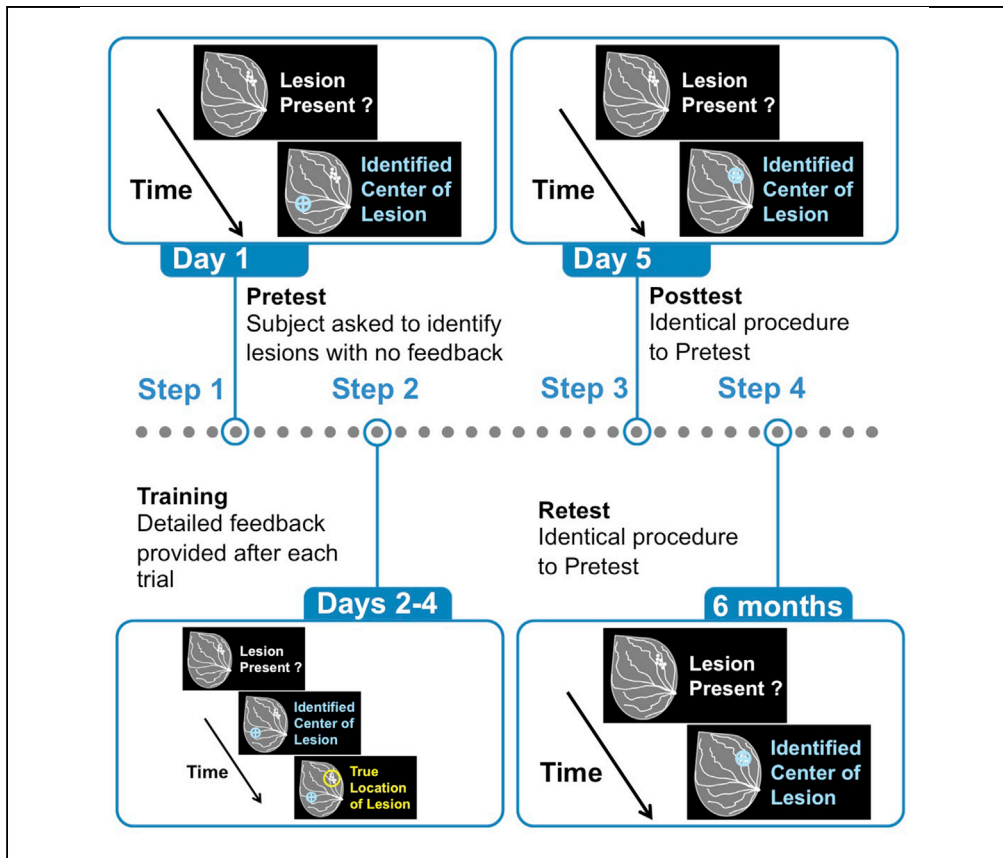


Protocol

A behavioral training protocol using visual perceptual learning to improve a visual skill



We describe a behavioral training protocol using visual perceptual learning (VPL) to improve visual detection skills in non-experts for subtle mammographic lesions indicative of breast cancer. This protocol can be adapted for the professional training of experts (radiologists) or to improve visual skills for other tasks, such as the detection of targets in photo or video surveillance.

Sebastian M. Frank,
Andrea Qi, Daniela
Ravasio, Yuka
Sasaki, Eric L. Rosen,
Takeo Watanabe

sebastian_frank@brown.
edu (S.M.F.)
takeo_watanabe@brown.
edu (T.W.)

HIGHLIGHTS

Behavioral training
using VPL induces
long-lasting
improvements of a
visual skill

Training should be
conducted using
detailed feedback
about response
accuracy

Training can be
conducted with
minimal technical
equipment

VPL protocol can be
used for clinical or
other professional
training

Frank et al., STAR Protocols 2,
100240
March 19, 2021 © 2020 The
Authors.
[https://doi.org/10.1016/
j.xpro.2020.100240](https://doi.org/10.1016/j.xpro.2020.100240)



Protocol

A behavioral training protocol using visual perceptual learning to improve a visual skill

Sebastian M. Frank,^{1,4,*} Andrea Qi,¹ Daniela Ravasio,¹ Yuka Sasaki,¹ Eric L. Rosen,^{2,3} and Takeo Watanabe^{1,5,*}

¹Brown University, Department of Cognitive, Linguistic, and Psychological Sciences, 190 Thayer St., Providence, RI 02912, USA

²Stanford University, Department of Radiology, 300 Pasteur Drive, Stanford, CA 94305, USA

³University of Colorado Denver, Department of Radiology, 12401 East 17th Avenue, Aurora, CO 80045, USA

⁴Technical contact

⁵Lead contact

*Correspondence: sebastian_frank@brown.edu (S.M.F.), takeo_watanabe@brown.edu (T.W.)
<https://doi.org/10.1016/j.xpro.2020.100240>

Summary

We describe a behavioral training protocol using visual perceptual learning (VPL) to improve visual detection skills in non-experts for subtle mammographic lesions indicative of breast cancer. This protocol can be adapted for the professional training of experts (radiologists) or to improve visual skills for other tasks, such as the detection of targets in photo or video surveillance. For complete details on the use and execution of this protocol, please refer to Frank et al. (2020a).

Before you begin

VPL can be defined as a long-term performance enhancement on a visual task after visual experience (Seitz and Dinse, 2007; Sasaki et al., 2010; Sagi, 2011; Doshier and Lu, 2017). VPL is a powerful tool to improve visual detection and discrimination abilities in healthy subjects (e.g., Ball and Sekuler, 1982; Karni and Sagi, 1991; Fahle and Edelman, 1993; Bang et al., 2018) and in subjects with impaired vision (e.g., Polat et al., 2004; Ding and Levi, 2011; Hussain et al., 2012). VPL has been primarily investigated using primitive visual features such as texture (Karni and Sagi, 1991), orientation (Schoups et al., 2001) or motion direction (Ball and Sekuler, 1982; Frank et al., 2020b) but it can also occur for more complex visual stimuli such as visual feature conjunctions (e.g., Frank et al., 2014, 2016).

Several factors modulate the development of VPL, including (but not limited to) (1) the total number of training sessions (e.g., Karni and Sagi, 1991), (2) the number of trials per training session (e.g., Amar-Halpert et al., 2017; Shmuel et al., 2020), (3) feedback signals during training (e.g., Fahle and Edelman, 1993; Herzog and Fahle, 1997; Frank et al., 2020a), (4) reinforcement signals during training (e.g., Seitz and Watanabe, 2003; Law and Gold, 2009; Roelfsema et al., 2010), (5) a night of continuous sleep (~6–8 h) between successive training sessions (e.g., Tamaki et al., 2020), and (6) subject expertise prior to training (e.g., Bejjanki et al., 2014). We will briefly discuss the effectiveness of each factor in the [Experimental design considerations](#). Professional visual trainings designed with respect to these modulating factors of VPL have a high potential to induce a long-lasting improvement of the trained visual skill (Karni and Sagi, 1993; Frank et al., 2018).

