Perception & Psychophysics
1976, Vol. 20 (3), 163-167

Lump detection in simulated human breasts

CALVIN K. ADAMS
Departments of Ophthalmology and Psychology, University of Florida, Gainesville, Florida 32601

DEBORAH C. HALL and H. S. PENNYPACKER
Department of Psychology, University of Florida, Gainesville, Florida 32601

MARK KANE GOLDSTEIN
Department of Psychology, University of Florida, and V.A. Hospital, Gainesville, Florida 32601

LARRY L. HENCH and MICHAEL C. MADDEN
Department of Materials Science and Engineering, University of Florida, Gainesville, Florida 32601

GERALD H. STEIN
Department of Medicine, University of Florida, and V.A. Hospital, Gainesville, Florida 32601

and

A. CHARLES CATANIA
Department of Psychology, University of Maryland Baltimore County, Baltimore, Maryland 21201

Sixteen observers palpated silicone models of human breasts containing lumps 1.6-12.1 mm in diameter. Detectability depended on the size of the lump, producing a systematic psychometric function. In eight observers who participated in three or more sessions, performance improved with practice, with most improvement occurring within one or two 26-trial sessions. Three-week retention measures disclosed no appreciable decrease in performance, but a significant correlation was found between the number of lumps detected and duration of trial ($p < .01$). There was no difference in performance between four observers who used their preferred hands and four observers who used their nonpreferred hands. These data establish that examination of breast models for the detection of lumps simulating cancer is a task amenable to experimental analysis.

The potential effectiveness of breast self-examination (BSE) as a screening procedure for early detection of breast cancer depends on the effectiveness of manual palpation. The external location of the breast, coupled with the softness of the tissue and the hardness of its backing, makes it an ideal organ for physical examination. In addition, 94% of all cancerous lesions of the breast are potentially palpable; the remaining 6% are considered to be poor candidates for early detection by manual palpation alone (Haagensen, 1971). There is general agreement that the earlier a breast cancer is detected, diagnosed, and treated, the greater are chances for survival (Adair, Berg, Joubert, & Robbins, 1974; Delario, 1959; Fisher, Slack, & Bross, 1969; Haagensen, 1971; Spratt & Donegan, 1967).

The authors wish to thank Drs. Bruce W. Brient, Bernard Cantor, Donald Greer, and William R. Pfaff for technical and medical consultation. We gratefully acknowledge the effort of Raymond C. Cole, Technical Director, Apparel Development, Division of Research and Development of the International Playtex Corporation, in supplying national sales data from the foundation garment industry. We are especially grateful to General Electric Corporation for providing the materials. Send reprint requests to Dr. Calvin K. Adams, Department of Ophthalmology, J-284, JHMHC, University of Florida, Gainesville, Florida 32610.

Although BSE has in the past few years received considerable attention as an early screening procedure, there has, to date, been only one known clinical demonstration that regular practice of BSE results in the detection of smaller tumors. Capraro (Note 1), a practicing physician who has advocated and taught BSE to his patients for years, conducted a study in which he compared the size of the breast tumor at surgery in three groups of 51 women each. He had instructed the women in Group 1 in BSE, and they had reported regular self-examination. The women in Group 2 were under the regular care of other physicians, some having had BSE instruction and some not. The Group 3 women had never been taught BSE. The results were significant, with virtually no overlap in the ranges of tumor sizes among the three groups at the time of surgery. The ranges for Groups 1, 2, and 3 were, respectively, .5-2.0 cm, 2.0-3.0 cm, and 5.0 cm and larger. These data clearly establish that BSE can lead to the detection of smaller tumors.

Wolfe (1974) presented clinical data on the palpability of various tumor sizes (<.5 cm to >2.0 cm). He noted that detectability of lumps increased systematically with lump size, but, more importantly,

that nearly 50% of tumors .6-1.0 cm in size were detectable on clinical examination. Thiessen (1971), in a review of the practice of BSE, pointed out the limitations of breast size and tissue characteristics. The "suitability" of a breast for potential detection of a small lesion on physical examination is, according to this author, primarily related to size—small to medium breasts being the most suitable. The second selection factor was the intrinsic support characteristics of the breast—pendulous or poorly supported breasts being more suitable than those that are firm or well supported. In a study of 150 women, he classed half of them as having breasts suitable for BSE. Regretably, Thiessen presented no quantifiable data regarding physical dimensions or support characteristics of suitable breasts.

