Applied Cognitive Psychology, Appl. Cognit. Psychol. **30**: 750–756 (2016)
Published online 19 July 2016 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/acp.3250

Training Melanoma Detection in Photographs Using the Perceptual Expertise Training Approach

BUYUN XU[1], LIAM ROURKE[2], JUNE K. ROBINSON[3] and JAMES W. TANAKA[1]*

[1]*Department of Psychology, University of Victoria, Canada*
[2]*Department of Medicine, University of Alberta, Canada*
[3]*Department of Dermatology, Northwestern University, USA*

Summary: Although a deadly form of skin cancer, melanoma is treatable if detected early. However, current rule-based training practices in melanoma detection are not effective. We assessed an innovative technique to train melanoma detection using the perceptual expertise principles. Participants in the training group were trained to categorize melanoma and benign lesions to 95% accuracy. Participants in the control group received no training. Prior to testing all participants reviewed the ABCDE rules. Training was evaluated by the pre and post tests using the Melanoma Detection Test where participants categorized images of melanoma and benign lesions. As compared to the control group, the training group showed significant improvement in melanoma detection and became less liberal (i.e., bias toward categorizing a lesion as melanoma), and both improvements maintained a week after the training. These findings indicate that perceptual expertise training is a promising approach to train melanoma detection.

INTRODUCTION

Although a potentially deadly form of skin cancer, melanoma is also one of the most treatable if the cancerous lesion is detected at an early stage. If diagnosed at an early stage when the lesion is limited to penetration of the tumor to <1.0 mm below the surface of the skin, the five-year relative survival rate for patients with melanoma is high (98.2%) (American Cancer Society, 2015). However, the survival outcome drops significantly to 59% if the cancer spreads to lymph nodes and only 15.5% if the cancer is detected after spreading to lymph nodes and internal organs such as the liver, lung or brain (Howlander et al., 2012). Therefore, an important strategy for reducing melanoma mortality rates is to improve the early detection, diagnosis and treatment of the cancer through medical training and education.

The conventional approach for teaching the diagnostic features of skin cancer is the ABCDE system where the participant is directed to five key features of a lesion regarding its asymmetrical shape (A), irregular border (B), variegated color (C), a size that is greater than 6 mm in diameter (D) and evolving appearance (E) (Figure 1). A skin lesion meeting the ABCDE diagnostic criteria is considered suspicious and the patient is referred to a dermatologist for further evaluation. Unfortunately, training programs designed to improve ABCDE diagnostic skills of physicians, medical students and practitioners have been largely ineffective (Rourke, Oberholtzer, Chatterley, & Brassard, 2015). Studies have shown that at-risk patients (Oliveria, Chau, Christos, Mushlin, & Halpern, 2004), medical students (Aldridge, Maxwell, & Rees, 2012) and primary care physicians (Chen, Bravata, Weil, & Olkin, 2001) perform poorly in discriminating melanoma from benign lesions. One limitation is that learners have difficulty judging the ABCDE features reliably and accurately. For example, one study showed that, among three independent observers, moderate inter-observer agreement was found in judging the lesion's color variegation (k statistics = 0.43) and pigment diffusion (k statistics = 0.58), and poor inter-observer agreement was found in judging the lesion's border irregularity (k statistics = 0.13) and contour irregularity (k statistics = −0.03) (Meyer et al., 1996).

Another weakness of the ABCDE approach is that benign lesions—lesions that are not melanomas—may present with ABCDE features. Common and harmless lesions like seborrheic keratoses or dysplastic nevi may be asymmetrical, have jagged borders, be multi-colored and be larger than 6 mm. Conversely, many melanoma lesions in their early stages—when they are curable—are missed by the ABCDE criteria. Melanoma and benign lesions belong to 'fuzzy' categories where the visual features are overlapping and probabilistic rather than exclusive and defining (Ashby & O'Brien, 2005; Rosch, 1973; Rosch & Mervis, 1975; Zadeh, 1965). In contrast to rule-based categories, fuzzy categories are acquired through perceptual expertise training where learners are exposed to a broad range of category examples and practice classifying the examples with supervised feedback (Posner & Keele, 1968).

Learning to discriminate different types of fuzzy categories involves a kind of perceptual expertise—the ability to make fast, accurate identifications of objects at specific levels of categorization (Tanaka & Taylor, 1991). The seasoned bird watcher, for example, quickly distinguishes a Nashville warbler from the closely related Tennessee warbler. The car enthusiast instantly knows the subtle differences between a 1955 Chevy Belaire from the 1956 model. Perceptual expertise also requires holistic perception where an object is not recognized by its individual features but by the synthesis of the features into a whole. Common to bird experts (Hagen, Vuong, Scott, Curran, & Tanaka, 2014; Hagen, Vuong, Scott, Curran, & Tanaka, 2015; Johnson & Mervis, 1997; Tanaka & Taylor, 1991), dog experts (Tanaka & Taylor, 1991) and car experts (Gauthier, Skudlarski, Gore, & Anderson, 2000) is the fast, specific

*Correspondence to: James Tanaka, Department of Psychology, University of Victoria, Victoria, BC, Canada.
E-mail: jtanaka@uvic.ca