In the study of Frank et al. (2020a), we used VPL to improve the visual detection abilities of naïve subjects for subtle targets in complex natural images. Specifically, we trained college-age subjects



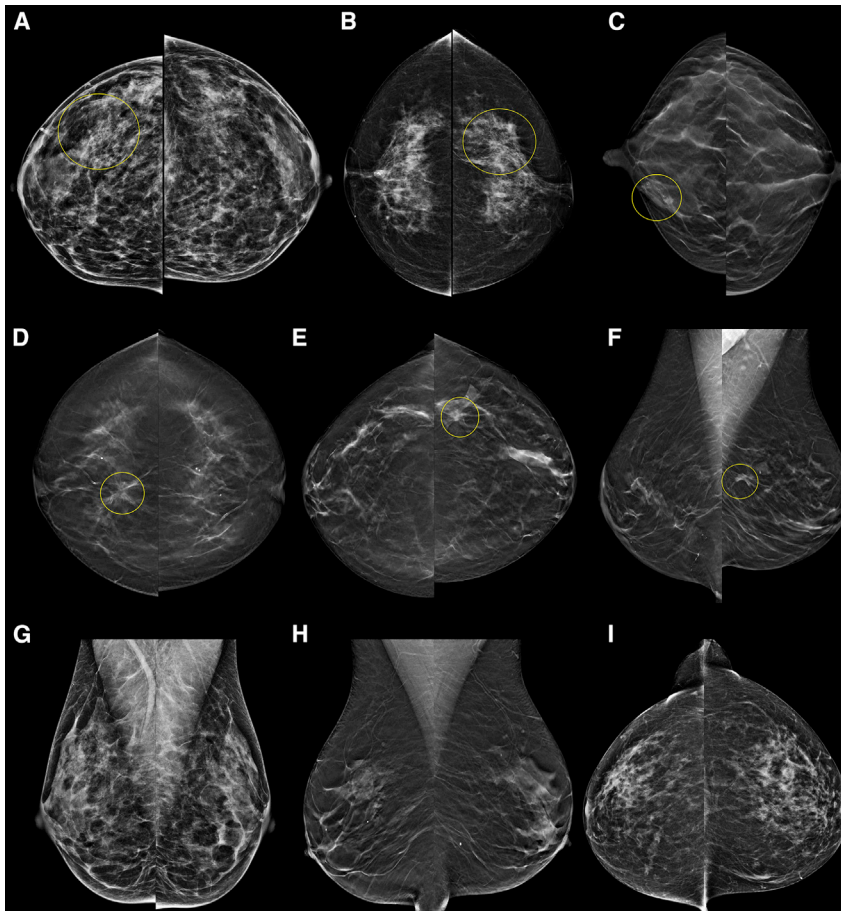


Figure 1. Different types of mammographic images used for training

(A–C) Example mammograms from three different patients. Each mammogram contains “grouped microcalcifications” (referred to as “calcification” lesion in the following), defined as fine white specks, tightly grouped together. Calcification lesions are indicative of breast cancer. For each patient the most representative slices of the left and right breasts were presented to subjects in the experiment. The yellow circle shows the location of the whole lesion and was only presented to subjects during detailed feedback after each trial of training.

(D–F) Same as (A)–(C) but example mammograms from three different patients with “architectural distortion” (referred to as “distortion” lesion in the following), defined as lines radiating to a central point, similar to the spokes of a wheel. This is another type of mammographic lesion indicative of breast cancer.

(G–I) Same as (A)–(C) but examples of mammograms without any lesion (referred to as “normal mammograms” in the following).

without any medical background to better detect one of two types of mammographic lesions, referred to as “grouped microcalcifications” and “architectural distortion” lesions (referred to as “calcification” and “distortion” lesions in the following; see Figure 1). These two types of lesions are indicative of breast cancer and they were chosen, because they are very difficult to detect by radiologists without years of professional experience (Bird et al., 1992; Birdwell et al., 2001; Rangayyan et al., 2010; Bahl et al., 2015). We designed a training schedule (Figure 2) consisting of three short training sessions (< 1 h per session) on separate days during which subjects were presented with mammograms from different patients on a computer screen and asked to decide whether a lesion was present in the mammogram and if so, where it was located. To facilitate the development of VPL the training was conducted by presenting subjects with detailed feedback about response accuracy after each training trial (Figure 3). We found that subjects developed VPL over the course of training, such that their visual detection abilities for the trained mammographic lesion were significantly increased after training compared with pretraining (Figure 6). Furthermore, the developed

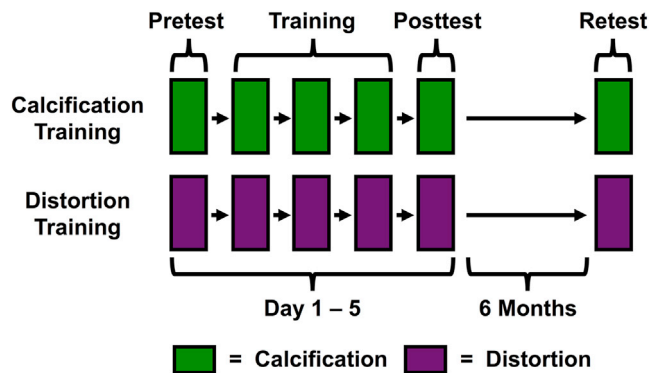


Figure 2. Study design

The study consisted of pretest, posttest, and retest sessions. Between pretest and posttest three training sessions were conducted. Six months after the posttest the long-term retention of VPL of trained subjects was assessed in a retest session. Each session of the experiment was conducted on a separate day and was about 45 min to 1 h in duration. Future experiments should consider including a greater number of training sessions to facilitate VPL. Two types of lesions were trained in different subject groups (calcification lesions, shown in green color, and distortion lesions, shown in purple color). Each group of subjects was only trained on one type of lesion (between-subject design). Additional groups of subjects can be included for other types of mammographic lesions in future studies.

VPL lasted for several months after training end, indicating that a long-lasting improvement of the visual skill to detect the trained type of lesion has occurred (see Frank et al., 2020a).

Here, we describe a simple, effective, and short protocol to train non-experts to detect lesions in mammograms that could potentially be used in the professional training of radiologists. This protocol can also be adapted to improve the visual detection skills for other targets, e.g., in photo or video surveillance.

