B. 20–21) Second, Sterling contends that the establishment (and substantial question) theory has improperly [22] shifted to it the burden of proof. (R.A.B. pp. 21–22) Third, it argues that the ALJ required it to possess substantiation "more extreme than the level of certitude required by Congress and the Food and Drug Administration for the marketing of drugs." (R.A.B. p. 22) Finally, respondent argues that the First Amendment "precludes the proscription or restriction of commercial speech in areas of good faith differences of opinion." (R.A.B. p. 23) We disagree with all of respondent's arguments for the reasons set forth below. We further hold that it did not establish the superiority of either Bayer or Cope.

As we explained in *Bristol-Myers,* the establishment theory is not a new theory of advertising substantiation. Slip op. at 18–19 [102 F.T.C. at 331–332]. It is based on the straightforward notion that when an advertiser represents in its ads that there is a particular level of support for a claim, the absence of that support makes the claim false. Therefore, the inquiry contemplated by *Pfizer* for reasonable basis claims does not conflict with the more narrowly focused inquiry involved where representations are made that a claim has been established or scientifically proven. Indeed, as we noted in *Bristol-Myers* (*Id.* at 18–19 [102 F.T.C. at 331–332]), a similar approach has been used in a number of post-*Pfizer* substantiation cases.

Respondent's argument that the burden of proof has been shifted to it is also incorrect. The complaint alleges that the establishment claims made by Sterling's ads are false. It is complaint counsels' burden to prove the falsity of those claims by proving that the claims have not been established.

Respondent further claims that it is being required to produce an

excessive level of support for its claims. It bases this argument upon the Committee Report accompanying the 1962 amendments to the Food, Drug, and Cosmetics Act which noted that substantial evidence of efficacy is necessary before a drug can be marketed. However, that Report also recognized that there will usually be differences of opinion among scientists regarding the drug. S. Rep. No. 1744, 87th Cong. 2d Sess. part 2 at 6 (1962). We do not prevent respondent from advertising its products when such differences of opinion exist, provided, of course, that it does not represent that the position supporting its products has been scientifically established. To support its establishment claims, we require respondent to have supporting evidence of the type and quantity that is acceptable to the scientific community.

Respondent's constitutional argument also must fail. As respondent has noted in its brief (R.A.B. p. 23), the Supreme Court has indicated that regulation of false, misleading or deceptive advertising is not barred by the First Amendment. *Bates* v. [23] *State Bar of Arizona,* 433 U.S. 350, 383 (1977); *Virginia State Board of Pharmacy* v. *Virginia Citizens Consumer Council, Inc.,* 425 U.S. 748, 771 (1976). To the extent that respondent claims in its ads that its products' superiority has been established when in fact it has not, the ads are false. There is, therefore, no constitutional impediment to the regulation of respondent's establishment claims.

1. Establishment of Bayer's therapeutic superiority.

A substantial portion of the briefs in this case is devoted to a discussion of the type of evidence necessary to establish the therapeutic superiority of one brand of aspirin over others. Sterling argues at great length that comparative superiority can be demonstrated without the use of well-controlled clinical tests.[34] (R.A.B. pp. 24–42) In support of Bayer's therapeutic superiority, Sterling has presented a substantial amount of nonclinical evidence, including studies comparing impurities in aspirin tablets and evidence regarding the manufacturing of Bayer. Complaint counsel argue that only well-controlled clinical studies can demonstrate the superiority of one brand of aspirin over another. The ALJ agreed with complaint counsel (F. 416) and determined that it has not been established that Bayer is superior to any other aspirin in terms of pain relief and freedom from side effects. (F. 474, 489)

The record contains the testimony of several expert witnesses who testified regarding the type and amount of evidence necessary to

[34] In a well-controlled clinical test, drugs are tested on real patients having actual symptoms. It is not disputed in this case that the elements of a well-controlled clinical test are the use of an appropriate pain model, replication of results, experienced unbiased investigator and adequately trained personnel, a written protocol, double-blinding, use of a placebo control, use of appropriate predetermined analytical techniques, and statistical and clinical significance of the results. (F. 417) *See Bristol-Myers,* slip op. at 24–27 [102 F.T.C. at 335–339].

establish the comparative superiority of one brand of aspirin over others. Based upon our analysis of that testimony, we conclude that at the present time, the relevant scientific community would not regard superiority as established unless supported by the results of well-controlled clinical tests. The record also shows that the only clinical study in the record comparing Bayer with another brand of aspirin (RX 450) does not show any clinically significant difference between Bayer and the other brand tested.

Numerous experts testified for respondent in this case regarding the type of evidence necessary to establish comparative superiority. However, only two of the experts, Drs. Alvan Feinstein [24] and William Fields were qualified as experts in the area of comparative testing of analgesics (F. 110, 115) and the record makes it clear that Dr. Fields had no experience in the testing of mild analgesics (such as the ones involved in this case) (Tr. 16573). Two of Sterling's other experts, Drs. Banker and Rhodes, were qualified only in the formulation and processing of drugs, and neither had any experience with well-controlled tests involving subjective response methodology. (Banker, Tr. 12872; Rhodes, Tr. 11095) Dr. Scoville was qualified only in FDA practices and procedures (F. 134), and he testified regarding the requirements of FDA regulations. The other two doctors who testified on behalf of respondent were Sterling employees who were not qualified as experts in any particular field (Drs. Tainter and Trout; F. 159, 162). Four experts testified for complaint counsel, two of whom, Drs. DeKornfeld and Moertel, were experts in the testing of analgesics (F. 34, 54). The third expert, Dr. Grossman, was qualified as an expert in the field of gastroenterology and aspirin side effects (F. 40) and the fourth expert, Dr. Miller, was qualified as an expert in the formulation and pharmaceutical analysis of drugs. (F. 49).

Although all of these experts have experience with analgesics the record makes it clear that some of them are not experienced in comparing analgesics for the purpose of evaluating relative efficacy. For example, respondent's expert, Dr. Banker, conducted tests of analgesics for the purpose of comparing drug delivery systems[35] (Banker, Tr. 12870), and Dr. Rhodes' tests compared features of drug manufacture such as compaction pressures (Rhodes, Tr. 11084). Neither tested comparative efficacy. Complaint counsel's experts Drs. Grossman and Miller were also not experienced in comparing analgesic efficacy. It is the testimony of experts with experience in the testing of comparative analgesic efficacy that must be given the greatest weight in determining the type of evidence necessary to establish the superiority of one brand of aspirin over others. That testimony (including testimony

[35] A "drug delivery system" is the form given to a dose of a drug. Examples of three delivery systems used for aspirin are tablets, capsules, and effervescent powder.

of respondents' witnesses Drs. Feinstein and Fields) provides strong evidence supporting the conclusion that experts will not regard superiority as established unless that conclusion is supported by clinical evidence.

The need for clinical tests to establish comparative superiority was clearly stated by Dr. DeKornfeld:

> A claim of comparative superiority, I feel quite strongly that a minimum of two carefully controlled clinical comparisons, both showing statistical significance in favor of one of the two compared drugs, is essential to establish the claim of clinical superiority. . . . [M]ost people working [25] in this area will accept two studies showing the same thing done under appropriate circumstances as establishing a claim. If the studies are lacking or if they are controversial, it would not be established. (Tr. 8388, 8391)

Dr. Moertel expressed the same idea. Speaking of clinical testing, he said: "it's the only way we know to properly establish therapeutic superiority. . . ." (Tr. 6255–56) Even respondent's expert Dr. Feinstein indicated that clinical tests provided the best evidence regarding patients' subjective responses. (Feinstein, Tr. 16413)

The experts also explained why clinical tests were necessary. There is general agreement among all the experts that no direct correlation has been demonstrated between the amount of aspirin appearing in the bloodstream at any time and the onset, intensity or duration of relief afforded by the aspirin. (Moertel, Tr. 6291; DeKornfeld, Tr. 8409; Banker, Tr. 12940; Feinstein, Tr. 16482; Danhof, Tr. 17269). Therefore, studies that examine the amount of drug in the bloodstream are not reliable for comparing analgesic performance of various brands of aspirin. Consumers' perceptions are not adequate because consumers cannot evaluate for themselves the efficacy of drugs. (DeKornfeld, Tr. 8421) The reason for this is that consumers' expectations regarding drug performance play a powerful role in influencing the response to drugs. (Feinstein, Tr. 16289) Indeed, this "placebo effect" may produce pain relief in 40–50% of the subjects in controlled tests who receive pharmacologically neutral substances. (Feinstein, Tr. 16322) Finally, the pain for which aspirin is taken is normally self-limiting—it will disappear regardless of what drug is taken. Thus, for these reasons, it is necessary to conduct well-controlled clinical studies in order to establish the superiority of a given brand of aspirin.

Respondent argues that although clinical tests are necessary when comparing different drugs, experts would not use controlled clinical trials to compare different formulations of the same drug. In support of this, respondent cites testimony of both Drs. Feinstein and Fields. Sterling also contends that no expert would ever recommend using a controlled clinical trial to compare different formulations of the same

drug because that would be akin to using a jet airplane to cross the street—"theoretically possible, but hardly the most efficient or sensible way to travel." (R.A.B. pp. 25–27) For this reason, respondent contends that the pharmaceutical evidence it has presented (evidence regarding rates of dissolution, etc.) can establish Bayer's therapeutic superiority. Nevertheless, the expert testimony in this case shows that at this time the consensus of experts would require clinical tests to establish the comparative superiority of any mild analgesic, even to substantiate the superiority of one brand of aspirin over another.

The reason that nonclinical evidence (such as a blood level study) is inadequate to establish comparative superiority was explained by Dr. Moertel: **[26]**

> We simply do not know at this point in time what value, if any, blood level studies have in determining comparative efficacy of mild analgesics because the studies to determine the correlation between blood levels and therapeutic effectiveness have simply not been conducted. . . . [R]ight now we do not know whether, for example, a high quick peak is good or bad in getting the most ideal therapeutic effect from salicylates because these studies have never been conducted. (Tr. 6291–92)

Dr. DeKornfeld held the same belief and stated, "I don't believe that blood levels can be directly translated at any time into establishing clinical effectiveness, unless clinical effectiveness is also measured independently." (Tr. 8409) He also stated that there is "very little, if any" relation between either the product formulation or pharmaceutical quality (size, shape, manufacturing process, presence of substances other than active ingredients) and the therapeutic superiority of any mild analgesic which meets the requirements that permit it to be marketed in this country. (Tr. 8414–15) Dr. DeKornfeld gave his opinion that the 223 Study (CX 448), a study which compared Bayer with 220 other brands of aspirin, could not be used to draw any conclusion regarding the therapeutic superiority of Bayer because it compared only physical and chemical characteristics and differences in those characteristics were not likely to have an impact on clinical effectiveness. (Tr. 8415–16)

An analysis of the testimony of respondent's experts shows that they, too, recognized the value of and need for well-controlled clinical tests. Dr. Feinstein agreed that the amount of a drug present in the blood is not well correlated with clinical analgesia. (Tr. 16413) He also conceded that there was not a high correlation between comparisons of pharmaceutical characteristics and clinical analgesia (Tr. 16415), and that in the absence of clinical trials he would have no way of saying that one drug was better than another (Tr. 16417). He did indicate that in some instances doctors must choose between two drugs (or between two brands of the same drug) and that in making

that sort of decision a doctor will use whatever evidence is available to make that choice even though the evidence might not demonstrate therapeutic superiority. (Tr. 16425–27) However, all that the doctor could then feel confident of would be that the chosen drug was equal to the one not selected. Whether or not it was superior could only be determined with clinical evidence. (Tr. 16427)

The other expert with experience in testing who testified for respondent was Dr. Fields. Respondent cites his testimony in support of the proposition that clinical tests are not necessary when comparing two formulations of the same drug. (R.A.B. p. 27) However, Dr. Fields has no experience in the testing of mild analgesics and his testimony cited by respondent related to a study in which he participated that measured aspirin's [**27**] effectiveness in preventing clotting and lowering the risk of stroke. Prior to conducting this study, a brand of aspirin had to be selected to administer to participants. Based upon nonclinical evidence (including the 223 Study, CX 448) Bayer was selected. Dr. Field's testimony makes it clear that the considerations of the scientists conducting the study were peculiar to the study. For example, they were concerned about factors which might tend to prevent double blinding and about whether the manufacturing process would encourage the decomposition necessary to prevent coagulation. (Fields, Tr. 16590, 16596) Although these considerations might be important in selecting an aspirin to test on potential stroke victims, they are not relevant to establishing the therapeutic superiority claimed by Sterling for Bayer.