All of the relevant reports cited have been clinical. That is, no study reporting an experimental analysis of factors influencing lump detection in the breast or any similar medium was found. Nevertheless, the task seems primarily a sensory one (detection of small forms in a "noisy" environment) that requires a systematic search (motor) pattern.

The present study is a feasibility study and was designed as such. It is the first in an experimental series the objective of which is to isolate and evaluate the factors that influence detection of lumps in simulated breast tissue. The approach is to determine psychophysically the detectability of lumps in a life-like model of the human breast. By allowing observers to palpate models in which various numbers of lumps of different sizes have been placed, it should be possible to determine detection thresholds. Further, such a procedure should also permit determination of the stability of thresholds over time and their susceptibility to reduction as a result of practice or training or both.

METHODS

Subjects

A total of 15 women and 1 man participated in this study as observers. All were unpaid volunteers. They ranged in age from 12 to 58 years, and 15 were right-handed (1 observer was left-handed, designated K). No attempt was made to select or exclude volunteers on the basis of familiarity with BSE. None had previously participated in a similar experiment.

Apparatus

Developing a model of the human breast that closely duplicated the size, shape, consistency, and physical characteristics of living tissue (normal and cancerous) was seriously hampered by the lack of quantitative data describing the breasts of any population of women. However, data were obtained from a brassiere manufacturer (Cole, Note 2) and included a relative frequency distribution of sales by size (32A, 36C, etc). The 34B brassiere accounted for most sales; progressively larger and smaller sizes accounted for progressively fewer sales. Quantitative data on the physical characteristics of breast skin and of adipose, glandular, and connective tissue (Cooper's ligaments) were also unavailable. Academic physicians in obstetrics and gynecology and in surgery were therefore asked to palpate the models and advise on their development. The models developed were judged to approximate closely the size and physical characteristics of the young, nongranular, well-supported, "B-cup" breast of a young woman. The lumps were judged to simulate accurately the characteristics of firm, well-fixed tumors.

Only three models were constructed for this initial investigation, because the models were expensive and it was imperative that consistent physical characteristics across models be maintained. Each model consisted of a thin (.25-.50 mm), clear silicone polymer skin surrounding a uniform clear silicone gel. Each model was a simple hemisphere of 2.25-in. (5.72-cm) radius (approximately the volume of a 34B brassiere cup). A hemispherical nipple of 1/8-in. (0.32-cm) radius was placed at the apex to serve as a tactile reference point. Tumors were simulated by steel spheres, attached to the back of the model by a thin (.2 mm) layer of silicone polymer. One model (the "clean" model) had no lumps. One had three large lumps (8.7, 11.1, and 12.7 mm in diameter) placed in the middle third of each of three radii and spaced 120° apart. The third model had five lumps (1.6, 2.4, 3.2, 4.8, and 6.4 mm in diameter), each placed in the middle third of each of five radii spaced approximately 72° apart.

Procedure

Testing was conducted in a quiet room under normal illumination. The observer was seated in a comfortable chair at a small table. On the table was a small platform, 36 in. long × 12 in. wide × 1 in. high (91.4 × 30.5 × 2.54 cm). Attached to the platform and facing the observer was a vertical frame, 30 in. wide × 30 in. tall (76.2 × 76.2 cm), on which was stretched a piece of heavy black felt (with armhole slits) to eliminate visual cues because the models were transparent. Centered on the platform was a 12 × 12 × 7/16 in. thick (30.5 × 30.5 × 1.1 cm) piece of polyurethane foam. Each model was placed at the center of this piece of foam for examination.

When the first session began, the examiner read the following instructions to each observer:

"(1) This is an experiment to see how accurately lumps in the breast can be felt. I'm going to present to you a series of breast models that may or may not have lumps in them, and I want you to find as many as you can. The models will be presented behind this screen because the lumps can be seen through the surface.

"(2) [Practice on clean breast.] Here is a representative model with no lumps. Examine it for a few minutes. Some of the lumps in the model to be presented may be deep against the back surface so press just hard enough to feel all the way through. Please don't jab with your fingernails or handle the model too roughly—don't do anything to the model that you wouldn't do to yourself. There is a nipple at the center that you can use as a reference point to tell you where you are in the search pattern. Examine in any manner you like; just make sure you cover the entire area. [Allowed observer to practice searching five times around the breast area.]