Experimental design considerations

Several factors are known to modulate VPL and are critical in designing an effective training protocol:

1. Total number of training sessions
 - a. VPL is facilitated by a greater number of training sessions (e.g., Karni and Sagi, 1991).
 - b. It is important to note that improvements in VPL often occur after a night of continuous sleep between successive training sessions (see below), thus training sessions here refers to training sessions on separate days.
 - c. The “optimal” number of training sessions depends on the training task. Tasks with greater difficulty may require a greater number of training sessions than easier tasks.

Note: The optimal number of training sessions may be determined in a pilot experiment in which training sessions are conducted until subjects reach a performance plateau meaning that further training does not yield any greater performance improvements.

2. Number of trials per training session
 - a. VPL may be more dependent on the number of training sessions (that is, the total number of training sessions conducted on separate days) than the duration (i.e., the number of trials) within a training session (see Amar-Halpert et al., 2017; Shmuel et al., 2020). In general, we recommend keeping the duration of a training session around 45 min to 1 h to avoid subject fatigue and disengagement. We did not include explicit breaks in our training paradigm. However, the training was self-paced, meaning that subjects could complete each training session using their own speed. With a self-paced training design, subjects have the opportunity to take a break after each training trial. If explicit breaks are included in the training paradigm we recommend limiting the breaks to a few minutes to minimize the occurrence

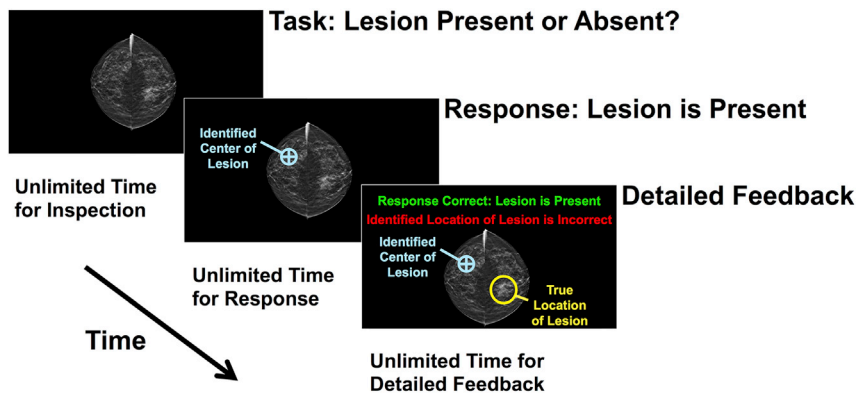


Figure 3. Example training trial

Subjects were shown a mammogram from a patient consisting of representative slices of the left and right breasts side-by-side. Subjects examined the mammogram for the presence or absence of a lesion (either calcification or distortion lesion in different subject groups). Subjects were given as much time as needed to respond. They could zoom into the mammogram to magnify anatomical details or zoom out. Furthermore, they could move the mammogram in different directions (left, right, up, down). Subjects responded by pressing one of two keys on the computer keyboard for lesion present or absent. If subjects indicated the presence of a lesion, they were further asked to indicate the center of the lesion. Therefore, subjects moved the cursor to the center of the lesion and clicked on the center with the cursor. Subjects could change their decision until making a final confirmation. Then, subjects were presented with detailed feedback about response accuracy. This feedback consisted of a written statement that could read, depending on the subject response and the presence or absence of a lesion: "Response correct: lesion is present." (corresponding to a hit, printed in green color). "Response correct: lesion is not present." (corresponding to a correct rejection, printed in green color). "Response incorrect: lesion is present." (corresponding to a miss, printed in red color). "Response incorrect: lesion is not present." (corresponding to a false alarm, printed in red color). Furthermore, subjects were shown the examined mammogram one more time with the indicated center of the lesion (shown by blue crosshair) and the true location of the lesion (shown by yellow circle if a lesion was present). In case of a hit the blue crosshair was inside the yellow circle. In case of a miss the mammogram was shown with the yellow circle only if the subject responded that a lesion was absent. If the subject responded that a lesion was present but the indicated center of the lesion was outside the true location of the lesion, this response was also considered a miss and the mammogram was presented with the blue crosshair and the yellow circle during detailed feedback. If the subject response was a correct rejection, the mammogram was shown one more time without any crosshair or circle. In case of a false alarm the mammogram was shown with the blue crosshair for the indicated center of the lesion but without any yellow circle during detailed feedback.

of stabilization of learning prior to the break, which could induce interference on new learning after the break.

3. Feedback signals during training
 - a. Feedback about response accuracy speeds up VPL (Fahle and Edelman, 1993) and reduces variability in the development and speed of VPL between different subjects (Herzog and Fahle, 1997). For complex visual stimuli such as lesions in mammograms response feedback might even be necessary to develop long-lasting VPL (Frank et al., 2020a).
4. Reinforcement signals during training
 - a. VPL is facilitated by reinforcement signals (Seitz and Watanabe, 2003; Law and Gold, 2009; Roelfsema et al., 2010). Reinforcement signals can be induced by reward during training (e.g., Seitz et al., 2009). Such rewards facilitate VPL and also motivate subjects during training.
5. A night of continuous sleep between successive training sessions
 - a. A night of continuous sleep between successive training sessions facilitates VPL by consolidating VPL against interference from other stimuli or tasks (e.g., Tamaki et al., 2020). Therefore, successive training sessions should be conducted on separate days with a night of continuous sleep in between.

Note: In our experience it is also beneficial to space successive training sessions not only by one night of sleep but by a total of two to three days.

6. Subject expertise

- a. Experts may show better performance than novices on a given task if this task is related to their expertise. Action video gamers may show faster VPL compared with nonvideo game players (e.g., [Bejjanki et al., 2014](#)). Therefore, expertise prior to the experiment may strongly influence the speed and potentially also the amount of VPL. For this reason, subjects with action video game or other types of expertise related to the given VPL training task are often excluded prior to participation in the experiment to make learning results from different subjects more comparable.

Note: We suggest that VPL for complex visual tasks should include several training sessions spaced by at least one night of continuous sleep between successive training sessions. Similarly, pretest and posttest sessions should be conducted with at least one night of continuous sleep before the first and after the last training sessions, respectively. Furthermore, detailed feedback about response accuracy should be provided during training to facilitate the development of VPL and to reduce variability in the speed and amount of VPL between different subjects. As a general recommendation, we suggest that a training session should not exceed the duration of ~45 min to 1 h with breaks to avoid subject fatigue and disengagement.