Sterling also attempts to rely for support on several scientific articles which it contends indicate that experts do not require clinical tests when comparing two formulations of the same drug. (R.A.B. p. 27) But respondent's own expert, Dr. Feinstein, admitted that both of the articles cited by respondent were generally quite positive about the importance of well-controlled clinical trials and that the point of those articles was that clinical trials must be done carefully. (Tr. 16472–73)

The evidence in this case also shows that in the past Sterling demanded that its competitors rely on clinical studies to support their comparative claims. In 1970, in response to advertisements by Bristol-Myers claiming superiority for Excedrin, Sterling's advertising agency wrote on Sterling's behalf to the television network and argued that such superiority claims should not be made unless substantiated by well-controlled clinical tests. (CX 347C-E) Although it is true that Excedrin has a different formula than aspirin, Sterling requested that the same clinical test standard be imposed on other marketers of analgesics who claim superiority over aspirin for any analgesic differing in any way from the standard aspirin tablet. This request was

made again in 1974 by Dr. Monroe Trout, the Senior Vice President and Director of Medical Affairs for Sterling. In an appearance before the FDA's OTC Analgesics Panel, he suggested, "... that OTC analgesic products containing aspirin, *with or without additional ingredients*" disclose on their labels that the product is not superior to two 5-grain aspirin tablets "unless the superiority claimed or implied by such variance is adequately established by well-controlled studies of pain relief, anti-pyresis, anti-inflammatory or side effects." (CX 456M; emphasis added) Dr. Trout made clear in his statement that the same clinical testing requirement should be imposed upon a manufacturer claiming superiority for an analgesic that was merely a larger-than-normal dose of aspirin. (CX 456M) Thus, Sterling would have required its competitors to perform clinical tests to establish the superiority of one analgesic containing only aspirin over another analgesic containing only (albeit a lesser amount of) aspirin. **[28]**

Respondent objects to these references to its prior statements and contends that the ALJ has used them to estop Sterling from asserting a different position in this case. (R.A.B. p. 31) We do not believe that statements made by Sterling in the past prevent it from expressing different views in this case. However, Sterling's prior statements do provide evidence of the fact that scientists involved in the manufacture of analgesics require well-controlled clinical studies to establish superiority. These statements also show that Sterling was aware of the significance of clinical testing. Thus, we find that Sterling's prior statements are relevant and probative of the need for clinical tests to establish Bayer's superiority.

Respondent argues that it would be inconsistent with FDA policy to hold that nonclinical evidence does not constitute a reasonable basis for its claims. Further, it argues that in some instances the FDA will permit a drug to be marketed based solely upon pharmaceutical studies. Respondent has cited several examples of drugs which were qualified by the FDA as safe and effective even though no clinical studies were submitted. (R.A.B. pp. 37–42) It is true that the FDA would permit some internal analgesics to be marketed without any clinical testing. (45 FR 77807–08, Nov. 24, 1980) However, these are analgesics whose ingredients are identical to some other drug on the market which has already been proven safe and effective with clinical studies. If it can be shown that the new drug ("generic") is absorbed into the body at basically the same rate and to the same extent as a drug already on the market, the new drug will be assumed equally safe and effective as the drug already on the market. Differences in the rate or extent of absorption are regarded as significant only if they "would result in therapeutic failure or hazard to the patient." 42 FR

1626, January 7, 1977.[36] Thus, this nonclinical evidence is only used by the FDA to support a conclusion that drugs are *equivalent,* not that one is superior to another. Indeed, the FDA Commissioner expressed the belief that if two drugs are manufactured in compliance with good manufacturing practice, if they contain identical amounts of the same active ingredients and if they do not present any problems regarding rate or extent of absorption, then it is reasonable to assume that the drugs are of equal efficacy. 42 FR 1625, January 7, 1977. This statement basically echoes the comment of Dr. Feinstein that nonclinical data could lead to a conclusion that two brands of aspirin are at least equally effective. (Tr. 16427) **[29]**

Although the F.D.A. has never directly considered superiority claims for aspirin, it did consider a somewhat analogous issue. The F.D.A. OTC Analgesics panel was presented with nonclinical evidence regarding buffered aspirin which showed that it was absorbed into the bloodstream more rapidly than unbuffered aspirin. (This is similar to the blood level data submitted in this case by Sterling to justify its claims of Bayer's superiority.) Despite this evidence, the panel stated that no conclusion could be drawn regarding whether buffered aspirin provides more rapid relief, greater relief or more prolonged relief than unbuffered aspirin. 42 FR 35470, July 8, 1977. Thus, the FDA panel believed that blood level data could not be used to determine the comparative superiority of buffered aspirin and that controlled clinical studies were necessary to support claims of superiority. Therefore, our determination that two well-controlled clinical studies would presently be required by experts to establish Bayer's superiority over other brands of aspirin is in no way inconsistent with FDA policy or regulations.

Finally, respondent argues that it would not be feasible to conduct a well-controlled clinical study comparing Bayer with all other brands of aspirin because there are more than 200 other brands and such a test would be prohibitively expensive. (R.A.B. pp. 27–28) It also argues that it would not be ethical to conduct such a clinical trial because it would not provide a clear benefit compared to the risk of the study. (R.A.B. p. 34) However, Dr. Feinstein's testimony indicates that it would not be necessary to test all brands of aspirin in a well-controlled clinical trial in order to establish therapeutic superiority. Pharmaceutical tests could be conducted of all brands of aspirin (as Sterling has already done). Then a clinical trial could be performed on two or three of the brands in order to demonstrate whether pharmaceutical differences correlate with therapeutic differences. This sort of scheme could clinically prove superiority and would be feasible. (Fein-

[36] *See U.S. v. Generix Drug Corp.*, 51 L.W. 4282 (1983), which restricts the number of generic drugs that can be marketed without new drug applications.

stein, Tr. 16462–63) Dr. Feinstein also felt that this sort of testing scheme would overcome any ethical barriers that might block a large-scale clinical study of analgesics. (Feinstein, Tr. 16462)

Thus, it is the consensus of the experts with experience in comparing analgesic efficacy who testified in this proceeding that at this time well-controlled clinical tests are necessary to establish the comparative superiority of one brand of aspirin over others. We emphasize that we are not attempting to decree what constitutes scientific establishment because this standard may change with time. Indeed, we recognize in this case, as we recognized in *Bristol-Myers,* slip op. at 69 [102 F.T.C. at 373], that relevant experts might, in some instances, regard a proposition as established even if the clinical tests do not meet all of the criteria set out above. But, as we discuss below, the substantiation possessed by Sterling was plainly inadequate to substantiate the claims it made regarding Bayer's superiority. **[30]**

The record in this case does contain the results of one well-controlled clinical study comparing Bayer with other analgesics. That is the Lasagna-DeKornfeld Study conducted by Drs. Louis Lasagna and Thomas DeKornfeld (who testified for complaint counsel in this proceeding).[37] Its results were published in 1962 in the *Journal of the American Medical Association.* The study was randomized, placebo controlled, and double-blinded. According to Dr. Robert John who was medical director of respondent's Glenbrook Laboratory from 1971–1974 (and who testified for complaint counsel), respondent was aware of the study (John, Tr. 5546–47). Furthermore, respondent relied on the study to support advertising claims made to the public and to support complaints to the FTC concerning competitors' advertising. (Admissions 713, 714 in CX 678)

The Lasagna-DeKornfeld Study tested the analgesic performance of five OTC analgesics. Two of the five analgesics tested, Bayer and St. Joseph's, were plain 5-grain aspirins. The other three, Anacin, Excedrin and Bufferin, were combination products. Doses were administered to test subjects suffering pain and then the amount of pain relief received by each subject was recorded at seven time intervals after administration. After examining the results of the study, it was the conclusion of respondent's expert Dr. Feinstein that the study did not show any difference in therapeutic effectiveness between any of the products tested. (Tr. 16397) Indeed, Dr. Feinstein agreed that the Lasagna-DeKornfeld Study did not indicate any clinically important difference between Bayer and St. Joseph's aspirin, and when asked which of the two brands he would choose, Dr. Feinstein indicated that he would select the cheaper. (Tr. 16438) The Lasagna-DeKornfeld

[37] This study was undertaken in 1960 at the request of the F.T.C. in order to evaluate superiority claims made in advertising for each of the tested analgesics. (DeKornfeld, Tr. 8332)

Study is significant because it shows that it is possible to conduct a well-controlled clinical study comparing different brands of aspirin. It is also the only such study introduced in this case and it failed to demonstrate a clinically significant difference between the brands tested. Thus, it clearly has not been established that Bayer is therapeutically superior to all other brands of aspirin.

2. Establishment of Cope's superiority.

The other establishment claim made by respondent was that Cope is more effective than any other nonprescription internal analgesic for the relief of nervous tension headaches. Respondent does not dispute that well-controlled clinical tests are necessary to establish the therapeutic superiority of one drug over a different drug. In fact, in its brief, respondent quotes a statement [31] from its expert, Dr. Feinstein, that pharmacokinetic information cannot be used to compare two different drugs. (R.A.B. p. 26) However, the clinical evidence that respondent has submitted is inadequate to establish that Cope is superior to other OTC internal analgesics for the relief of nervous tension headache.

First, respondent has submitted evidence regarding the contents of a Cope tablet.[38] Sterling contends that the larger amount of aspirin in Cope (842 mg. compared to 650 mg. in a standard dose) necessarily provides increased analgesia. None of the studies submitted by respondent actually compare Cope's dose of aspirin with a standard dose. Furthermore, the Parkhouse Study discussed by respondent's witness, Dr. George Goldstein, demonstrated no statistically significant difference between aspirin dosages of 600 mg. and 1200 mg. (Tr. 15614) This finding is in accord with the conclusion of the FDA's OTC analgesics Panel which stated that:

> [T]here are no data available to show that multiple dosages greater than 650 mg. will provide any greater clinical benefit for analgesic and antipyretic effects. 42 FR 35364, July 8, 1977.

Thus, the mere fact that Cope has a greater than normal amount of aspirin does not establish its superiority.

Respondent has also submitted four clinical studies conducted on Cope. Only two of those compared Cope with other analgesics. (The other two compared different versions of Cope's formula with a placebo only and therefore cannot establish Cope's superiority.) Of the two remaining studies, the first compared Cope with aspirin. In this study the formulation of Cope differed slightly from the marketed version.

[38] A Cope tablet contains 421 mg. of aspirin, 32 mg. of caffeine, 50 mg. of magnesium hydroxide, 25 mg. of aluminum hydroxide gel, and 12.5 mg. of methapyrilene fumarate. (F. 799)

(Moertel, Tr. 6342) The study does not employ a placebo control so it is impossible to evaluate the sensitivity of the testing procedure. (Moertel, Tr. 6344) In addition, in the course of performing the study, the investigators changed their method of statistical analysis when they discovered that as originally designed their study would not demonstrate any difference between Cope and aspirin. Complaint counsel's witness, Dr. Moertel, an expert in analgesic testing, referred to this as "a gross and obvious example of statistical manipulation, and this is simply not acceptable scientific methodology." (Tr. 6345–46) Dr. Moertel noted that by changing methods of analysis, the investigators were able to generate results for some parameters that showed Cope to be superior to aspirin. [32] (Tr. 6348) However, because of the numerous flaws in the study, Dr. Moertel concluded that the study did not provide any evidence to establish that Cope is superior to aspirin and "certainly offers no evidence that it is superior to all other analgesics." (Tr. 6348)

Dr. Moertel also reviewed the other clinical study, which compared Cope with Anacin, and it was his conclusion that "this study offers very strong evidence that there is no difference at all between Cope and Anacin, and offers no evidence of any superiority of Cope over any other marketed analgesic." (Tr. 6349) He based this opinion upon the fact that this study contained the same flaws as the study comparing Cope with aspirin (different formulation tested, no placebo control, shift in analytical techniques). (Tr. 6349–50) Also, in this study, there was no significant difference for any of the parameters analyzed. (Tr. 6350)

Our examination of the ads in this case shows that respondent made claims of established superiority for Bayer and for Cope. The evidence shows that, in fact, relevant experts would not regard either product as having been established as superior. Therefore, respondent's advertised claims are false.

3. Bayer's superior pharmaceutical quality.

In addition to the claims of established superior efficacy which Sterling made for Bayer and Cope, Sterling also represented that it had tested Bayer against 220 other brands of aspirin and that this test demonstrated that Bayer was superior to the other brands with respect to quality, purity, freshness, stability, and speed of disintegration. Although grouped apart from the other establishment claims, these claims, alleged in paragraphs 20 and 21 of the complaint are, in fact, akin to establishment claims because the advertisements claim that Bayer's superior qualities were demonstrated by the test comparing it with 220 other brands. In the previous sections of this decision, we have found that experts require claims of superior effica-

cy to be substantiated with well-controlled clinical tests. However, Sterling contends that claims of superior manufacturing quality may be substantiated with various chemical and sensory tests.[39]

In 1968 Sterling conducted a test entitled, "Quality Comparison of Bayer Aspirin and Competitive Aspirin Products of the American Market," (the "223 Study") which compared Bayer with 220 other brands of aspirin and aspirin-based analgesics. In this test, 30 different **[33]** product attributes were tested by nonclinical means. Several of the attributes tested were related to purity, freshness, stability, and speed of disintegration. Sterling claims that this study substantiates its claims related to Bayer's quality. However, the ALJ concluded that this test did not demonstrate that Bayer was purer, fresher, more stable, or quicker to disintegrate than the other brands tested. Furthermore, he concluded that the test did not demonstrate that Bayer was superior in overall pharmaceutical quality to the other tested brands. From these conclusions respondent has appealed. (R.A.B. pp. 51–56)

Sterling argues that the ALJ's analysis of the 223 Study is invalid and that it is improper to examine specific pharmaceutical parameters in isolation. It contends that the 223 Study should be used only to make a judgment regarding overall pharmaceutical quality. (R.A.B. pp. 43, 56, 85–86) Although it may originally have been Sterling's intent to use the 223 Study only to reach a conclusion regarding pharmaceutical quality, its ads represent much more than that. Consumers could reasonably interpret the ads discussed above (pp. 8–14) to indicate that Bayer had been compared to 220 other brands and that the tests showed Bayer to be purest, freshest, most stable, and quickest to disintegrate.[40] Thus, it is valid to examine the individual attributes tested in the 223 Study to determine whether that evidence supports the specific claims made in the ads.