"(3) [Explanation of trial procedure.] You will be allowed one search around on each trial; always start here at the 3 o'clock position [for right-handed searches, and 9 o'clock position for left-handed searches] and move around clockwise. I will position your hand at the proper place at the beginning of each trial and indicate that the trial has begun. When you think you feel a lump, you may probe a bit with your fingers (again please watch your nails); but do not try to move the lump or handle it roughly. When you think that you feel one, indicate that to me. Do not move your hand from the spot. For each lump that you feel, I will ask you to estimate your confidence on a scale of 1 to 5, where 1 indicates little confidence and 5 indicates virtual certainty. When you have indicated detection and confidence, continue to search the rest of the breast, indicating in the same manner any other lumps you find. I will tell you when you have completed one search, and the trial will be terminated.

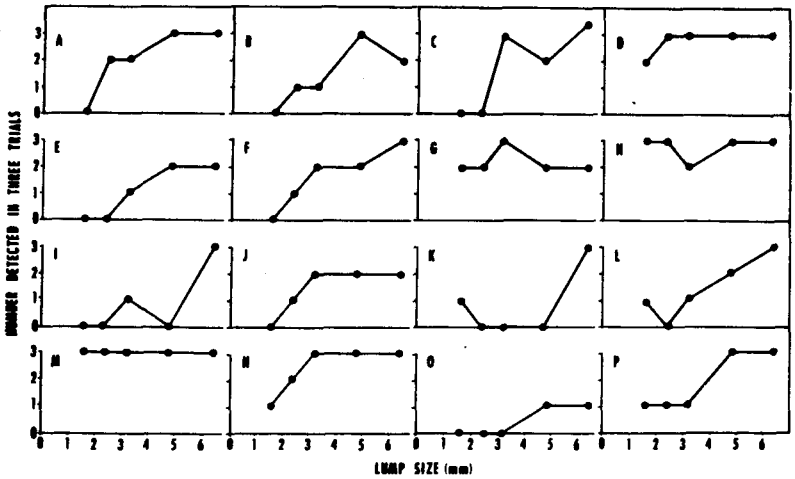

Figure 1. The number of lumps detected in the first three trials of the first session as a function of lump size for the 16 individual observers. Data for the five smallest lumps only are included.

If you find no lumps, I will again tell you when you have completed one search and ask you to estimate your confidence that no lumps were present. Any questions?

"(4) [Demonstration of hand position for search. Model was removed and observer was seated comfortably in front of the table and screen.] Please place your preferred (or nonpreferred) hand through this slot and rest it comfortably on the table off to the right (or left). [Positioned hand.] Between trials please rest your hand comfortably on the board directly behind the screen off to the side like this. I will place the models directly centered on the foam and will place your hand on the correct starting point when the trial is to begin. When the trial is over, please rest your hand to the side again so that the next model may be placed correctly. [Demonstrated positioning of hand.] Any questions?"

Each session consisted of 26 trials—10 presentations of each of the models with lumps and 6 presentations of the clean model. Order of presentation was random, but with the restriction that no model was presented more than twice in succession, and that the clean model was presented at least once in each block of six trials. Radial orientation of each model was varied randomly in 120° increments from trial to trial. Intertrial intervals were not strictly specified, but usually ranged from 5 to 15 sec. The models were maintained at room temperature (21° ± 1°C).

The use of only three models (rather than the nine required for an unmodified method of constant stimuli) necessitated careful design of the use of three. It was clearly recognized from the beginning that this restriction increased the possibility that with practice some observers might recognize one or more of the models, and this could induce position preferences in their search patterns or otherwise contaminate their performance with cognitive bias. The procedures of restricting the search to once around the model, rotating the radial orientation of each model on each trial, and using spherical rather than naturally shaped models were specifically introduced to minimize these possibilities. The psychophysical procedure was thus a modified method of constant stimuli with group presentation of the stimuli. The order of stimuli in each model was fixed, but the starting point in the order was varied.

Initially, eight observers (seven women and one man) were tested for a single session (all using their preferred hands) to establish that systematic data could be obtained and that the range of lump sizes was appropriate. Then data were obtained from the eight remaining observers, in a replicated baseline procedure designed to investigate practice effects, bilateral equivalence, and retention over time. These observers were given a single daily session until performance (% detected) had stabilized. No sustained increase in performance was observed. Four were tested using the preferred hand, and four using the nonpreferred hand. Three observers (two using the preferred hand; one, the nonpreferred hand) participated in only three sessions (two inadvertently saw the models, and one became unavailable for further testing). After 3 weeks, the remaining five observers were retested for a single session, using the same hand they had used in their previous sessions.