Instructions prior to training

- △ **CRITICAL:** An important component in training is the amount of practice or instruction prior to training. In the study by [Frank et al. \(2020a\)](#) we have used highly informative instruction slides (see [Figure 4](#)) to familiarize subjects with the diagnostic features of the different types of lesions used for training. These instruction slides were created by an expert (E.L.R., M.D., > 20 years of experience in radiology)

7. Show the instructions to subjects prior to each test session of the experiment (see below). This instruction, especially prior to the pretest, is critical to the success of the experiment, because subjects need to have the necessary information to do the task. In addition to the instruction slides we gave the following verbal instruction to subjects in the experiment:

- a. **Pretest instructions:** Dear participant, welcome to the experiment! In the following you will see multiple mammograms from different patients, some of which contain a lesion [either a calcification or a distortion lesion depending on subject groups]. Such a lesion can be an indication of breast cancer. It is therefore important to identify the lesion in a mammographic screening. Please be aware that the majority of the mammograms does not contain any lesion. If there is a lesion, it will be [either a calcification lesion, defined as fine white specks, tightly grouped together, like a cloud of small white dots, or a distortion lesion, defined as lines radiating to a central point, similar to the spokes of a wheel, looking like a star or a cartoon depiction of the sun, distorting the regular and coherent architecture of a breast]. If a lesion is present, the lesion will be contained only in one of the breasts. Sometimes you might see geometric shapes in a mammogram such as a circle or a triangle. Those are markers and do not reflect any abnormality.

Please review each mammogram carefully. You can take as much time as you need to examine the mammogram. Move the mammogram in different directions using the up/down/left/right arrows on the keyboard. You can and should magnify anatomical details in the mammogram, because it makes the detection of a lesion easier. Press the “k” key on the keyboard to zoom in and the “m” key to zoom out. If you think that a lesion is present, press “j” and if you think no lesion is present, press “n” on the keyboard.

All buttons with assigned functions can be found on the printed sheet next to the keyboard for you to review during the experiment.

If you respond that a lesion is present you will be asked to indicate the center of the lesion by mouse clicking on the center of the lesion in the mammogram. A blue crosshair will show up when you click on the center. You can change the location of the center by dragging the

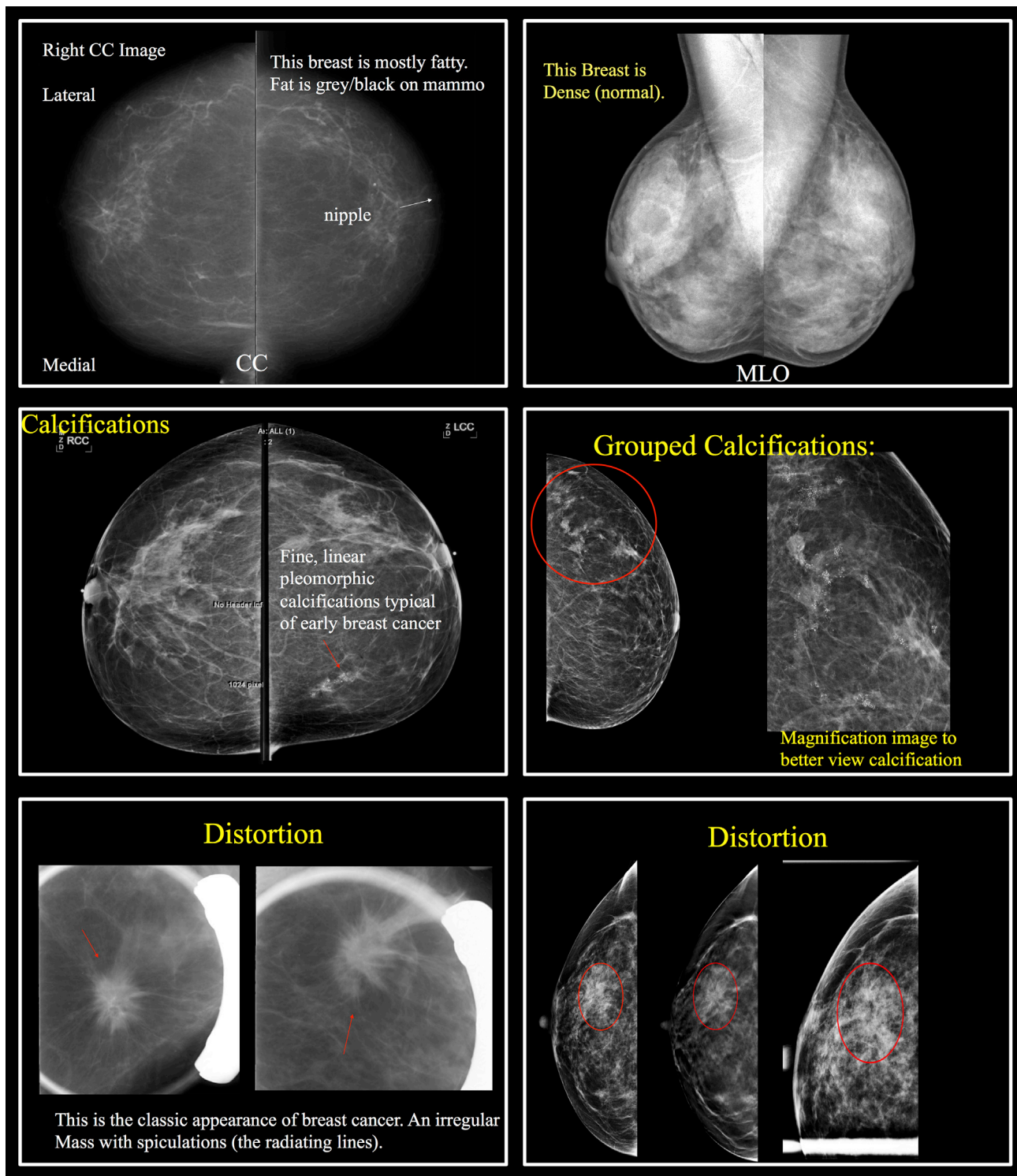


Figure 4. Example introductory slides on screening mammography

The slides were created by E.L.R., a radiology expert with more than 20 years of experience in screening mammography.

crosshair to a new location. Also, if you click on the new location the crosshair will show up at this location. Please take as much time as you need to indicate the center of the lesion.

In this task, only response accuracy matters, reaction time is not relevant. Therefore, please try to respond as accurately as possible. If you responded that a lesion is present but you want to change your decision, you can do so any time by pressing “n” for lesion absent on the keyboard. Confirm your final decision for the center of the lesion by pressing space bar. You can then move on to the next trial with a different mammogram. If you need a break, you can wait as long as you wish after the end of each trial before moving on to the next trial. Again, please bear in mind that the majority of mammograms will not contain any lesion.”

Note: A similar but shortened instruction is given to subjects during posttest and retest.