First, a facial examination of the evidence shows that Bayer is not quicker to disintegrate than all other tested brands of aspirin. Two of the thirty tests in the 223 Study involved speed of disintegration. The results (CX 430A, B) show that all tested samples of Bayer passed both tests. But so did all tested samples of at least 22 other brands.[41] Thus, with respect to speed of disintegration, it is impossible to conclude that the 223 Study shows Bayer to be superior to all other tested brands. In addition, Sterling also had in its possession several other

[39] As we noted above (*supra* pp. 11–12) claims of pharmaceutical (or manufacturing) quality are linked to and often imply therapeutic quality. However, if appropriately qualified, an ad may make a claim solely regarding such characteristics not directly related to therapeutic quality.

[40] Not every ad which mentioned a tested attribute necessarily implied that Bayer was superior with respect to that attribute. In some ads, attributes were mentioned only as examples of the tests performed to demonstrate Bayer's superior overall quality. *See* CX 48.

[41] Among the brands that disintegrated as rapidly as Bayer were McKesson, Norwich, Parke-Davis, Rexall, and St. Joseph. (CX 430 A)

studies of rates of tablet disintegration which show that some other brands disintegrate as rapidly as Bayer. (F. 634, 645)

The results of the 223 Study also do not show that Bayer is purer than all other tested brands of aspirin. Expert witnesses testified that an aspirin tablet may contain a veritable alphabet soup of impurities including FSA, ASAN, ASSA, and SSA (substances whose full names are free salicylic acid, aspirin anhydride, acetylsalicylsalicylic acid, and salicylisalicylic acid; *see* Rhodes, Tr. 11159; Falliers, Tr. 13346, 13361, 13363). [34] The 223 Study only tested for one of those impurities, FSA. Two of the thirty tests measured the level of FSA in the various brands. All of the tested samples of Bayer passed one of the FSA tests but one Bayer sample failed the second test. However, all tested samples of Parke-Davis and Safeway brands passed both of the FSA tests. Thus, the 223 Study does not show that Bayer is the purest of all brands tested.

Some Bayer ads also represented that Bayer was the most stable of all brands tested. According to the charts reporting the results of the 223 Study (CX 430A, B), five of the thirty tests measure product stability (*i.e.,* tendency to decompose). Since FSA is created when aspirin decomposes, the two tests measuring the presence of that impurity measure the extent of decomposition. Also relevant (according to the charts) are tests to determine if tablets are off-color, have acetic odor, and if the cotton wadding in the bottle has decomposed. The results of the 223 Study show that no Bayer samples had acetic odor, were off color or had decomposed wadding. However, as indicated above, the Bayer samples were unable to pass one of the tests for FSA content. The Parke-Davis sample tested passed all five tests. In addition, according to the testimony of respondent's witness, Jerome Winig, a chemist who was involved in the manufacture of Bayer aspirin for more than 40 years (I.D. pp. 50–51), acetic odor and off-white color may be caused by the manufacturing process and may not be an indication of decomposition. He testified that a test for the presence of FSA is a much more accurate measure of decomposition. (Tr. 14231, 14242) Examining test results for FSA levels, both Parke-Davis and Safeway brands performed better than Bayer. Thus, the 223 Study did not demonstrate that Bayer was the most stable brand tested.

Although Sterling's ads claimed that Bayer had been tested against 220 other brands of aspirin and been found the freshest, "freshness" does not appear to have been a product attribute specifically tested in the 223 Study. An examination of the 223 Study results shows that no test or group of tests is listed as a test of freshness. Indeed, Dr. Rhodes, respondent's expert in the formulation of drugs stated that "freshness" is not a technical term used in the industry and that it

is just an imprecise "layman's term." (Tr. 11441) Nonetheless, he believed that the 223 study did test for freshness because it tested "the odor of the tablet, the amount of FSA present, the appearance, the integrity of the seal." (Tr. 11442) Dr. Rhodes stated that in his opinion the 223 Study showed Bayer to be superior. As we indicated above, there are four tests related to FSA, tablet odor and appearance. Bayer and Safeway passed three of those tests and Parke-Davis passed all four. While one might dispute whether integrity of the seal directly relates to freshness, two of the thirty tests evaluated how well the tested products were sealed. Bayer, Safeway, and Parke-Davis [35] each passed and failed one test. Consequently, even according to the criteria which Dr. Rhodes used to determine freshness, Bayer appears to be no fresher than at least two other brands tested in the 223 Study.

Does the 223 Study demonstrate that Bayer is superior in overall pharmaceutical quality to the other 220 brands tested? That is the final issue presented by paragraphs 20 and 21 of the complaint. Sterling argues that Bayer performed better overall than any other brand in the 30 tests that composed the 223 Study. Since Bayer samples failed only five of the 30 tests and every other brand failed more, Sterling contends the study demonstrated that Bayer is the superior quality aspirin. The ALJ did not agree. He determined that, for several reasons, the 223 Study was inadequate to demonstrate Bayer's superior quality. First, the protocol was not adequate. (F. 1164) Second, there were inadequate records of the means used to collect and handle the samples tested. (F. 1166) Third, the study failed to control for the age of samples. (F. 1168) Fourth, a Sterling employee selected the parameters to be tested. (F. 1169) Fifth, all tests, including unblinded sensory tests, were conducted by Sterling employees. (F. 1170) Sixth, Bayer samples were transported by a different means than other samples. (F. 1174) Seventh, some of the Bayer samples came from Sterling's warehouse and not from the store shelves. (F. 1175) Eighth, some minor brands were grouped together in a somewhat arbitrary fashion. (F. 1177) Ninth, there were no tests for statistical significance of demonstrated differences. (F. 1179) In addition to these methodological flaws, the ALJ noted that Bayer was not superior to all other brands in all respects tested. (F. 1179) Thus, he determined that the 223 Study was not capable of demonstrating qualitative superiority.

Respondent objects to this conclusion and argues that the ALJ's reasoning was not supported by expert testimony. It notes that Sterling's experts uniformly agreed that the 223 Study showed Bayer to be qualitatively superior but that none of complaint counsel's experts testified to flaws in the study. (R.A.B. pp. 51–56) Complaint counsel concede that none of their experts testified regarding the 223 Study

but they contend that the study fails to meet standards required by relevant experts. Complaint counsel contend that these standards were derived from the expert testimony of Sterling's own witnesses. (C.An.B. p. 36) [36]

It would, of course, be possible for complaint counsel to meet its burden of proof on this issue by showing that the 223 Study did not live up to standards set forth by respondent's experts. Thus, it is not essential that complaint counsel present testimony from their own experts on every issue. However, after considering the arguments of counsel and the testimony, it does not appear to us that complaint counsel have been able to meet their burden of proof on this issue. For example, complaint counsel quote testimony from respondent's experts to show that nonclinical tests must be subjected to statistical analysis. (C.An.B. pp. 37–38) Specifically, they cite testimony from Drs. Horner, Feinstein, Danhof, and Banker. Yet, an examination of the testimony cited shows that in no instance does it truly apply to the 223 Study. Dr. Horner noted that in general it is important to test for statistical significance (Tr. 10850), but he did not indicate that the results of the 223 Study lacked statistical significance. Dr. Feinstein stated that in evaluating clinical studies, it is important to test for statistical significance. (Tr. 16428) But he did not indicate that the same sort of statistical analysis should be applied to the sorts of tests conducted in the 223 Study. The portion of Dr. Danhof's testimony cited by complaint counsel discussed a study of gastric bleeding in which no statistically significant differences had been shown. (Tr. 17241–42) Dr. Danhof did not mention the 223 Study. Finally, in his testimony, Dr. Banker stated that if a study shows a large difference, a test for statistical significance may not be necessary. Only when small differences are demonstrated must they be tested for statistical significance. (Tr. 12904–05) Dr. Banker does not indicate whether the differences demonstrated in the 223 Study were large or small.

Other testimony regarding other alleged flaws cited by complaint counsel also suffers the same problem—it is not clearly related to the 223 Study. The testimony adduced by complaint counsel from respondent's experts was of a general nature. None of those experts indicated that the 223 Study was flawed or that conclusions should not be drawn from it. Thus, although the shortcomings of the 223 Study might prevent it from demonstrating Bayer's superior quality, we do not think that complaint counsel have met their burden of proof on this issue. [37]

C. *The Substantial Question Issue.*

Paragraph 12 of the complaint restates four superiority claims

which respondent made regarding Cope and Vanquish.[42] Paragraph 13 alleges that a substantial question as to the validity of these claims exists among experts. Then paragraph 14 charges that even in those instances in which the ads did not indicate that the claims had been established, the failure to disclose the existence of the substantial question rendered the ads deceptive. The ALJ found the ads deceptive and entered an order provision which would require respondent to substantiate every comparative performance or freedom from side effects claim with a type and quantity of evidence necessary to establish the claim among experts. If that degree of substantiation was lacking, the ALJ's order would require respondent to disclose in the ad the existence of a substantial question as to the claim's validity. However, as our decision in *Bristol-Myers* indicates, slip op. at 38–44 [102 F.T.C. at 348–355], we no longer endorse the substantial question theory of liability. For reasons set forth in that decision we dismiss all the allegations of paragraphs 12–14.[43] **[38]**

D. *Lack of a Reasonable Basis*

Unlike *American Home Products* and *Bristol-Myers,* the complaint in this case alleges that for certain of its comparative performance claims, Sterling lacked a reasonable basis. Specifically, paragraph 10 of the complaint alleges that Sterling claimed Bayer is therapeutically superior to any other aspirin and that Bayer Children's Aspirin is therapeutically superior to any other children's aspirin. Paragraph 11 alleges Sterling lacked a reasonable basis for those claims. As we indicated above, our examination of the challenged advertisements in this case did not reveal any ad in which Sterling represented that Bayer Children's Aspirin was therapeutically superior to other brands (*supra* pp. 15–16). Furthermore, all advertisements which represent that Bayer is therapeutically superior to other brands also represent that Bayer is therapeutically superior to other brands also represent that Bayer's superiority has been established (*supra* pp. 12–13).[44] For that reason, we have considered the amount of support necessary to establish these claims among members of the relevant

[42] The four comparative claims are: (1) Cope is a more effective reliever of "nervous tension headache" than any other OTC internal analgesic; (2) Vanquish is a more effective pain reliever than aspirin or buffered aspirin; (3) Vanquish will result in less stomach discomfort than any unbuffered OTC internal analgesic; and (4) Vanquish is a more effective pain reliever than the largest selling "extra strength" tablet. As we discussed in part A, respondent's ads made all four of these claims.

[43] Application of the substantial question theory would require a conclusion that consumers expect the same level of support for a claim regardless of whether Sterling represents the claim has been established. Had complaint counsel presented evidence showing that consumers held such a belief—*i.e.,* that consumers believed that all superiority claims were scientifically established—then the failure to disclose the lack of evidence establishing such claims might well have been deceptive. However, as in *Bristol-Myers,* no evidence of consumer beliefs was presented here.

[44] We therefore find it unnecessary to decide whether the evidence Sterling had would have been sufficient to provide a reasonable basis for an unembellished superiority claim (*i.e.,* one not representing that superiority had been scientifically established).

scientific community. In determining whether respondent possessed a reasonable basis to support these very same claims, we must apply the same standard. According to *Pfizer,* one of the factors to consider in determining whether an advertiser possesses a reasonable basis to support a claim is the nature of the claim. 81 F.T.C. at 64. Since Sterling has represented in its ads that Bayer's superiority has been established, it could not have a good faith belief in the truth of that claim unless adequate evidence existed to establish Bayer's superiority. Thus, in this instance, the reasonable basis approach requires the same level of support as the establishment theory. Since we have already determined that Bayer's superiority was not established, we therefore also find that Sterling lacked a reasonable basis for this claim.

III. TENSION RELIEF CLAIMS

Paragrah 15 of the complaint in this case alleges that Sterling represented that Bayer, Cope and Midol relieve nervous tension, stress, fatigue and depression. Paragraph 16 charges that Sterling lacked a reasonable basis for these claims. The ALJ found that respondent made the claims without a reasonable basis and he entered an appropriate order provision. We agree with respect to the claims for Cope and Midol, but disagree with respect to the claims for Bayer. **[39]**

A. *The Advertisements.*

The ALJ cites ten ads which represent that Cope relieves tension.[45] Each of these ads presents Cope as a product especially designed to relieve nervous tension headache, and each ad indicates that Cope has an ingredient that relieves pain and an ingredient that relieves nervous tension. For example, CX 273 states:

> This is the most common headache there is, the nervous tension headache. As you can see, it's a two fold problem. Nervous tension and pain. For relief, try the two fold approach you get with Cope. Cope alone of all the leading headache remedies you can buy gives you a powerful pain reliever plus a proven relaxer. . . .