RESULTS

Figure 1 shows the number of times each of the 16 observers detected each lump on the first three trials of the first test session. Only data obtained from the model with five small lumps are presented. These data demonstrate that from the beginning of testing the detectability of lumps was primarily dependent on the size of the lump. Figure 1 reveals

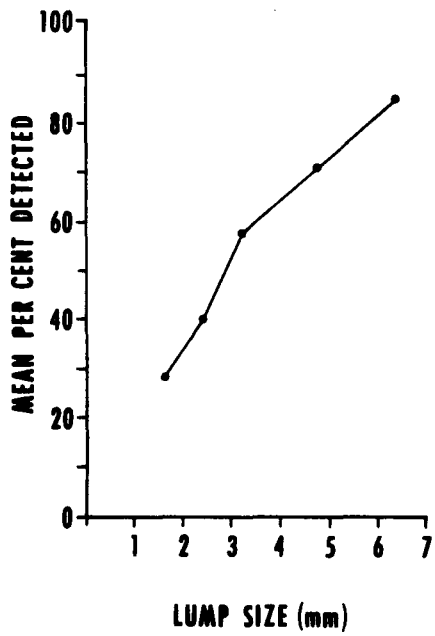

Figure 2. Percentage of lumps detected in the first three trials of the first session as a function lump size. Data are the average of the 16 observers presented in Figure 1.

that the number detected decreased as lump size decreased for all but three observers (G, H, M), who showed essentially errorless performance from the outset.

The pooled data from all 16 observers on the first three trials are presented in Figure 2. The estimated mean threshold, based on linear interpolation to 50%, is 2.4 mm. The relationship between lump size and detectability is systematic.

To evaluate changes in performance as a function of practice, the data from the eight observers who participated in three or more sessions were pooled and averaged across five trial blocks (each block equals half a session, since each model with lumps was presented 10 times each session). These data are presented in Figure 3, and they again show systematic psychometric functions, with rapid increases in the percentage of small lumps detected. The inset in Figure 3, which shows the estimated average thresholds, based on linear interpolation to 50%, demonstrates that stable performance had been reached after only 1½ sessions of testing. The data in Figure 3 also indicate that because these observers improved markedly in detecting small lumps, evaluating their performance by using threshold as the dependent measure after the first session was not an optimal strategy. That is, the performance of all but one of the observers had by then exceeded the range of performances the instrument was designed

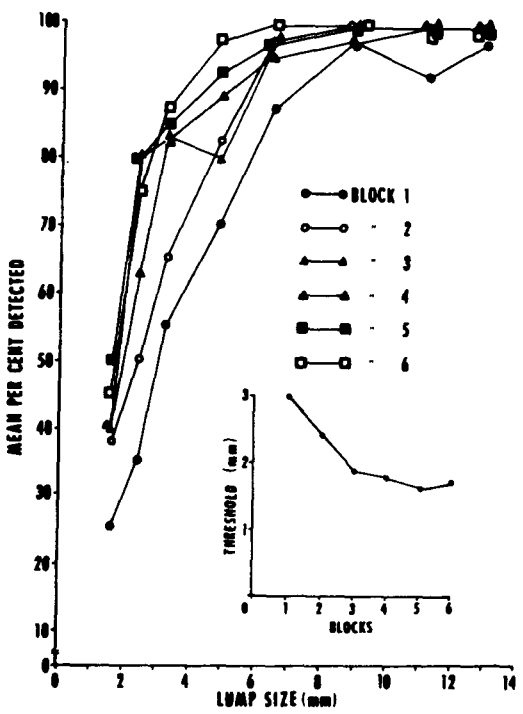

Figure 3. Average detectability functions for the eight observers who participated in three or more sessions as a function of lump size. Parameter is blocks of trials (each block is five trials, or one-half session). Inset shows estimated mean thresholds as a function of blocks.

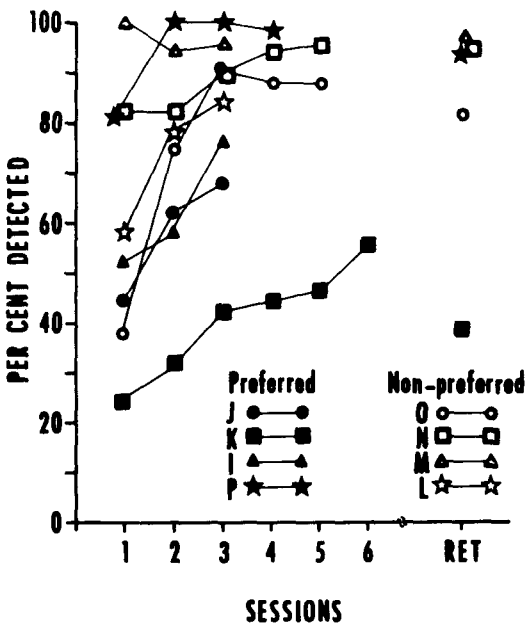

Figure 4. Individual detectability functions for the eight observers who participated in three or more sessions as a function of sessions. Open symbols represent observers tested with the nonpreferred hand, and closed symbols represent observers tested with the preferred hand. Retention interval was 3 weeks.

to evaluate with maximum sensitivity using threshold as the dependent measure. To minimize the resulting loss in sensitivity and to use all of the data, rather than only those from the two smallest lumps, subsequent analyses of the data were based on absolute number of lumps detected.