- b. **Training instructions.** Same instructions as during Pretest except that subjects are explicitly instructed to examine the detailed feedback slide at trial end: “At the end of each trial you will be presented with feedback about your response on this trial. You will learn whether your response on this trial was correct or incorrect and you will be shown again the mammogram you examined on this trial. The feedback mammogram will include the true location of the lesion (if any) enclosed by a yellow circle. Your blue crosshair response (if any) will also be shown on the mammogram in addition to the yellow circle for comparison. Please take as much time as you need to examine the feedback slide. You can move the mammogram during feedback in different directions and zoom in and zoom out as during the trial. Please zoom in to magnify anatomical details of the lesion (if present).”

Training regimen design

As shown in [Figure 2](#) (a follow-up of [Frank et al., 2020a](#)), the experiment consists of “pretest,” “posttest” and “retest” sessions, as well as several “training” sessions between pretest and posttest. Each session is conducted on a separate day. During pretest, subjects’ baseline performance on the task is assessed. During posttest, subjects’ VPL as a result of training is assessed. Pretest and posttest are conducted before the first and after the last training sessions, respectively. Improvements in detection performance for the trained type of lesion should be long-lasting without any need for further practice. To test whether this is the case, a retest session of trained subjects should be conducted at least several weeks, or months, or even years after the posttest. Subjects should not be exposed to the trained stimuli or perform the training task in between the posttest and the retest to avoid any confounds in the assessment of the stability of VPL over time without any further practice. During each test session, no feedback about response accuracy is provided to subjects to avoid any further feedback-guided learning during test sessions. In each training session subjects are trained on the learning task with detailed feedback about response accuracy after each training trial ([Figure 3](#)). The number of training sessions was kept at a minimum in the study by [Frank et al. \(2020a\)](#) to make the training applicable to the busy schedules of professional radiologists. However, greater numbers of training sessions are predicted to yield greater amounts of VPL.

In [Frank et al. \(2020a\)](#), training was conducted on a standard computer system, consisting of a computer, screen (LCD monitor, screen resolution 1,680 × 1,050 pixels; image resolution: 2,000 × 1,125), viewing distance: 30 cm (but could be smaller or greater since no chin rest was used and subjects could adjust their seating position during the experiment), keyboard and mouse ([Figure 5](#)). The room lights were turned off during the experiment and no chin rest was used. Subjects could search each mammogram for the presence or absence of a lesion by making eye movements (overt search).

Note: All mammograms in the study by [Frank et al. \(2020a\)](#) came from the same pool. That is, each subject was exposed to the same set of mammograms, however, each mammogram was randomly sorted into one of the sessions (pretest, three training sessions, posttest, retest) prior to the experiment to avoid any order effects or any interactions between the mammograms and different sessions of the experiment.



Figure 5. Experimental setup

A standard computer system, consisting of a computer, screen, mouse, and keyboard was used. For this photo the room lights were turned on. However, in the real experiment, the room lights were turned off and subjects were alone in the room sitting comfortably in a chair in front of the computer screen.

Note: In each session of test and training we kept the total number of mammograms with a lesion relative to the total number of normal mammograms (that is, the prevalence rate) low, to reflect the low frequency of the trained lesion in routine radiology examinations of asymptomatic women (Bird et al., 1992; Birdwell et al., 2001; Rangayyan et al., 2010; Bahl et al., 2015) and to imitate conditions in a real clinical setting (Evans et al., 2013).

Note: In Frank et al. (2020a) we examined VPL of different types of lesions (calcification and distortion lesions, see Figure 1). To this aim, we included two groups of subjects (each group trained on a different type of lesion). For pilot experiments we recommend to have separate groups of subjects for each training condition (that is, we recommend to use a between-subject design).

Note: In follow-up experiments it might be interesting to use a within-subject design and to investigate whether subjects can learn two or more types of lesions simultaneously. However, in our opinion, such as within-subject design could be problematic for pilot experiments, because interference between VPL of different types of lesions might occur, which might make it more difficult for VPL of either type of lesion to develop.

8. A broad summary of the [Frank et al. \(2020a\)](#) study is outlined below:
 - a. Subjects conduct three days of training with a pretest day before and a posttest day afterwards. There is also a retest 6 months after the posttest session. Each of these days included 40 trials (test) or 50 trials (training) with differing mammogram images.
 - b. For the pretest session, subjects are asked to indicate the center of lesion with no feedback given.
 - c. In the training session, subjects are asked to indicate the center of lesion, with detailed feedback given after each trial.
 - d. In the posttest and retest sessions, subjects are given the same regimen as the pretest.

Note: This was a between-subject design, so subjects were randomly assigned to either calcification or distortion test and training groups. Subjects were recruited from the community of Brown University.

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Original data	Frank et al. (2020a)	https://doi.org/10.17632/9szpfjyssp.1
Experimental models: organisms/strains		
Human subjects	Community of Brown University	N/A
Software and algorithms		
MATLAB 2012a	The MathWorks	https://www.mathworks.com/
Psychtoolbox 3.0.10	Brainard (1997) ; Pelli (1997)	http://psychtoolbox.org/

Step-by-step method details

Pretest (day 1)

⌚ Timing: 1 session of 1 h

Subjects are presented on each trial with a different mammogram and are asked to decide whether the mammogram contains the trained type of lesion.

1. Instruction slides
 - a. Give subjects explanation as to the purpose of the study and the methods used in the study. Subjects will be committing to one session every day for 5 days and then a retest 6 months after the last session. Subjects do not require any background knowledge of reading mammograms or any medical background.
 - b. Subjects are shown instruction slides that detail the importance of the study and background information on breast cancer and mammographic lesions (see [Figure 4](#) and [Experimental design considerations](#)).
2. Pretest session
 - a. Subjects are assigned to the trained type of lesion prior to the experiment.
 - b. Mammograms with the trained type of lesion and normal mammograms are presented in random order.
 - c. Task:

- i. All the available keystrokes (moving in four directions, zoom in, zoom out) are printed in large font next to the keyboard.
- ii. Make sure subjects are familiar with the keyboard and mouse before starting by asking whether they know the available keystrokes for the study and asking them to move the cursor around on the screen. No practice trials are given. The experimenter should remain in the testing room with the subject for the first five trials to make sure that the subject understands the task and knows which buttons to press.
- iii. Subject is presented with a mammogram slide, then asked to indicate if there is a lesion using the "J" or "N" keys on the keyboard for "yes" or "no."
- iv. If the subject chooses "J," then the subject is asked to click on the center of the lesion using the mouse cursor.

Note: Upon pressing the "J" key a crosshair, enclosed by a circle, will appear on the screen and the subject can adjust to where they believe the center of the lesion is located by moving and clicking the cursor. The circle has a diameter of 2 degrees visual angle. The subject cannot adjust the circle size to indicate the lesion area. If the subject zooms into the mammogram to magnify anatomical details the crosshair and circle are similarly magnified. The subject can still change the decision by pressing the "N" key.