The implication of this ad is that a nervous tension headache is a combination of two problems, tension and headache. Cope is represented as having an ingredient to cure each of these problems. Thus consumers could reasonably infer that if they were tense, they could take Cope because one of Cope's ingredients is specifically intended to relieve nervous tension. The other nine ads cited by the ALJ make a similar representation.

[45] CX 272-276, 283, 287, 292-294.

The ALJ found that 17 ads for Midol represent that it relieves nervous tension. (F. 390) We agree with respect to 16 of those ads.[46] In all of these ads, Midol is presented as a product that relieves menstrual cramps. In addition to relieving cramps, each of the ads also represents that Midol can cure a number of other ills. For example, CX 296A states:

> Midol goes to work fast to help relieve a woman's discomforts. Like low backache, headache, calms jumpy nerves too. And the over all action of Midol chases the blues away.

CX 306Z035 indicates that Midol will "soothe irritability" and CX 306Z037 states that Midol contains "a special mood-brightener that gives you a real lift." In CX 306B, a woman is described as tense before she takes Midol and happy afterwards. All of these ads represent that Midol will relieve tension ("calm jumpy nerves," "soothe irritability") and most of them also portray [**40**] Midol as able to relieve depression because it is a "mood brightener." Thus, respondent did represent that Midol relieves tension, depression and stress.

Finally, the ALJ determined that five advertisements for Bayer represent that it will relieve tension (CX 29, 30, 33, 141, 151). Although it is a close call, we find that complaint counsel have not met their burden of showing that any of these ads represents that Bayer will relieve tension. Typical of these ads is CX 30. It is a television ad which shows a mother on a hot summer day attempting to tend numerous noisy neighborhood children in a backyard pool. As the scene progresses, the mother becomes more fatigued and irritable. The announcer states:

> As the day wears on, "Hot Weather Headache" can make you tense, irritable, out of sorts. And that's when Bayer works wonders. . . . Take two Bayer tablets and put your feet up. In just minutes, headache's gone.

At the close of the ad, the mother appears refreshed and happy. Although this ad does mention tension and does depict a potentially tension provoking scene, the message of the ad appears to be that the hot, hectic day gives the mother a headache and the headache, in turn, makes her tense. The ad then states that Bayer will relieve the headache. The implication is that headache relief causes the tension to disappear.

We considered a somewhat similar ad in *Bristol-Myers*. In that case, CX 53 depicted a tense confrontation between a student and a college dean. The audio portion of the ad stated that Bufferin provided relief from headache. We ruled that CX 53 did represent that the product

[46] CX 296A, 297–300, 306, 306A-C, 306R, 306Z005, 306Z011, 306Z035, 306Z037, 306Z041, 306Z045, 306Z053.

would cure tension, but in that case we had copy tests showing that more than 50% of viewers received the impression Bufferin relieved tension. In this case, since the claim is not apparent to us from a careful examination of the ads, we are unwilling in the absence of extrinsic evidence to find that consumers infer from these ads that Bayer will relieve tension. No extrinsic evidence was presented and therefore we find that Sterling did not represent that Bayer relieves tension.

B. *Evidence on Tension Relief.*

Cope and Midol both contain aspirin. In addition, Cope contains buffers, caffeine and an antihistamine, methapyrilene fumarate. Midol contains (in addition to aspirin) caffeine and an antispasmodic, cinnamedrine hydrochloride. Sterling has presented no evidence indicating that caffeine, buffers or [41] cinnamedrine hydrochloride produce any tension relieving effect and there is expert testimony confirming that none of these three ingredients relieves tension. (Rickels, Tr. 7974, 8019, 8021). Indeed, caffeine, an ingredient in both Cope and Midol is contraindicated for the relief of tension. (Rickels, Tr. 7974) However, respondent did present evidence that aspirin and methapyrilene fumarate have tension relieving properties, and it is to that evidence we now turn.

The evidence presented by Sterling regarding its tension-relief claims must be considered in light of the factors set forth in *Pfizer* for determining what constitutes a reasonable basis. The ads at issue here advise consumers to take aspirin-based analgesics for the relief of tension. If the products cannot provide that relief, then consumers may forego effective remedies and are needlessly being encouraged to consume aspirin, a drug with potentially hazardous side effects. (*see* infra p. 47) Furthermore, as with other performance claims regarding mild analgesics, it is virtually impossible for consumers to verify for themselves whether the product can relieve tension. All of these considerations are relevant to determining whether Sterling's evidence constitutes a reasonable basis for its claim.[47]

As we mentioned in *Bristol-Myers,* tension can be caused or exacerbated by headache pain. Since aspirin can relieve a headache, it could relieve or lessen the tension caused by a headache. However, respondents' ads represented that Cope and Midol could relieve tension which exists independent of headache pain. Thus, the mere fact that aspirin relieves pain does not by itself support claims that aspirin has tension-

[47] Sterling appears to argue (in connection with its claims for Bayer) that the Commission is precluded from anything more than a superficial determination of a scientific test's adequacy. We disagree and will, in appropriate situations, consider evidence regarding the details of scientific tests in order to determine whether the advertiser's reliance thereon was reasonable. *See Bristol-Myers,* slip op. at 32–37 [102 F.T.C. at 343–347]; *Porter & Dietsch,* 90 F.T.C. at 870–871; *Firestone,* 81 F.T.C. at 445–449.

relieving properties. (Rickels, Tr. 8102–03)[48] [42]

To support its claim that aspirin can relieve tension, Sterling has relied on various studies and reports in medical literature. The first item was a 1965 study by Krumholtz and Merlis. This study was also submitted into evidence in *Bristol-Myers,* and once again, Dr. Rickels, complaint counsel's expert on psychopharmocology and tension, criticized flaws in the study, especially the failure of the authors to identify the symptoms of the test population. It is therefore possible that changes experienced by the test subjects resulted from the analgesic effects of aspirin. (Tr. 8115–16) Furthermore, the study was designed to measure aspirin's effect on depression rather than tension, and the authors recognized that additional study was necessary to test the tranquilizing action of aspirin. (Goldstein, Tr. 17977)

Much of the other evidence presented regarding aspirin related to its ability to overcome insomnia. However, sleeplessness can be caused by many factors other than tension. (L. Goldstein, Tr. 17900) Thus, a mere showing that aspirin has hypnotic (sleep-inducing) effects does not constitute a reasonable basis for tension relief claims unless it is shown that the sleeplessness was caused by tension. Sterling submitted a report of a 1959 study by Boyd, *et al.,* which reported that buffered aspirin had hypnotic effects. However, the record does not show that any of the insomniacs who participated in the study suffered from tension. Furthermore, the record shows that a substantial majority of subjects in Boyd's study were receiving other medication including barbiturates (a sleep-inducing drug). (Goldstein, Tr. 17981) Finally, Dr. Rickels noted that the study had methodological flaws. (Tr. 8178)

Sterling also submitted three reports from medical literature and a chapter from a 1969 textbook supporting aspirin's efficacy as a tension reliever. The three reports refer to potential sedative effects of aspirin, not to tension relief. (Goldstein, Tr. 17849, 17851, 17852) Thus, these reports do not make clear whether aspirin relieves sleeplessness caused by tension or whether aspirin relieves pain thereby permitting the subject to sleep. (Goldstein, Tr. 17983–84) References to medical texts, without more, are not given much weight by scientists. Normally, they look to the underlying data and not to the text. (Rickels, Tr. 7978) Indeed, neither the textbook reference nor the three reports contain any underlying data.

Sterling additionally offered several electroencephalagram ("EEG") studies to support its tension-relief claims. However, Dr. Rickels noted that because EEG studies are not able to directly meas-

[48] Dr. Rickels gave the following analogy to illustrate the point: A person with a bladder infection may have to urinate frequently during the night and be unable to sleep the whole night through. An antibiotic which cures the infection would permit the person to sleep the entire night but the antibiotic would not, therefore, be considered a sleeping aid.

ure a drug's ability to relieve tension, they only give a preliminary indication of what a drug can do. (Rickels, [43] Tr. 7970, 8184–85) Indeed, in some instances EEG's have indicated that a change was taking place in a subject when clinical studies have been unable to show any change. (Rickels, Tr. 8185) In addition, EEG results can be misleading because of variability among test subjects. (Goldstein, Tr. 17959–61) Respondent's witness Dr. Goldstein criticized four EEG studies submitted by Sterling (including three studies which he had conducted) based upon the failure to take variability into account. He indicated that by failing to take variability into account, he "was mixing apples and oranges and expressing their averages in terms of bananas." (Tr. 17959) The only other EEG study was conducted in 1978, long after Sterling had made the claims at issue in this case. Although the authors concluded that aspirin did improve the sleep of insomniacs, Dr. Goldstein noted that there were individual differences among the eight test subjects and that two received no benefit at all. He concluded that this was caused by the fact that the test subjects were insomniacs and there was no indication as to what the cause of each patient's insomnia was. (Tr. 17900) Finally, Dr. Goldstein conceded that the FDA would not rely solely on EEG studies to demonstrate the sedative property of a drug. (Tr. 17987)[49]

In addition to the evidence submitted by respondent, complaint counsel submitted evidence on two studies which show that aspirin does *not* relieve tension. The first was a study conducted by complaint counsel's expert Dr. Rickels in 1971 and it showed that a 500 mg. dose of aspirin (the normal dose is 650 mg.) was no more effective as a tension reliever than a placebo. (Rickels, Tr. 7951) The second study was also well-controlled, and it tested a normal dose of aspirin. It was the authors' conclusion that aspirin was not able to relieve tension. (Rickels, Tr. 8195) These studies indicate that at the time respondent was making tension relief claims for Cope and Midol, it was possible to conduct well-controlled studies measuring tension relieving capacity and that such studies showed aspirin was not a tension reliever. Furthermore, a letter from Sterling's files (CX 358) shows that as of May 1969 (prior to making most of the tension relief claims), Sterling was well aware that aspirin would not relieve tension. Thus, respondent did not have a reasonable basis for believing aspirin would relieve tension. [44]

Methapyrilene fumarate, an ingredient in Cope, is an antihistamine (a substance which combats infection). Sterling presented a variety of evidence (but no expert testimony) regarding methapyrilene

[49] In addition to the tests and studies discussed above, Sterling presented evidence designed to show that ingestion of aspirin increases the body's level of tryptophan, an amino acid which must be present at certain levels for sleep to begin. However, as explained above, producing sleep is not the same as reducing tension and thus, this evidence does not support tension relief claims regarding Cope and Midol.

(including two clinical studies). However, none of this evidence provides a reasonable basis to substantiate the tension relief claims for Cope. First, most of the participants in the two clinical studies had headaches. (F. 938) By not separating those participants who had only tension from those who had pain, it is impossible to conclude that Cope had tension relieving properties separate from its analgesic properties. (Rickels, Tr. 8007) Since the Cope ads represent that Cope can relieve this "free-floating" tension, these tests do not support those claims. The journal excerpts submitted by Sterling (F. 945) suffer from the same flaw.

Second, since methapyrilene fumarate is an antihistamine, it is also a hypnotic. However, as explained above, inducing sleep is not the same as relieving tension. Indeed, drowsiness would be undesirable, even dangerous, when produced as a side effect of a drug taken during the day. (Rickles, Tr. 8183) For that reason, eight studies submitted by Sterling demonstrating methapyrilene's sleep inducing properties do not support tension relief claims. The same is true of the bibliographic material listed in F. 946.