Figure 4 presents, session by session, the individual detection functions for each of the eight observers, where each observer's data for each session are represented by a single data point. All data are based on detection of the five smallest lumps only. Figure 4 portrays two major findings. First, all but one of the observers showed marked improvement in performance during the first three sessions, and that subject (M) could not improve because she began at an errorless level. Second, performance after a 3-week retention interval showed no marked change for three observers (M, N, P), and a minor decrease for two observers (K, O).

Only 2 observers generated more than one "false alarm" (detection response in the clean breast), and 12 emitted none. The confidence rating data are not included because all but one observer responded with ratings of 5 after the first few trials.

DISCUSSION

The present data establish that examination of breast models for the detection of lumps simulating cancer is a task amenable to experimental analysis. From the beginning of practice, detectability of lumps depended primarily on their size. But

a marked practice effect occurred during the first few sessions. Finally, after a retention interval of 3 weeks, little or no decrease in performance occurred.

Secondary analyses of the data disclosed that an observer's age is not a good predictor of performance. For 16 observers, the Pearson product-moment correlations were all low and nonsignificant—age with number of lumps detected in the first three trials, $r = .24$; age with threshold, $r = -.12$; or age with mean trial duration in seconds, $r = -.11$. Conversely, trial duration was related to detection performance ($r = .67$, $p < .01$) between mean trial duration and number of lumps detected in the first three trials. Further, there were no reliable warm-up or fatigue effects within sessions; and no meaningful differences emerged in the practice effect, the number detected, or retention scores between the groups of observers who practiced with their preferred vs. nonpreferred hands.

Finally, the data gave no indication of the observer's being aware that only three models were being used. Specifically, there were no step-function changes in percentages detected or trial durations to indicate recognition of any model. Although only one circular search of the model was allowed on each trial, observers frequently reported more lumps than were present, detecting some lumps more than once during a trial. Other observers never found all of the lumps consistently; thus they had no way of knowing how many lumps were in each model because no feedback on performance was ever given to them. Although the results are not presented here, further testing (using the opposite hand) was conducted on the five observers who completed the study. When first tested with the opposite hand, two observers improved slightly, one showed no change, and two showed considerable decrement in performance. These results would not be expected if the observers had learned to recognize the models.

In summary, the findings of this feasibility study warrant the conclusion that the practice of BSE is a behavioral phenomenon amenable to controlled investigation. Continued experimentation will almost surely yield data that far more sensitively and exactly describe the functional relationships existing between detection and several medically significant physical factors. Once these relations are described and understood, substantial improvements in the technology of mass instruction and screening may confidently be expected.

REFERENCE NOTES

1. Capraro, V. C. Personal communication, August 1, 1975.
2. Cole, R. Personal communication, September 15, 1975.

REFERENCES

ADAIR, F., BERG, J., JOUBERT, L., & ROBBINS, G. Long-term follow up of breast cancer patients: The 30-year report. *Cancer*, 1974, **33**, 1145-1150.

DELARIO, A. J. *Breast cancer: Factors modifying prognosis*. New York: Macmillan, 1959.

FISHER, B., SLACK, N. H., & BROSS, I. D. Cancer of the breast: Size of neoplasm and prognosis. *Cancer*, 1969, **24**, 1071-1080.

HAAGENSEN, C. D. *Diseases of the breast*. Philadelphia: W. B. Saunders, 1971.

SPRATT, J. S., & DONEGAN, W. L. *Cancer of the breast*. Philadelphia: W. B. Saunders, 1967.

THIESSEN, E. Breast self-examination in proper perspective. *Cancer*, 1971, **28**, 1537-1545.

WOLFE, J. Analysis of 462 breast carcinomas. *American Journal of Roentgenology, Radium Therapy and Nuclear Medicine*, 1974, **121**, 846-853.

(Received for publication January 23, 1976;
revision accepted June 16, 1976.)