Note: Future studies may consider including a circle adjustable in size such that subjects can indicate the lesion area.

- v. Subject presses space to confirm decision, no feedback is provided, and the next mammogram slide is shown.

Note: The difficulty level of different mammograms with and without any lesion was not equated across pretest, training, posttest, and retest. Instead, mammograms were randomly assigned to different sessions for each subject to cancel out differences in difficulty between different mammograms across subjects. Future studies should consider piloting all mammograms in the study in a different group of subjects to calculate the difficulty to detect the presence or absence of a lesion in each mammogram prior to test and training.

Training (days 2–4)

⌚ Timing: 3 sessions of 1 h per day over 3 days

Each training session is conducted on a separate day. Subjects are presented with a different mammogram on each trial and are asked to decide whether the mammogram contains the trained type of lesion.

Note: In the training phase, each trial is immediately followed by detailed feedback about response accuracy (see below and [Figure 3](#)).

Note: Subjects are notified which lesion they are assigned for training before the training session starts.

Note: A set of new, previously unseen mammograms is used for each training session.

3. Instruction
 - a. Subjects are told that there will be detailed feedback about response accuracy after each slide.
4. Training session
 - a. Mammograms with the trained type of lesion and normal mammograms are presented in random order.

- b. Subjects are asked to indicate whether there is a lesion or not using the same steps as in the pretest.
- c. If there is no lesion in the mammogram:
 - i. Subjects will be notified with green text if they responded correctly (correct rejection), and red text if they responded incorrectly (false alarm).
 - ii. The mammogram will be shown again so that the subject can review the slide. Subjects can control how long they want to review the feedback slide and press space to move on from the feedback slide when they are ready. There is nothing indicated on the slide, except a blue crosshair of the indicated center of the lesion if the subject committed a false alarm.
- d. If there is a lesion in the mammogram:
 - i. Subjects will be notified with green text if they responded correctly (hit), and red text if they responded incorrectly (miss).
 - ii. The mammogram will be shown again with the area of the lesion circled in yellow (corresponding to the true location of the lesion) and the blue crosshair corresponding to the center of the lesion reported by the subject (if the subject indicated the presence of a lesion).

Note: When indicating where the lesion is, subjects can click anywhere within the yellow-circled area (Figure 1) and be counted as correct; the subject does not need to click exactly in the center of the lesion.

- e. Subject presses space to confirm when finished with review, and the next mammogram slide is shown for the next trial of training.

△ **CRITICAL:** Subjects are to be reminded before beginning of training to examine the feedback image for every slide carefully, for both normal mammograms and lesions. Furthermore, subjects are to be informed that the percentage of normal mammograms is a lot higher than of mammograms with the trained type of lesion.

Posttest (day 5)

⌚ **Timing:** 1 session of 1 h

Posttest is identical with the pretest except that a set of new, previously unseen mammograms is used.

5. Instruction slides
 - a. Subjects are again familiarized with the diagnostic features of the trained type of lesion by using the same instruction slides as prior to the pretest.
 - b. Subjects are told that there will not be feedback after each slide during the test.
6. Posttest session
 - a. Exactly the same procedures as in the pretest session are used.

Retest (6 months after posttest)

⌚ **Timing:** 1 session of 1 h

The long-term stability of VPL is assessed by conducting a retest session, which is identical with the pretest and posttest sessions except that a set of new, previously unseen mammograms is used.

△ **CRITICAL:** Only trained subjects who completed each of the preceding sessions of the experiment should be recruited for the retest.

7. Instruction slides

- a. Subjects are asked to review all slides in the instruction slides from the pretest session.
8. Retest session
 - a. Exactly the same procedures as in the pretest and posttest sessions are used.

Expected outcomes

Successful VPL on this task should lead to improved detection of the trained type of lesion, if it is present in a mammogram, in the posttest, as shown by an increase in the number of hits compared with the pretest. Furthermore, the erroneous detection of the trained type of lesion, if it is not present in a mammogram, should decrease, as shown by a decrease in the number of false alarms in the posttest compared with the pretest. In the study by [Frank et al. \(2020a\)](#), across a group of 12 subjects who trained with detailed feedback on calcification lesions, we observed a mean \pm standard error of the mean (SEM) number of 2.83 \pm 0.55 hits for calcification lesions during pretest, which increased to 4.58 \pm 0.73 hits during posttest. The mean \pm SEM number of false alarms decreased from 1.83 \pm 0.58 to 0.92 \pm 0.50 from pretest to posttest (see Figure S3A in [Frank et al., 2020a](#)). For a different group of 12 subjects who trained with detailed feedback on distortion lesions, the mean \pm SEM number of hits increased from 1.92 \pm 0.36 during pretest to 4.33 \pm 0.63 during posttest and the mean \pm SEM number of false alarms decreased from 6.00 \pm 1.03 during pretest to 2.08 \pm 0.61 during posttest (see Figure S3B in [Frank et al., 2020a](#)). Hit and false alarm rates in each session of the experiment can be combined into an observer sensitivity score (referred to as d') for further statistical analyses.

[Frank et al. \(2020a\)](#) demonstrate that detailed trial-by-trial feedback about response accuracy is necessary to achieve long-lasting VPL for the trained type of lesion ([Figure 6](#)). In addition, in the absence of feedback, no significant learning occurred for either type of lesion across subjects. Furthermore, the results of this study showed that the difficulty to detect the two different types of lesions varied. Specifically, due to their slightly more distinct visual features, calcification lesions were easier to detect than distortion lesions across sessions (see [Figure 6](#)). Therefore, greater observer sensitivity is expected for calcification than distortion lesions across test and training sessions.

Limitations

Although feedback-guided VPL may be a cost-effective training especially in fields that require visual expertise (such as radiology), there are clear limitations. [Frank et al. \(2020a\)](#) conducted the study with novice viewers who have little background knowledge about mammograms and mammographic lesions beyond the instruction slides, not with radiologists that have seen many similar mammographic images and are aiming and motivated to make a diagnosis. Therefore, it is necessary to investigate in future studies whether feedback-guided VPL is similarly effective in improving visual detection skills in more experienced subjects. Furthermore, only two types of mammographic lesions (calcification and distortion lesions) were examined in the study by [Frank et al. \(2020a\)](#). It is important to investigate whether feedback-guided VPL also improves the detection of other types of mammographic lesions such as masses, asymmetries, and focal asymmetries. Another limitation of this study is that subjects were presented with the most representative images of the left and right breasts, whereas in breast tomosynthesis many images are collected, which the radiologist has to search through image-by-image to make a decision about the presence or absence of a lesion. In addition, the current study was conducted in a psychophysics laboratory and it is necessary to determine in future investigations whether VPL works similarly in a real clinical setting and on professional workstations (see also [Evans et al., 2013](#)).