Finally, Sterling relied on numerous works of Dr. Arnold Friedman. Dr. George Goldstein,[50] a Sterling employee, said these articles implied that an ingredient with sedative properties is appropriate to treat tension. (Tr. 15508) However, Dr. Friedman's writings concerned the effects produced by a combination of an analgesic and a barbiturate and as Dr. Rickels, an expert in pharmacology, indicated, it would not be proper to draw any conclusions about an antihistamine (such as methapyrilene fumarate) from data regarding a drug containing a barbiturate. (Rickels, Tr. 8016)

Although respondent has presented some evidence to show that Midol will brighten a user's mood, it presented no such evidence to the FDA. In its submission, it claimed only that the ingredients in Midol would relieve menstrual pain. (George Goldstein, Tr. 15603) However, in this proceeding, respondent has claimed that the caffeine in Midol acts as a mood brightener. The claim is based in part on the testimony of Dr. George Goldstein. Dr. Tainter, another Sterling employee, disagreed and indicated that the presence of caffeine in Midol might heighten the user's pain and would, in any event, be too small a dose to affect the user's mood. (CX 417B) Also, Drs. Goodman and Gilman, who were recognized by respondent's witness Dr. George Goldstein as "among the leading lights of American pharmacology" (George Goldstein, Tr. 15590) do not list brighter mood among the [**45**] effects of a therapeutic dose of caffeine (George Goldstein, Tr. 15593–94). Final-

[50] Two Drs. Goldstein testified for respondents. Up to this point, all references have been to the testimony of Dr. Leonide Goldstein, an expert in the biological basis of human behavior (F. 121). Dr. George Goldstein was qualified as an expert only in the use of Sterling's products.

ly, the rest of the evidence presented by Sterling makes it appear that the amount of caffeine in Midol is too small to have any therapeutic effect. (Rickels, Tr. 7974; George Goldstein, Tr. 15587, 15592)

The evidence presented by respondent thus is inadequate to substantiate the mood altering claims it made for Cope and Midol. None of the evidence which Sterling presented regarding aspirin separates its unquestioned analgesia from any tension-relieving effect it may possess. Other evidence (such as EEG studies and the tryptophan theory) is inconclusive, especially in light of the well-controlled studies presented by complaint counsel which show that aspirin relieves tension no more effectively than a placebo, and the letter from Sterling's files showing that it knew substantiation did not exist for the claim that aspirin relieved tension. The evidence presented regarding methapyrilene fumarate is either not helpful because any tension-relieving effect Cope might produce was not isolated from Cope's analgesic effect, or of questionable relevance (*i.e.,* the studies regarding barbiturates). The evidence regarding the caffeine in Midol does not clearly show that caffeine can brighten one's mood and also seems to indicate that under any circumstance, Midol does not contain enough caffeine to have any effect. Thus, in light of the claims made by respondent for its products, in light of the testimony, we conclude that respondent did not possess a reasonable basis for its claims that Cope and Midol can affect a user's mood.

IV. FAILURE TO DISCLOSE THE PRESENCE OF ASPIRIN; REPRESENTATION THAT MIDOL CONTAINS OTHER THAN ORDINARY ASPIRIN

Paragraphs 23 through 25 of the complaint charge that Sterling failed to disclose in its advertising that Vanquish, Cope, and Midol contain aspirin. An examination of the challenged advertisements for these products shows that none disclose that aspirin is an ingredient. However, as we explained in *Bristol-Myers,* slip op. at 54 [102 F.T.C. at 361], we are unprepared to hold that the mere failure to disclose the presence of aspirin in advertising for aspirin-based analgesics renders that advertising materially misleading. In *Bristol-Myers,* we found that advertisements for Bufferin and Excedrin created the impression that those products did *not* contain aspirin. In light of that advertising, we held that the failure to disclose the presence of aspirin was materially misleading. In this case, there are no allegations that Sterling's advertising created the impression that Vanquish and Cope do not contain aspirin. For that reason, we find that the failure to disclose the presence of aspirin in Vanquish and Cope does not violate the FTC Act and we dismiss all charges related to Paragraphs 23 and 24. On the other hand, Paragraph 26 alleges that advertisments for Midol represented that it does not contain aspirin and caffeine. These

representations would make the failure to disclose the presence of aspirin a violation of the FTC Act. [46]

The only active ingredients in Midol are aspirin, caffeine, and an anti-spasmodic, cinnamedrine hydrochloride. Nonetheless, numerous ads for Midol do create the impression that Midol does not contain aspirin. This impression is created by ads which state that Midol and its ingredients are special and out-of-the-ordinary. For example, CX 302 states:

> Midol starts to work fast with an exclusive formula that helps stop periodic pain . . . and its medically approved ingredients gives effective relief from headache and low backache. All in all, Midol's unique formula gets you through those days in comfort.

The impression created by this ad is that Midol, and everything about it, is special and different. Its formula is described as "unique" and "exclusive." In the midst of language that creates an aura of uniqueness about the product, the ad states that Midol has a "medically-approved ingredient" to relieve pain. Consumers could reasonably infer that the "medically-approved ingredient," a part of the unique formula, is also unique. As respondent concedes, however, the "medically-approved" pain reliever is aspirin (Hartman, Tr. 9166), an ingredient," which is anything but unique or special (*see American Home Products,* 98 F.T.C. at 362), an ingredient which is familiar to most consumers. It would have been a simple matter for respondent to clarify its ads so that consumers would realize that it is the anti-spasmodic ingredient which is unique to Midol.

In another ad Midol is contrasted with "ordinary pain relievers." The ad (CX 296B) states, "An exclusive formula with medication ordinary pain relievers don't give you, Midol relieves the pain of backache, headache and other discomforts mature women can get." Although respondent contends that no other product has exactly the same formula as Midol and that ordinary pain relievers do not contain cinnamedrine hydrochloride, the ad has blurred the distinction between ingredients that relieve pain and those that perform another function. The impression created thereby is that the pain relieving medication in Midol is not contained in ordinary pain relievers. Aspirin is a common (*i.e.,* ordinary) pain reliever. Therefore, CX 296B creates the false impression that Midol does not contain aspirin.[51]

The ads cited above could reasonably mislead consumers into believing that Midol does not contain aspirin. A misleading claim or omission will violate the FTC Act only if the omitted information (in this case, that Midol contains aspirin) would be a material factor in

[51] Similar ads are CX 297, 300. On the other hand, CX 305 indicates that it is the anti-spasmodic that is not contained in ordinary pain relievers and CX 306 makes it clear that it is the anti-spasmodic that is exclusive.

a consumer's purchasing decision. *F.T.C.* v. *Colgate-Palmolive Co.*, 380 U.S. at 392. Section 15 of the FTC Act provides that an omission of fact may be material "in the light of representations made or suggested . . . or . . . with respect to [47] consequences which may result from the use" of the product. The failure to disclose the presence of aspirin is material in light of "representations made or suggested" in respondent's ads which create the impression that the pain reliever in Midol is unique. It is this false impression which would lead a consumer to look no further for a non-aspirin-based analgesic. It is this same false representation of uniqueness which would discourage a consumer from looking for a less expensive analgesic.

Expert testimony in this case explains one reason why consumers look for non-aspirin-based pain relievers. Aspirin may cause numerous side effects. According to Dr. Grossman, an expert in gastroenterology, aspirin may cause dyspepsia and gastrointestinal bleeding, and it may exacerbate or even cause ulcers. (Tr. 7471, 7479, 7720, 7722) According to Dr. Stevenson, an expert in asthma and immunology, aspirin, even in minute doses, can cause asthmatics to suffer attacks which may be severe or even life-threatening. (Tr. 1480, 1489) In addition, aspirin can cause skin reactions such as hives or swelling. (Stevenson, Tr. 1511–12)

Respondent argues that the ALJ greatly inflated the number of consumers who would suffer adverse reactions from aspirin. (Sterling Reply Brief, pp. 48–50) However, as we found in *American Home Products*, 98 F.T.C. at 367, the number of consumers who suffer adverse reactions to aspirin is significant. Immunologists generally warn all asthmatics to avoid aspirin, regardless of whether they are known to be aspirin-sensitive (Farr, Tr. 2606) and respondent concedes that there are at least two to six million asthmatics in the United States (Sterling Reply Brief, p. 50). Thus, some consumers avoid aspirin for medical reasons. For them, Sterling's failure to disclose aspirin's presence is material in the context of ads which create the impression Midol is aspirin-free.

The nondisclosure of aspirin is also material for economic reasons. At the time Sterling disseminated the challenged Midol advertising, Midol was substantially more expensive than most other aspirin-based analgesics. In fact, the *wholesale* price of Midol was twice as expensive as the *retail* price of Bayer and more than four times as expensive as the average retail price for non-Bayer aspirin. (F. 6, 11) Since Midol ads create the false impression that its pain reliever is unique and, therefore, not available in other products, consumers are not as likely to consider less expensive aspirin-based analgesics available on the market as potential alternatives to Midol. Thus, failure

to disclose aspirin's presence is material in the context of ads which imply that the analgesic in Midol is available in no other product. [48]

We also find that consumers are not already aware of the ingredients in the analgesics which they use. This lack of knowledge is demonstrated by several studies in the record including an informal study conducted by Dr. Moertel (Tr. 6355–60) and a 1970 Analgesic Segmentation Study (CX 394) performed at Sterling's request, which showed that 82% of those surveyed were unable to name any of the ingredients in the brand-name headache remedy they normally use. Sterling argues that the results of these studies do not prove that consumers are unaware of the contents of analgesics because the results are not projectible to the general population. (Sterling Reply Brief, p. 49) While it is true that these survey results may not be statistically projectible, they at least suggest that consumers are unaware of the ingredients in analgesics. (Pernica, Tr. 1998) In light of the fact that respondent has offered no contrary evidence, we are unwilling to conclude that consumers generally know which analgesics contain aspirin.

We stress that we find a violation of the FTC Act only in those instances in which Sterling affirmatively represented (either expressly or by implication) in its advertising that Midol did not contain aspirin. As we indicated in *Bristol-Myers,* slip op. at 54 [102 F.T.C. at 361], we do not find that the mere failure to disclose the presence of aspirin in advertising for aspirin-based analgesics is misleading. And, indeed, there are numerous ads in the record which do not misrepresent Midol's (or Cope's, or Vanquish's) aspirin content. These ads are silent on the subject. After viewing ads such as these, consumers would not necessarily know (based upon the ads) whether the advertised product contained aspirin. If consumers viewing these ads were concerned about aspirin content, they would have to look elsewhere for the information, but at least they would not be discouraged from doing so. On the other hand, in the context of the affirmative claims made in the Midol ads, the failure to disclose the presence of aspirin is a material omission of fact which renders the advertisement false.[52] [49]

V. COPE'S UNIQUE FORMULA

Paragraph 22 of the complaint charges that respondent falsely represented that Cope's formula is unique. The ALJ found that this representation had been made, and we agree. For example, CX 274 states:

[52] Paragraph 26 of the complaint also alleges that Sterling represented that the stimulant in Midol is not caffeine. However, the record does not contain any evidence indicating that knowledge of the presence of caffeine in an OTC analgesic is material to consumers, the ALJ made no findings on that issue, and complaint counsel have not appealed that point. Therefore, we dismiss that portion of paragraph 26 that relates to caffeine.

Cope looks different, is different. Besides a powerful pain reliever, Cope gives you a gentle relaxer. The others don't. . . . A unique formula for really effective relief of nervous tension headache. And you get it only in Cope.

CX 275 also represents that the Cope formula is "unique." These ads were shown to the public from December 1969 through March 1970. (CX 633)

The Cope ads mentioned above state that the Cope formula is unique because only Cope contains both "a powerful pain reliever" and also "a gentle relaxer." The clear message is that no other analgesic contains both types of ingredients. The ingredients in Cope are aspirin, caffeine, methapyrilene fumarate, and two antacids. Thus, the "powerful pain reliever" is aspirin and the "gentle relaxer" is methapyrilene fumarate, an antihistamine which, as explained above (*supra* p. 44) produces drowsiness. At the same time as Sterling was advertising Cope, Bristol-Myers was promoting Excedrin P.M., an OTC analgesic which contained, among other things, both aspirin and methapyrilene fumarate. (CX 357B; *see also Bristol-Myers,* slip op. at 50. [102 F.T.C. at 358]) Thus, Excedrin P.M. also contained the "powerful pain reliever" and the "gentle relaxer." Furthermore, the evidence shows that at the time Sterling was advertising that only Cope contained a "powerful pain reliever" and "a gentle relaxer," it was aware that Bristol-Myers was marketing Excedrin P.M. (CX 678 Admission 1069) and it was aware that Excedrin P.M. contained aspirin and methapyrilene fumarate. (CX 357) Therefore, advertisements that portrayed Cope as unique were false and misleading. As explained in Part IV above, such false representations of uniqueness are material because they discourage consumers from shopping for potentially less expensive alternatives.

VI. INCONSISTENT CLAIMS

Paragraph 17 of the complaint charges that respondent made several mutually inconsistent claims for its products. Specifically, it charges that respondent claimed in its ads that: (1) Bayer is as effective for the relief of headache pain as any **[50]** other OTC analgesic, and Vanquish is a more effective reliever of headache pain than aspirin; (2) Bayer will cause stomach upset no more frequently than any other OTC analgesic, and Vanquish will cause less stomach upset than unbuffered aspirin; and (3) Bayer is as effective a reliever of nervous tension headache as any OTC analgesic, and Cope is more effective for the relief of nervous tension headache than any other OTC analgesic. Paragraph 29 charges that the making of inconsistent contemporaneous claims is unfair or deceptive. The ALJ disagreed and dismissed paragraph 17 and that portion of paragraph 29 that

relates to inconsistent claims. From this determination, complaint counsel have appealed.

Complaint counsel's argument is based upon the Commission's decision in *Pfizer.* They argue that in order to substantiate advertising claims *Pfizer* requires an advertiser to possess a reasonable basis consisting of evidence that "would satisfy a reasonable and prudent businessman, acting in good faith, that such representation is true." 81 F.T.C. at 64, quoted at C.A.B. p. 64. This standard, complaint counsel contend, precludes the making of inconsistent claims because, having made a claim for one product, no advertiser could reasonably make an inconsistent claim for a competing product which it also sold to the public. Reasonableness, complaint counsel argue, cannot be judged solely in terms of the quantity of evidence supporting a given claim. (C.A.B. p. 65) Such conduct would be both unfair, because it places on the consumer the burden of taking the gamble as to which claim is correct, and deceptive, because the making of a claim implies the existence of a reasonable basis.