Troubleshooting

Problem 1

Subject pace is extremely fast and finishes each session quickly (steps 2, 4, 6, and 8).

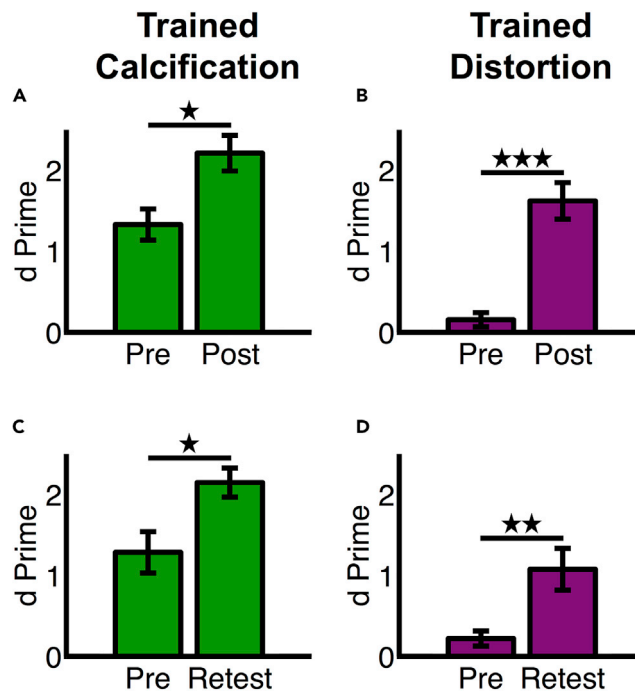


Figure 6. Results of the study

(A) Mean \pm standard error of the mean (SEM) observer sensitivity (d') in the pretest (Pre) and posttest (Post) for a group of 12 subjects who trained on calcification lesions with detailed feedback. The asterisk shows that the increase in d' from pretest to posttest was statistically significant (paired-sample t test between posttest and pretest). * $p < 0.05$.

(B) Same as (A), but for a different group of 12 subjects who trained on distortion lesions with detailed feedback. *** $p < 0.001$.

(C) Same as (A), but for the subgroup of 9 subjects trained on calcification lesions from (A) who were available for a retest session (Retest) on calcification lesions six months after the posttest. The asterisk shows that the increase in d' from pretest to retest was statistically significant (paired-sample t test between retest and pretest). No significant difference in d' was observed between retest and posttest ($p > 0.05$). This indicates that subjects' performance improvements from pretest to posttest were long-lasting. * $p < 0.05$.

(D) Same as (C), but for the subgroup of 9 subjects trained on distortion lesions from (B) who were available for a retest session on distortion lesions six months after the posttest. Subjects' performance during retest was significantly greater compared with pretest. However, performance during retest decreased significantly than during posttest ($p < 0.05$). This indicates that performance improvements were partially long-lasting. The slight decrease in performance from posttest to retest might have occurred due to the greater difficulty to detect distortion lesions than calcification lesions. ** $p < 0.01$. For further details see Frank et al. (2020a).

Potential solution

It is crucial that the subject is taking the time to analyze each mammogram slide and answering to the best of their ability. For example, if a subject quickly moves past each slide on to the next training or test trial, then it is highly unlikely that any perceptual learning will occur because of a lack of attention to the stimulus and task. To avoid this issue, the researcher needs to emphasize the importance of taking time during the instruction slides prior to pretest to become familiar with the distinct visual features of the trained type of lesion (since the subject has no previous exposure to this lesion or mammograms more generally). Furthermore, the researcher should remind subjects prior to each session of the experiment that there is no time constraint and that subjects should take as much time as necessary on each trial to respond as accurately as possible. In the study by Frank et al. (2020a) we found that the mean \pm SEM response time for lesion present or absent (calculated as the median response time across all trials per session for each subject and averaged across subjects) was 21.8 ± 4.37 s during pretest, 18.9 ± 4.01 s during posttest and 17.1 ± 2.15 s during retest for the calcification training group with detailed feedback during training (12 subjects for pretest and

posttest, 9 subjects for retest). The mean \pm SEM response time for the distortion training group with detailed feedback during training was 12.9 ± 1.85 s during pretest, 13.3 ± 1.70 s during posttest and 10.1 ± 2.11 s during retest (12 subjects for pretest and posttest, 9 subjects for retest). If only partial feedback about response accuracy was provided (meaning that only a written statement about response accuracy was given for feedback without any opportunity to review the examined mammogram one more time), then the results for the calcification/distortion training groups were 16.4 ± 3.16 s/ 14.1 ± 3.12 s for pretest and 10.4 ± 1.68 s/ 13.9 ± 2.67 s for posttest (12 subjects for each training group). Subjects trained on calcifications achieved 6.92 ± 1.04 s during retest (9 subjects). When no feedback about response accuracy was provided, then subjects in the calcification/distortion training groups achieved mean response times of 13.3 ± 2.28 s/ 15.1 ± 3.38 s during pretest and 9.79 ± 2.06 s/ 8.79 ± 1.74 s during posttest (12 subjects per training group).

Problem 2

Subject pace is extremely slow and finishes each session slowly (steps 2, 4, 6, and 8).

Potential solution

As mentioned before, a training session should not exceed 45 min to 1 h to avoid exhaustion and disengagement. If a subject is taking too long on the slides, then although they are putting forth focus and effort, they may lack the stamina to continue through all 50 slides. To avoid this issue, the researcher needs to make clear to the subject that there will be 50 slides total per training session, and to answer as accurately as possible on all of the slides. The researcher should remind subjects of this prior to each session of the experiment and that they have as much time necessary on each trial to respond as accurately as possible. It is important here to not rush the subject if their pacing is too slow, because only the subject can personally know their own stamina and prompting them to rush through the rest of the slides can result in less accurate data and more confounding variables. It is a suggestion that training sessions should not exceed 1 h; however, this is not a fixed rule and it is possible for subjects to extend past this if they feel that the time is necessary for them to fully analyze each slide.

Problem 3

Subject forgets which keys on the keyboard correspond to which action (steps 2, 4, 6, and 8).

Potential solution

The test and training protocols involve many keystrokes, and only a few are listed on the choice screen during the session. If a subject does not remember that they can zoom in or zoom out, this can affect potential perceptual learning and the accuracy of the subject's results. Therefore, it is important to tell the subject about the keystroke options available before the study and provide a printout list of all the choices available and their corresponding actions. The researcher can also stay in the room for the first mammogram session in the pretest to verify the options available.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Takeo Watanabe (takeo_watanabe@brown.edu).

Materials availability

Instruction slides for subjects are available from the Lead Contact upon request.