Finally, complaint counsel argue that on numerous occasions in the past the Commission has determined that the making of mutually inconsistent claims constituted a violation of the F.T.C. Act. To support this, they cite cases such as *Rudolph R. Siebert,* 49 F.T.C. 1418 (1953); *Montgomery Ward & Co., Inc.,* 70 F.T.C. 52 (1966), *aff'd,* 379 F.2d 666 (7th Cir. 1967); and *Sears, Roebuck and Co.,* 95 F.T.C. 406 (1980), *aff'd,* 676 F.2d 3985 (9th Cir. 1982).

After carefully considering complaint counsel's arguments, we conclude that it would be inappropriate for us to find that respondent violated the F.T.C. Act merely because it made inconsistent advertising claims. We believe the inconsistent claims theory would be a new theory of advertising substantiation which would shortcut and be contrary to principles of law set forth in *Pfizer* and its progeny. Thus, for the reasons set forth below, we agree with the ALJ that paragraph 17 and that portion of paragraph 29 of the complaint that relate to inconsistent claims should be dismissed. **[51]**

Complaint counsel argue that *Pfizer* requires not only an adequate quantity of support to substantiate claims but also support of a type that would satisfy a reasonable businessman that the claim is true. Inconsistent claims, complaint counsel contend, could never be reasonably substantiated because a reasonable businessman would never believe that two inconsistent claims were both true. The reasonable businessman standard quoted by complaint counsel actually comes from *H.W. Kirchner,* 63 F.T.C. 1282, 1294, and was restated by the Commission in *Pfizer,* 81 F.T.C. at 64. After quoting the substantiation standard from *Kirchner,* the opinion in Pfizer makes it clear that the *Pfizer* "test evaluates both the reasonableness of an advertiser's

actions and the adequacy of the evidence upon which such actions were based." 81 F.T.C. at 64. Thus, the reasonable basis standard in *Pfizer* subsumes the *Kirchner* standard. Each individual advertising claim alleged in paragraph 17 could have been evaluated under the *Pfizer* standard. The performance of this analysis would have determined whether Sterling's reliance on the substantiation it possessed was reasonable. However, the reasonable basis analysis does not determine whether a claim is true. *See Pfizer,* 81 F.T.C. at 67 n. 22. Therefore, it is at least theoretically possible that two inconsistent claims could both be substantiated with a reasonable basis. In effect, the approach recommended by complaint counsel would bypass the *Pfizer* analysis in favor of a rule finding liability based solely upon the wording of advertising, regardless of the substantiation possessed by the advertiser.

The inconsistent claim theory would also produce an anomalous result described in respondent's answering brief. (R.An.B. p. 54) If an advertiser made inconsistent claims regarding products A and B, the advertiser would have automatically violated the F.T.C. Act. If, however, one advertiser made the claim regarding product A and a competing advertiser made the inconsistent claim regarding product B, the claims would be judged under a reasonable basis standard and it is possible that neither advertiser would be found to have violated the law. By applying the reasonable basis standard to all claims, regardless of the advertiser, the above result can be avoided.

Finally, we find that the previous Commission cases cited by complaint counsel do not actually apply the inconsistent claim theory proposed by complaint counsel. A good example is the *Sears* case. In that case, Sears advertised that no pre-soaking was necessary prior to washing dishes in the Lady Kenmore dishwasher. The owner's manual, however, said just the opposite. The Commission held that Sears lacked a reasonable basis for its claim and entered an order which, among other things, prohibited Sears from making any claims in its advertising that were contradicted in an owner's manual given to a consumer after purchase. Complaint counsel contend that this is an example of the inconsistent claim theory. **[52]** However, the Commission (and the ALJ) in *Sears* carefully examined the substantiation possessed by Sears and determined that it did not have a reasonable basis for its claim. 95 F.T.C. at 426–469, 514. The fact that the owner's manual contained inconsistent statements was merely a factor considered in determining whether the advertising was adequately substantiated. The analysis performed in *Sears* was not the abbreviated inconsistent claim analysis proposed by complaint counsel.

Thus, we find that the inconsistent claim theory is not appropriate for analyzing advertising substantiation. The advertising claims set

forth in paragraph 17 of the complaint could have been each individually challenged for lacking a reasonable basis under *Pfizer*. Since this was not done, we agree with the ALJ that complaint paragraph 17 and that portion of paragraph 29 relating to inconsistent claims should be dismissed.[53] **[53]**

VII. LIABILITY OF LOIS HOLLAND CALLAWAY

The complaint in this case charges the advertising agency, Lois Holland Callaway, Inc., with responsibility for certain advertising claims regarding Vanquish. We dismiss each of these charges since we have ruled that the ads in question were not deceptive.

Specifically, paragraph 8 charges the agency with falsely representing it has been established that Vanquish is more effective than aspirin and less likely to cause stomach upset than unbuffered analgesics. However, as we indicated above, (*supra* pp. 16–19) the challenged advertising contains no establishment representations regarding Vanquish. In paragraph 12, Lois Holland Callaway is charged with representing that Vanquish is a more effective pain reliever and less likely to cause stomach upset, without disclosing that a substantial question existed among experts as to the validity of these claims. For reasons explained above (*supra* p. 37), we dismissed all claims based on the substantial question theory. In paragraph 23, Lois Holland Callaway is charged with failing to disclose in advertising the presence of aspirin and caffeine in Vanquish. As we explained above (*supra* p. 45), this would constitute a violation of the FTC Act only if the advertising also represented or implied that Vanquish did not contain aspirin. We found that Vanquish advertising created no such implication. We therefore have dismissed all unfair or deceptive advertising claims with which Lois Holland Callaway has been charged.

---

[53] Sterling has also raised objections to several of the ALJ's evidentiary rulings. First, it objects because the ALJ refused to admit into evidence approximately 160 scientific articles. (R.A.B. pp. 71–74) We find that this decision was an appropriate exercise of the ALJ's duty to manage a complex lawsuit. Although 160 articles were excluded, approximately 60 were accepted and the ALJ gave respondent the opportunity to select which articles would be admitted. (Tr. 11937–38, 18055) Furthermore, the ALJ permitted respondent's experts to quote from any of the articles and to read any passage into the record. (Tr. 11938)

Sterling also appealed the ALJ's exclusion of six unpublished scientific studies from the record. (R.A.B. 74–77) Two of the studies, RX 195 and 207, involve nonclinical tests submitted by Sterling to show it possessed a reasonable basis for superiority claims regarding Bayer. However, as we explained above (*supra* p. 38), because these claims were embellished with representations of establishment, only well-controlled clinical tests can constitute a reasonable basis for Bayer's superiority. Respondent was not harmed by the ALJ's rejection of RX 415 because that exhibit was submitted to support the claims of Bayer's superior quality and we held that complaint counsel failed to meet their burden of proof on that issue. RX 197 was rejected as being duplicative of other evidence on the record. The ALJ did, nonetheless, permit testimony regarding the study (*e.g.* Fields, Tr. 16758–61). There is no evidence that respondent was prejudiced by this evidentiary decision. RX 190 was rejected by the ALJ because it was unpublished, the author was reporting on a study by someone else, and Sterling called no foundation witness. Without a proper showing of reliability, rejection of RX 190 was proper. *Bristol-Myers*, 85 F.T.C. 688, 743–744 (1975). Finally, respondent was not prejudiced by the rejection of RX 422. The record shows that RX 422 was not complete (Tr. 17921–27) and, although given the opportunity to have the study's author testify (Tr. 15082), respondent chose not to call him. Nonetheless, the ALJ did admit an abstract of the complete study. (Tr. 17926–27) Thus, there was no error in the rejection of RX 422.

VIII. RELIEF

The order we enter today proscribes the violations of the FTC Act committed by respondent. It also prohibits related violations in order to assure respondent's future compliance. *F.T.C.* v. *Ruberoid Co.*, 343 U.S. 470, 473 (1952); *American Home Products*, 98 F.T.C. at 398.

However, there are differences between our order and the one entered by the ALJ. First, we dismissed those allegations of the complaint based upon the substantial question theory of advertising substantiation. The ALJ's order would have required respondent either to substantiate all comparative efficacy claims with well-controlled clinical tests, or to disclose in ads making those claims that the claims had not been proven. Our order imposes the clinical testing requirement only on those ads which claim that an analgesic's superiority has been established.

Second, the ALJ's order required that respondent possess a reasonable basis for any claim it makes regarding any nonprescription drug. We decline to enter so broad an order provision. However, since we have found instances in which respondent lacked an appropriate level of substantiation for both comparative and noncomparative therapeutic performance claims, our order requires Sterling to possess a reasonable basis for any therapeutic performance claim it makes regarding an OTC internal analgesic. [54]

The ALJ also entered an order provision imposing specific substantiation requirements for all claims regarding pharmaceutical quality. (I.D. pp. 353–354) As we discussed above, we found that complaint counsel failed to meet its burden of showing that respondent lacked adequate substantiation for its claims regarding Bayer's overall pharmaceutical quality. Accordingly, we have limited this order provision to claims regarding specific product attributes. These were the only claims related to pharmaceutical quality which complaint counsel showed respondent failed to substantiate. Additionally, our order narrows the scope of the aspirin disclosure required by the ALJ, limiting the disclosure of aspirin's presence to those ads for analgesics which contrast the advertised product with aspirin. Finally, we have limited the product coverage of some of the order provisions. Each of these modifications is discussed below.

A. *Establishment Claims.*

Paragraph I of the order sets forth the level of support which Sterling must possess before it can advertise that the superior effectiveness of any nonprescription internal analgesic product has been established. These types of claims must be supported by two well-controlled clinical studies meeting the criteria set forth in subpara-

graphs A-C of paragraph I. Testimony in this case shows that experts require studies to meet these criteria in order to establish an analgesic's superior efficacy. (*Supra* pp. 21–32) Indeed, we imposed the same testing requirement in *American Home Products* and *Bristol-Myers* based upon the expert testimony elicited in those cases. However, as in *Bristol-Myers,* slip op. at 67 [102 F.T.C. at 372], we have included paragraph I D in order to avoid penalizing Sterling for purely technical instances of noncompliance with the detailed provisions of paragraph I, if Sterling can show that the scientific community would not regard the violation as affecting the adequacy of support for the claims.

The ALJ's order applied the clinical testing requirement to establishment claims made by Sterling regarding any nonprescription internal analgesic product and we agree with this product coverage. Complaint counsel have argued that the clinical testing requirement should apply to establishment claims regarding all OTC drugs. (C.A.B. pp. 27–35) However, we rejected similar arguments in both *Bristol-Myers* and *American Home Products.* As we held in *American Home Products,* 98 F.T.C. at 402–403, it is possible that claims of superiority for other drug products may be established by other than two well-controlled clinical tests. Indeed, there is testimony in this case that nonclinical tests can establish the efficacy of antacids. (Scoville, **[55]** Tr. 14476–81) Thus, it would not be appropriate for us to apply paragraph I of the order to all OTC drugs and we, therefore, limit its applicability to OTC internal analgesics.

On the other hand, we do not believe that application of this part of the order should be limited to the specific brands involved in this case, *i.e.,* merely to Bayer and Cope. In determining the breadth of an order provision, we must consider the extent of violations the transferability of the violations to other contexts, and any past history of violations. *See Sears, Roebuck and Co.* v. *F.T.C.,* 676 F.2d 385, 391–396 (9th Cir. 1982).

Respondent's violative ads were widely disseminated over several years. For example, the advertisements representing that Bayer's therapeutic superiority had been established were disseminated more than 2,600 times over a 29–month period. Also, respondent Sterling does have a previous history of dealings with the F.T.C.[54] Finally, and

[54] In 1950, the Commission entered a litigated order against Sterling based in part upon false advertising representations made regarding Bayer Aspirin. *Sterling Drug, Inc.,* 47 F.T.C. 203 (1950). In that case the Commission entered an all products order against Sterling. 47 F.T.C. at 214. On four occasions subsequent to that litigation, Sterling has consented to the entry of cease and desist orders. 49 F.T.C. 1635 (1953) (efficacy claims regarding "Campho-Phenique"); 61 F.T.C. 1008 (1962) (false representations regarding the safety of "Isuprel," a drug for oral inhalation); 73 F.T.C. 979 (1962) (false representations regarding the efficacy of a dietary supplement, "Super Ironized Yeast"); 84 F.T.C. 547 (1974) (false claims regarding medical benefits from using the spray disinfectant, "Lysol"). We do not take these consent orders as evidence of prior guilt. However, they are relevant for determining the appropriate scope of relief. Each of those consent orders applied not only to the product which had allegedly been falsely advertised, but also to other similar products. Thus, in no instance has an order entered against Sterling been limited in its scope solely to the product which was the subject of the proceeding.

most important, it is clear that respondent could easily change the names of some of its products or make inadequately substantiated claims of established superiority regarding other OTC internal analgesics. The evidence shows that respondent made inadequately substantiated establishment claims regarding both Bayer and Cope. Although we did not find a violation, the record also shows that respondent made claims of therapeutic superiority for Vanquish. Thus, our order requires all claims of established comparative efficacy made by Sterling regarding OTC internal analgesics to be substantiated by two well-controlled clinical tests. **[56]**

In numerous previous cases the Commission has issued (and courts have upheld) cease and desist orders applying to all of a company's products based upon violations committed in the advertising of only one, or a few, products. *E.g., Litton Industries, Inc.*, 97 F.T.C. 1 (1981), *aff'd*, 676 F.2d 364 (9th Cir. 1982) (misrepresentations regarding microwave oven, order applied to any product used for personal or household purposes); *Sears, Roebuck and Co.*, 95 F.T.C. 406 (1980), *aff'd*, 676 F.2d 385 (9th Cir. 1982) (misrepresentations regarding dishwasher, order applied to 12 major home appliances); *Jay Norris Corp.*, 91 F.T.C. 751 (1978), *aff'd*, 598 F.2d 1244 (2d Cir. 1979), *cert. denied*, 444 U.S. 980 (1979) (misrepresentations regarding six products, order covered general mail order merchandise); *Porter & Dietsch, Inc.*, 90 F.T.C. 770 (1977), *aff'd*, 605 F.2d 294 (7th Cir. 1979), *cert. denied*, 445 U.S. 950 (1980) (one product was misrepresented, order covered any good, drug, cosmetic or device); and *I.T.T. Continental Baking Co.*, 83 F.T.C. 865 (1973), *modified on other grounds*, 523 F.2d 207 (2d Cir. 1976) (Wonder Bread was misrepresented, order applied to any food product). The coverage of this order provision is much more narrow, applying only to other OTC internal analgesics.[55]

Paragraph II of our order applies to establishment claims which respondent makes regarding Bayer's superior pharmaceutical quality. Complaint counsel argue that these claims must be substantiated with well-controlled clinical studies because all claims of superior pharmaceutical quality imply therapeutic superiority. (C.A.B. pp. 9–27) However, as we explained above, we do believe that consumers can understand pharmaceutical or manufacturing quality (if properly characterized) as a concept separate from therapeutic superiority. It is true that many of respondent's ads which claim that Bayer is pharmaceutically superior also impliedly represent that it is therapeutically superior. But there are other ads which make representations regarding Bayer's pharmaceutical quality only. As we discussed in the liability section, Sterling's ads claimed that its tests demon-

[55] According to material submitted by Sterling to the Physician's Desk Reference (36th ed. 1982), as of 1982 Sterling manufactured 5 OTC internal analgesics.

strated both Bayer's overall pharmaceutical superiority to other aspirin as well as its superiority in terms of four specific attributes (purity, freshness, stability and speed of disintegration). Complaint counsel failed to meet its burden of showing that the tests Sterling performed do not establish Bayer's overall [57] pharmaceutical superiority. They also failed to show what sort of evidence experts require to establish the pharmaceutical superiority of an analgesic. It is, therefore, inappropriate for us to enter any order provision detailing a specific level of substantiation which Sterling must possess when it represents that Bayer is pharmaceutically superior.

Nonetheless, even a facial examination of respondent's support for its claims shows that other brands of aspirin were at least as pure, as fresh, as stable and as quick to disintegrate as Bayer. Thus, Sterling does not possess support demonstrating or establishing Bayer's superiority with respect to those attributes. Paragraph Part II of our order accordingly requires that when respondent represents that it has been established that an OTC internal analgesic is fresher, purer, more stable, or quicker to disintegrate than others, it must possess support for that claim which would satisfy relevant experts that the product has the superiority attributed to it. However, at this time we reach no conclusion as to what type or what quantity of evidence is necessary to satisfy that burden.

B. *Reasonable Basis Provision.*

Paragraph III of our order requires Sterling to possess a reasonable basis for all therapeutic performance claims regarding OTC internal analgesics. As a practical matter, this paragraph applies primarily to those claims that are not presented (either expressly or implicitly) as claims whose truth has been scientifically established. For establishment claims, Sterling will be held to the more specific standards set forth in Paragraph I, so Paragraph III adds nothing to Sterling's obligations with respect to establishment claims. The purpose of Paragraph III is to hold Sterling to a more general reasonable basis standard for all other "non-establishment" claims.

A similar provision was entered in our order against Bristol-Myers (slip op. at pp. 70–73) [102 F.T.C. at 374–377], and our reasons for entering one here are very similar. While we are unwilling to go as far as the ALJ's order, which would have imposed a reasonable basis requirement for all claims for all nonprescription drugs, we believe that a requirement limited to therapeutic performance claims for OTC analgesics is reasonably related to the violations found. Most of the claims in the case were establishment claims and we found that Sterling did not possess adequate substantiation for any of those claims. Our concern is that this violation, the making of inadequately

substantiated claims, can easily be transferred to other sorts of claims, including non-establishment claims. [58]

Moreover, in this case the record shows that respondent has already made a number of non-establishment claims without possessing the evidence required to satisfy the reasonable basis standard. We found ten such violations in its tension relief advertisements for Cope, and another 16 violations in similar advertisements for Midol (*supra* at p. 39). As we noted in *Bristol-Myers,* such a record (combined with the other factors dismissed in the previous section, such as the history of prior violations) might well justify an order extending to all claims or to all nonprescription products. Instead, we are limiting this provision to therapeutic performance claims for OTC analgesics—*i.e.,* to the exact product category and claims involved in this case. Paragraph III, thus, has a much closer relation to the violations involved here than did the broader reasonable basis provision that was struck down on appeal in *American Home Products* v. *FTC,* 695 F.2d at 710–711.

Finally, we note that paragraph III specifies that two well-controlled clinical tests will always be sufficient to constitute a reasonable basis, but it also permits Sterling to satisfy this requirement with any other "competent and reliable evidence" sufficient to provide a reasonable basis for the challenged claim. While this more general standard does leave some ambiguity regarding the absolute minimum level of evidence required to satisfy paragraph III, some flexibility is inherent in any reasonable basis order. For the reasons already discussed at length in our *Bristol-Myers* opinion (slip op. pp. 71–73) [102 F.T.C. at 375–377], we believe that the flexibility provided by paragraph III represents an appropriate balance between the need for clear standards and the need to prevent repeated violations. Should Sterling ever be in doubt about the level of evidence required for any further claim, it can always: (a) take advantage of paragraph III's "safe harbor" by conducting two well-controlled clinical tests; (b) request an advisory opinion from the commission pursuant to Rule 2.41; or (c) qualify its advertising claim to make consumers aware of the lower level of substantiation.

C. *Ingredient Claims and Omissions.*

Sterling's advertisements falsely represented that the pain reliever in Midol was special or unique and that Cope was the only OTC analgesic containing both a pain reliever and a sedative. (*supra* pp. 45–49) Under paragraph IV of the order, Sterling may not represent that a product contains any special, unusual or unique ingredient or ingredients when the same ingredients are used in other nonprescription drug products intended for the same purpose. This is the only provision of our order which we believe should apply not only to

claims regarding analgesics but to advertising claims made by respondent regarding any nonprescription drug. In determining the scope of this provision, we have considered [59] the same factors discussed in connection with paragraphs I and II of the order. First and foremost, as we discussed in *Bristol-Myers,* a false claim regarding ingredients could be made for any drug product. Second, documents from Sterling's files show it was fully aware that Cope's ingredients were not unique and that another analgesic on the market, Excedrin P.M., contained both a pain reliever and a sedative. Third, as we discussed above, this is the second time a cease and desist order has been entered against Sterling regarding the advertising of its OTC analgesics. These reasons justify entry of a broad order provision applying to all nonprescription drugs advertised by Sterling.[56]

In its appeal brief, respondent argues that entry of any order provision applying to all drugs is unjustified because there is no showing that Sterling has a history of past violations. In addition, Sterling assures us that its advertising claims were all made in good faith. (R.A.B. pp. 77–80) However, in determining the appropriate scope of order provisions, we consider all the factors discussed in *Sears.* Taken in conjunction, the ease with which the violation could be transferred to other drugs, Sterling's past history of violative advertising, and the fact that it appears Sterling knew its Cope ads were false all justify entry of an order applying to all drugs.

The purpose of paragraph V is to prevent respondent from passing off its aspirin-based analgesic products as being different from aspirin or from otherwise misrepresenting the identity of any analgesic ingredient. The principal means by which this deception has been accomplished in the past, has been to create the impression that some analgesic ingredient in respondent's product is different from the ingredient in any competing analgesic. To prevent this practice, paragraph V prohibits any misrepresentation that the analgesic ingredient in an aspirin-containing product is different from aspirin. To prevent closely related violations, the order prohibits misrepresentations regarding the identity of any analgesic ingredient in respondent's products. The order also makes clear that any attempt to contrast with aspirin the ingredient in an aspirin-based analgesic without disclosing that the ingredient in respondent's product is also aspirin, will violate the order.

The ALJ's order would have required the disclosure of aspirin's presence in any ad for an aspirin-based analgesic. However, nondisclosure of aspirin constitutes a violation only in those instances in which respondent falsely represents that the advertised product does

[56] According to material submitted by Sterling to the *Physician's Desk Reference for Nonprescription Drugs* (1st ed. 1980) as of 1980, Sterling manufactured 36 nonprescription drug products.

not contain aspirin. Thus, paragraph V is specifically tailored to prevent the sort of violation committed by respondent. **[60]**

D. *Labeling*

The order entered by the ALJ in this case applied not only to respondent's advertising, but also to the labeling of its products. As we stated in both *American Home Products,* 98 F.T.C. at 411, and *Bristol-Myers,* slip op. at p. 76 [102 F.T.C. at 380], our liaison agreement with the FDA recognizes that primary responsibility for labeling of nonprescription drugs rests with that agency. For the reasons we set forth in *American Home Products,* the order which we enter does not apply to labeling.

E. *Corrective Advertising.*

In their appeal brief, complaint counsel request that we impose a corrective advertising requirement on Sterling and require it to include a notice in its advertising disclosing that Bayer has not been proven therapeutically superior to other aspirin. (C.A.B. pp. 37–60) The ALJ declined to require corrective advertising and we agree that it would not be appropriate in this case.

Corrective advertising is a remedy available to the Commission. *Warner-Lambert Co.* v. *F.T.C.,* 562 F.2d 749, 756–759 (D.C. Cir. 1977), *cert. denied,* 435 U.S. 950 (1978). Two inquiries must be made in order to determine if it is appropriate: (1) Did the advertisements in question play a substantial role in creating or reinforcing a false belief in the public's mind regarding the product; and (2) Will the belief remain after the advertising ceases? *Warner-Lambert Co.* v. *F.T.C., Id.* at 762. Based upon our analysis of the evidence in this case, we agree with the ALJ that it is not clear that respondent's advertising played a substantial role in creating or reinforcing a false belief regarding Bayer in the public's mind.

The record contains the results of several surveys which attempted to assess the public's image of Bayer. (F. 1240–1310). Two of these surveys are particularly significant. The first is the Assets and Liabilities Study, CX 395, which measured consumer attitudes regarding Bayer in 1967, prior to the dissemination of any of the ads which are the subject of this case. The other is the Zeisel Image Study, CX 521 which was conducted in 1975, after the dissemination of challenged ads. Although a comparison of the results of these two studies is apparently somewhat difficult to perform, complaint counsel concede that the results show that Bayer's image remained relatively stable throughout the eight-year period between the two studies. (C.A.B. p. 51) Based upon this, it is difficult to conclude that Sterling's advertising created or reinforced **[61]** the public's belief regarding Bayer's

superiority. Complaint counsel argue that even though the public's image of Bayer may have remained stable during the period, it nonetheless became "sharper," that is, more of those surveyed in 1975 had opinions than in 1967. To support this they cite the testimony of Dr. Brock, an expert, in the analysis of image studies. (Tr. 5155–59) Respondent countered Dr. Brock's testimony with the testimony of its own expert, Dr. Amstutz (Tr. 10164–90). We find this evidence regarding "sharpness" inconclusive and do not believe that it supports imposition of corrective advertising.

Complaint counsel also argue that the need for corrective advertising may be inferred directly from the advertising. We decline to draw such an inference in this case. Although the ads representing Bayer's comparative superiority were disseminated on several thousand occasions, we do not think that in this case that is adequate to justify corrective advertising. Indeed, the violative ads represent only a portion of the Bayer ads which appeared during the early 1970's. (The record contains nearly twice as many Bayer ads which did not represent its therapeutic superiority.) In light of survey evidence which appears to indicate no need for corrective advertising, we find such a need may not be inferred directly from the ads.

Numerous factors contribute to a product's image. Included among these are consumers' experience with the product, publicity regarding the product, longevity and visibility in the market, amount of advertising (regardless of content), advertising content and other sources. (Miles, Tr. 9355–59) The longer a brand has been in existence, the less its image stems from one particular advertising campaign. (Miles, Tr. 9366) For a brand such as Bayer, which has been on the market for many years, familiarity is the primary influence on brand image. (Haley, Tr. 10569) Complaint counsel contend that this case is similar to *Warner-Lambert* in which the Commission ordered corrective advertising regarding Listerine, a well-established brand. (C.A.B. p. 42) However, in *Warner-Lambert,* the record showed that the respondent had been making false claims regarding Listerine in its advertising for more than 50 years and that throughout that period the false claim had been a major theme of the advertising, 86 F.T.C. 1398, 1501. The corrective advertising was designed to correct that false advertising. In this case, there has been no showing that the false advertising was so extensive. Sterling's false advertising was disseminated during only a 29–month period, and the public's image of Bayer remained stable during that period. Thus, since it has not been shown that Sterling's advertising created or reinforced the public's image of Bayer, corrective advertising is an inappropriate remedy.