Data and code availability

Original data from Frank et al. (2020a) are available (Mendeley DOI: <https://doi.org/10.17632/9szpfjyssp.1>). This study used standard, custom-built MATLAB programmed scripts that are available from the Lead Contact upon request.

Acknowledgments

This work was supported by NIH R21EY028329, NIH R01EY019466, NIH R01EY027841, United States – Israel Binational Science Foundation BSF2016058.

Author contributions

S.M.F., E.L.R., and T.W. designed the study. E.L.R. provided and classified mammograms. S.M.F., A.Q., and D.R. collected data. S.M.F. analyzed the data. All authors wrote the manuscript.

Declaration of interests

The authors declare no competing interests.

References

- Amar-Halpern, R., Laor-Maayan, R., Nemni, S., Rosenblatt, J.D., and Censor, N. (2017). Memory reactivation improves visual perception. *Nat. Neurosci.* *20*, 1325–1328.
- Bahl, M., Baker, J.A., Kinsey, E.N., and Ghate, S.V. (2015). Architectural distortion on mammography: correlation with pathologic outcomes and predictors of malignancy. *Am. J. Roentgenol.* *205*, 1339–1345.
- Ball, K., and Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science* *218*, 697–698.
- Bang, J.W., Shibata, K., Frank, S.M., Walsh, E.G., Greenlee, M.W., Watanabe, T., and Sasaki, Y. (2018). Consolidation and reconsolidation share behavioural and neurochemical mechanisms. *Nat. Hum. Behav.* *2*, 507–513.
- Bejjanki, V.R., Zhang, R., Li, R., Pouget, A., Green, C.S., Lu, Z.L., and Bavelier, D. (2014). Action video game play facilitates the development of better perceptual templates. *Proc. Natl. Acad. Sci. U S A* *111*, 16961–16966.
- Bird, R.E., Wallace, T.W., and Yankaskas, B.C. (1992). Analysis of cancers missed at screening mammography. *Radiology* *184*, 613–617.
- Birdwell, R.L., Ikeda, D.M., O'Shaughnessy, K.F., and Sickles, E.A. (2001). Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection. *Radiology* *219*, 192–202.
- Brainard, D.H. (1997). The psychophysics toolbox. *Spat. Vis.* *10*, 433–436.
- Ding, J., and Levi, D.M. (2011). Recovery of stereopsis through perceptual learning in human adults with abnormal binocular vision. *Proc. Natl. Acad. Sci. U S A* *108*, E733–E741.
- Dosher, B., and Lu, Z.L. (2017). Visual perceptual learning and models. *Annu. Rev. Vis. Sci.* *3*, 343–363.
- Evans, K.K., Birdwell, R.L., and Wolfe, J.M. (2013). If you don't find it often, you often don't find it: why some cancers are missed in breast cancer screening. *PLoS One* *8*, e64366.
- Fahle, M., and Edelman, S. (1993). Long-term learning in vernier acuity: effects of stimulus orientation, range and of feedback. *Vis. Res.* *33*, 397–412.
- Frank, S.M., Bründl, S., Frank, U.I., Sasaki, Y., Greenlee, M.W., and Watanabe, T. (2020b). Fundamental differences in visual perceptual learning between children and adults. *Curr. Biol.* <https://doi.org/10.1016/j.cub.2020.10.047>.
- Frank, S.M., Greenlee, M.W., and Tse, P.U. (2018). Long time no see: enduring behavioral and neuronal changes in perceptual learning of motion trajectories 3 years after training. *Cereb. Cortex* *28*, 1260–1271.
- Frank, S.M., Qi, A., Ravasio, D., Sasaki, Y., Rosen, E.L., and Watanabe, T. (2020a). Supervised learning occurs in visual perceptual learning of complex natural images. *Curr. Biol.* *30*, 2995–3000.
- Frank, S.M., Reavis, E.A., Greenlee, M.W., and Tse, P.U. (2016). Pretraining cortical thickness predicts subsequent perceptual learning rate in a visual search task. *Cereb. Cortex* *26*, 1211–1220.
- Frank, S.M., Reavis, E.A., Tse, P.U., and Greenlee, M.W. (2014). Neural mechanisms of feature conjunction learning: enduring changes in occipital cortex after a week of training. *Hum. Brain Mapp.* *35*, 1201–1211.
- Herzog, M.H., and Fahle, M. (1997). The role of feedback in learning a vernier discrimination task. *Vis. Res.* *37*, 2133–2141.
- Hussain, Z., Webb, B.S., Astle, A.T., and McGraw, P.V. (2012). Perceptual learning reduces crowding in amblyopia and in the normal periphery. *J. Neurosci.* *32*, 474–480.
- Karni, A., and Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proc. Natl. Acad. Sci. U S A* *88*, 4966–4970.
- Karni, A., and Sagi, D. (1993). The time course of learning a visual skill. *Nature* *365*, 250–252.
- Law, C.T., and Gold, J.I. (2009). Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nat. Neurosci.* *12*, 655–663.
- Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* *10*, 437–442.
- Polat, U., Ma-Naim, T., Belkin, M., and Sagi, D. (2004). Improving vision in adult amblyopia by perceptual learning. *Proc. Natl. Acad. Sci. U S A* *101*, 6692–6697.
- Rangayyan, R.M., Banik, S., and Desautels, J.L. (2010). Computer-aided detection of architectural distortion in prior mammograms of interval cancer. *J. Digit. Imaging* *23*, 611–631.
- Roelfsema, P.R., van Ooyen, A., and Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends Cogn. Sci.* *14*, 64–71.
- Sagi, D. (2011). Perceptual learning in vision research. *Vis. Res.* *51*, 1552–1566.
- Sasaki, Y., Nanez, J.E., and Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nat. Rev. Neurosci.* *11*, 53–60.
- Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature* *412*, 549–553.
- Seitz, A.R., and Dinse, H.R. (2007). A common framework for perceptual learning. *Curr. Opin. Neurobiol.* *17*, 148–153.
- Seitz, A.R., Kim, D., and Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron* *61*, 700–707.
- Seitz, A.R., and Watanabe, T. (2003). Is subliminal learning really passive? *Nature* *422*, 36.
- Shmuel, D., Frank, S.M., Sharon, H., Sasaki, Y., Watanabe, T., and Censor, N. (2020). Early visual cortex stimulation modifies well-consolidated perceptual gains. *Cereb. Cortex*, bhaa215.
- Tamaki, M., Wang, Z., Barnes-Diana, T., Guo, D., Berard, A.V., Walsh, E., Watanabe, T., and Sasaki, Y. (2020). Complementary contributions of non-REM and REM sleep to visual learning. *Nat. Neurosci.* *23*, 1150–1156.