X. CONCLUSION

For the reasons set forth above, the initial decision of the administrative law judge is modified as described. An appropriate order is appended.

SEPARATE STATEMENT OF COMMISSIONER PERTSCHUK CONCURRING IN PART AND DISSENTING IN PART*

For the reasons stated in my separate opinion in *Bristol-Myers* (D. 8917) [102 F.T.C. at 382], announced today, I dissent from that portion of the Commission's opinion which reverses the "substantial question" doctrine developed in *American Home Products,* 98 F.T.C. 136 (1981), *aff'd,* 695 F.2d 681 (3d Cir. 1982). Therefore, I dissent from the Commission's decision to dismiss paragraphs 12 through 14 of the complaint.

I also dissent from the portion of the Commission's opinion which dismisses complaint paragraphs 17 and 29, which allege that Sterling violated Section 5 by making contemporaneous inconsistent claims for its OTC internal analgesic drug products. The Commission dismisses these charges, not because Sterling did not make such claims, but because it sees the basis of the charge as a "new theory of advertising substantiation which would shortcut and be contrary to principles of law set forth in *Pfizer* and its progeny." Slip op. at 50. [102 F.T.C. at 358]

I disagree. The inconsistent contemporaneous claims allegation stems directly from the reasonable basis doctrine set out in *Pfizer.* In my view, application of the reasonable basis doctrine to an examination of the claims made by Sterling in this case leads inexorably to the conclusion that Sterling has made unsubstantiated claims in violation of Section 5.

The Commission agrees that Sterling represented that Vanquish was better than aspirin in relieving pain and in avoiding stomach upset (slip op. at 16, 18) [102 F.T.C. at 329, 331], and that Cope was superior to any OTC analgesic for the relief of nervous tension headache (slip op. at 20) [102 F.T.C. at 332]. At the same time it was making those claims, however, Sterling was *also* claiming that Bayer aspirin was just as good as *any* internal analgesic in relieving pain and nervous tension headaches, and avoiding stomach upset. (F. 398–402)

There is simply no way those statements can be reconciled. Sterling's claims that Vanquish and Cope were more effective than aspirin plainly conflict with Sterling's contemporaneous claim that Bayer

* Statements by Chairman Miller and Commissioners Bailey and Douglas concerning this order were issued with the Final Order in *Bristol Meyers Co., et al. See* 102 F.T.C. 381, 386, 389.

aspirin was just as effective as any OTC internal analgesic—presumably, including Vanquish and Cope. Both statements can not be true at the same time.

Nevertheless, the Commission declines to find a violation on the ground that a reasonable basis analysis does not determine whether a claim is true, and that therefore it is "theoretically possible that two inconsistent claims can both be substantiated with a reasonable basis." Slip op. at 51. [102 F.T.C. at 358] [2]

While it might be theoretically possible for two inconsistent claims to be adequately substantiated, the problem with the Commission's rationale is that it fails to consider whether it is even theoretically possible for each claim made by Sterling in *this case* to be adequately substantiated. It appears obvious to me that they cannot. If Sterling has a reasonable basis for a claim that Vanquish provides superior pain relief to aspirin, it cannot have a reasonable basis for a claim that Bayer aspirin relieves pain just as well as all OTC internal analgesics. Conversely, if Sterling has a reasonable basis for a claim that aspirin relieves pain just as effectively as all OTC internal analgesics, it cannot have a reasonable basis for a claim that Vanquish relieves pain better than aspirin. Where an advertiser makes an objective and verifiable claim that its product performs better than *any* other product, adequate substantiation for that claim necessarily precludes the advertiser from having a reasonable basis for a claim that another product works better than, or as well as, the one advertised.

The Commission seems troubled, however, by the application of an "inconsistent contemporaneous claims" theory. It notes the apparent discrepancy between the case where a single advertiser is held liable for making inconsistent claims, and the case where the same claims are made separately by two different advertisers and the Commission finds each adequately substantiated. In fact, such a result would not be anomalous. Indeed, it would be perfectly consistent with the reasonable basis doctrine, which takes into account not only the sufficiency of the evidence on which an advertiser relies but also "the reasonableness of the advertiser's action and his good faith." *National Dynamics Corp.*, 82 F.T.C. 488, 553 (1973). In considering an advertiser's reasonableness, the Commission routinely considers information in the advertiser's possession which might give the advertiser reason to question the evidence relied upon to substantiate a claim. Clearly, an advertiser possessing data which directly contradicts a claim cannot have a reasonable belief in the truth of that claim. On the other hand, if the contradictory evidence exists but the advertiser is unaware of it and would have no reason to know about it, the advertiser would not be precluded from making the claim. In other words,

whether or not there is liability depends, at least in part, on the advertiser's knowledge. The application of the inconsistent contemporaneous claims theory simply is one example of the effect of this standard, and accordingly reflects no deviation from the established reasonable basis doctrine.

It is true, as the majority notes, that we could have proceeded to determine *which* of Sterling's claims was the one that lacked a reasonable basis. But where the conclusion is inescapable, as it is here, that one claim or the other lacked a reasonable basis, it seems like a waste of resources to require both sides to go through the full panoply of evidentiary exchanges just to find out which claim was the one to violate [3] Section 5. Accordingly, I would have sustained the allegations of the complaint with respect to the making of contemporaneous inconsistent claims.

FINAL ORDER

The matter has been heard by the Commission upon the appeal of counsel for respondent Sterling Drug, Inc., and complaint counsel and upon briefs and oral argument in support of and in opposition to the appeals. The Commission, for reasons stated in the accompanying Opinion, has granted a portion of respondent's appeal and denied that of complaint counsel. Therefore

It is ordered, That the initial decision of the administrative law judge be adopted as the Findings of Fact and Conclusions of Law of the Commission except as is otherwise inconsistent with the attached opinion.

Other Findings of Fact and Conclusions of Law of the Commission are contained in the accompanying Opinion.

It is further ordered, That the following Order to Cease and Desist be entered. [2]

ORDER

I

It is ordered, That Sterling Drug, Inc., its successors and assigns, and its officers, agents, representatives and employees, directly or through any corporation, subsidiary, division or other device, in connection with the advertising, offering for sale, sale or distribution of "Bayer Aspirin," "Bayer Children's Aspirin," "Vanquish," "Cope," "Midol," or other nonprescription internal analgesic product, in or affecting commerce, as "commerce" is defined in the Federal Trade Commission Act, do forthwith cease and desist from:

Making any representation, directly or by implication, that a claim concerning the superior effectiveness of such product has been established or proven unless such representation has been established by two or more adequate and well-controlled clinical investigations, conducted by independent experts qualified by training and experience to evaluate the comparative effectiveness of the drugs involved, on the basis of which it could fairly and responsibly be concluded by such experts (1) that the drug will have the comparative effectiveness that it is represented to have, and (2) that such comparative effectiveness is demonstrated by methods of statistical analysis, and with levels of confidence, that are generally recognized by such experts. The investigations shall be conducted in accordance with the procedures set forth below.

At least one of the adequate and well-controlled clinical investigations to evaluate the comparative effectiveness of the drug shall be conducted on any disease or condition referred to, directly or by implication, or, if no specific disease or condition is referred to, then the adequate and well-controlled clinical investigations shall be conducted on at least two conditions or diseases for which the drug is effective. The clinical investigations shall be conducted as follows:

A. The subjects must be selected by a method that:

1. Provides adequate assurance that they are suitable for the purposes of the investigation, and the diagnostic criteria of the condition to be treated (if any); [3]

2. Assigns the subjects to the test groups in such a way as to minimize bias; and

3. Assures comparability in test and control groups of pertinent variables, such as age, sex, severity or duration of disease or condition (if any), and use of drugs other than test drugs.

B. The investigations must be conducted double-blind, and methods of double-blinding must be documented. In addition, the investigations shall contain a placebo control to permit comparison of the results of use of the test drugs with an inactive preparation designed to resemble the test drugs as far as possible.

C. The plan or protocol for the investigations and the report of the results shall include the following:

1. A clear statement of the objective of the investigation;

2. An explanation of the methods of observation and recording of results, including the variables measured, quantitation, assessment of any subject's response and steps taken to minimize bias on the part of the subject and observer;

3. A comparison of the results of treatments or diagnosis with a

control in such a fashion as to permit quantitative evaluation. The precise nature of the control must be stated and an explanation given of the methods used to minimize bias on the part of the observers and the analysts of the data;

4. A summary of the methods of analysis and an evaluation of data derived from the study, including any appropriate statistical methods.

D. A test or investigation which is not conducted in accordance with these procedures may be used to establish a claim only if respondent can show that, notwithstanding the failure to satisfy these procedures, the test or investigation would still be generally accepted by the relevant scientific community as sufficient to establish the truth of the claim. [4]

II

It is further ordered, That respondent Sterling Drug, Inc., a corporation, its successors and assigns, and its officers, agents, representatives and employees, directly or through any corporation, subsidiary, division or other device, in connection with the advertising, offering for sale, sale or distribution of "Bayer Aspirin," "Bayer Children's Aspirin," "Vanquish," "Cope," "Midol," or any other nonprescription internal analgesic product, in or affecting commerce, as "commerce" is defined in the Federal Trade Commission Act, do forthwith cease and desist from making any representation, directly or by implication, that the superior freshness, purity, stability, or speed of disintegration of such product has been established, demonstrated, or proven unless at the time such representation is made, respondent possesses and relies upon competent and reliable scientific evidence which would permit qualified experts to conclude that the product has the comparative pharmaceutical qualities it is represented to have.

III

It is further ordered, That respondent Sterling Drug, Inc., its successors and assigns, and its officers, agents, representatives and employees, directly or through any corporation, subsidiary, division or other device, in connection with the advertising, offering for sale, sale or distribution of "Bayer Aspirin," "Bayer Children's Aspirin," "Vanquish," "Cope," "Midol" or any other nonprescription internal analgesic, in or affecting Commerce, as "commerce" is defined in the Federal Trade Commission Act, do forthwith cease and desist from making any therapeutic performance claim for such product unless respondent possesses a reasonable basis for making that claim. A

reasonable basis for such a claim shall consist of competent and reliable scientific evidence supporting that claim. Well-controlled clinical tests conducted in accordance with the criteria set forth in Order Paragraph I shall be deemed to constitute a reasonable basis for a claim.

IV

It is further ordered, That respondent Sterling Drug, Inc., its successors and assigns, and its officers, agents, representatives and employees, directly or through any corporation, subsidiary, division or other device, in connection with the advertising, offering for sale, sale or distribution of "Bayer Aspirin," "Bayer Children's Aspirin," "Vanquish," "Cope," "Midol," or any other nonprescription drug product in or affecting commerce, as "commerce" and "drug" are defined in the Federal Trade Commission Act, do forthwith cease and desist from making any representation, [5] directly or by implication that such product contains any unusual, special or unique ingredient or ingredients when such ingredient or ingredients are commonly used in other nonprescription drug products intended for the same use or uses as the product advertised by respondent.

V

It is further ordered, That respondent Sterling Drug, Inc., its successors and assigns, and its officers, agents, representatives and employees, directly or through any corporation, subsidiary, division or other device, in connection with the advertising, offering for sale, sale or distribution of "Bayer Aspirin," "Bayer Children's Aspirin," "Vanquish," "Cope," "Midol," or any other nonprescription internal analgesic in or affecting commerce, as "commerce" is defined in the Federal Trade Commission Act, do forthwith cease and desist from falsely representing that the analgesic ingredient in an aspirin-containing product is different from aspirin or otherwise misrepresenting the identity of any analgesic ingredient. It shall be a violation of this paragraph to contrast the analgesic ingredient of a product which contains aspirin with the analgesic ingredient of another product if that product also contains aspirin, unless respondent discloses clearly and conspicuously that the analgesic ingredient in its product is aspirin.

VI

It is further ordered, That respondent Sterling Drug, Inc., shall notify the Commission at least thirty (30) days prior to any proposed change in the corporation such as a dissolution, assignment or sale resulting in the emergence of a successor corporation, the creation or dissolution of subsidiaries or any other change in its corporation which may affect compliance obligations under this Order.

VII

It is further ordered, That the respondent herein shall within sixty (60) days after service of this Order upon it and at such other times as the Commission may require, file with the Commission a written report setting forth in detail the manner and form in which it has complied or intends to comply with this Order. [6]

Complaint paragraphs Eight A.2, Eight B, Eight C, Ten B, Twelve, Thirteen, Fourteen, Fifteen A, Seventeen, Twenty-Three, Twenty-Four, and that portion of Twenty-Nine which refers to Seventeen are hereby dismissed.