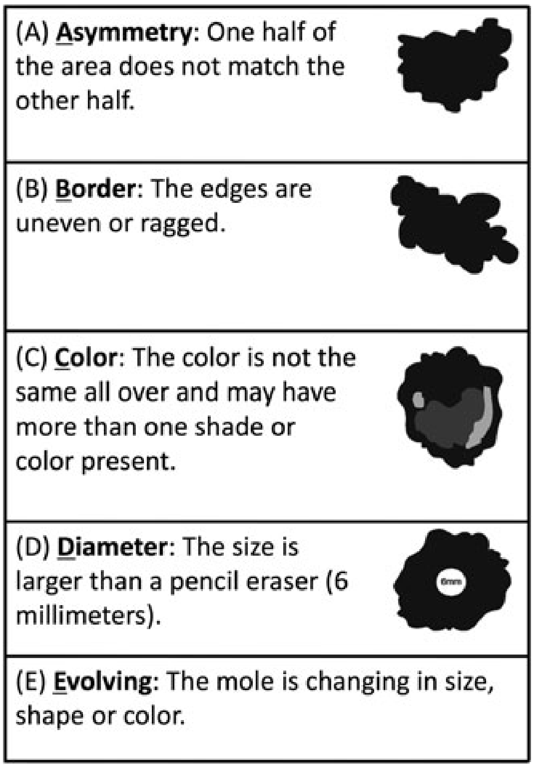

Figure 1. Images from an ABCDE instructional pamphlet describing how to detect a melanoma by its Asymmetry, Border, Color, Diameter and Evolving appearance

and holistic recognition of objects in their domain of expertise. According to their self-reports, experienced dermatologists, like other types of perceptual experts, employ holistic processes for categorizing skin lesions (Gachon et al., 2005).

An important question is whether the perceptual expertise for melanoma detection can be acquired through systematic training. Although real-world perceptual experts acquire their expertise over years or even decades of practice, it is possible to train expert-like performance in the laboratory under a shorter, more compressed time schedule. Following the protocol of past laboratory training studies (Gauthier, Williams, Tarr, & Tanaka, 1998; Scott, Tanaka, Sheinberg, & Curran, 2008; Tanaka, Curran, & Sheinberg, 2005; Tanaka & Pierce, 2009), effective perceptual expertise training requires the following active ingredients. First, participants should be exposed to multiple exemplars of each category to ensure that learning transfers from trained stimuli to novel stimuli. Second, the set of training stimuli should represent the range and frequency of exemplars of the category. That is, typical or common category exemplars should appear more frequently during training than less common exemplars. Third, expertise training should involve immediate feedback so that the learners can update their category representation after each response and performance should meet a specified criterion of mastery. For example, in a typical training trial the participant receives immediate feedback after they make a decision and progresses to new exemplars only after they achieve a specified criterion of performance (e.g., 95% correct). Finally, training should occur over multiple days to allow perceptual learning to consolidate. Studies following these rules have successfully produced expert-like recognition for birds (Tanaka et al., 2005), cars (Scott, Tanaka, Sheinberg, & Curran, 2008), artificial Greeble objects (Gauthier, Williams, Tarr & Tanaka, 1998; Gauthier et al., 1998) and faces (Heptonstall, Tanaka, & Hoven, 2013; Tanaka & Pierce, 2009). Critically, participants in these studies demonstrated reliable transfer effects to novel images and novel categories. In the study by Tanaka et al. (2005), for example, participants who learned to identify different species of owls (e.g., great grey owl, great horned owl, barred owl) demonstrated superior abilities to discriminate new images of these learned species. Moreover, perceptual performance transfers to new categories that are closely related to the trained expert categories. Participants in the owl and wading bird study, for instance, were better able to discriminate related species of owls (e.g., barn owl) that were not learned during the training (Tanaka et al., 2005). In addition, the effects of expert training are long lasting, where expert recognition was still robust a week later after training (Scott et al., 2008).

In the current study, we introduced, the first time, the perceptual expertise training approach in melanoma detection training. The purpose of the current study was to evaluate whether receiving perceptual expertise training over a relatively short period of time is effective in improving the performance of melanoma detection, as compared to participants receiving no training at all. Naïve observers were recruited to participate in perceptual expertise training protocol and another group served as the no-training control group to control for the test-retest gains (Roediger & Karpicke, 2006). As a measure of pre-training performance, participants in both groups completed the Melanoma Detection Test (MDT) to assess their pre-training ability to discriminate melanoma and benign lesions. Following completion of the pre-test, participants in the perceptual expertise training group received four sessions of training in which they learned to categorize varieties of melanomas and benign lesions. After training, the MDT was re-administered to both groups and their pre- and post-training MDT scores were compared. Based on the perceptual expertise hypothesis we predicted that, as compared to the performance change of the control group, the perceptual expertise training method could improve participants' melanoma detection. Second, given the robustness of perceptual expertise training effects (Scott et al., 2008), we predicted that participants in the perceptual expertise group will maintain their performance a week after training without the benefit of additional practice.

METHODS

Participants

Twenty-seven undergraduate students from University of Victoria participated in the study. All of the participants had normal or corrected-to-normal vision, and none of them have received formal medical training. Among them, thirteen participants (three males) were randomly assigned into the training group, and the rest 14 participants (one male) were assigned into the control group. The average age was 21.82 for the training group and 20.57 for the control group, which were not significantly different ($t_{25} = 1.49$, $p = 0.15$).

Stimuli

Four types of melanoma (acral lentiginous, lentigo maligna, nodular, superficial spreading) and four types of benign pigmented lesion (blue nevi, lentigo, melanocytic nevi, seborrheic keratoses) were used in the current study. The images were collected from dermatological image repositories (e.g., dermnetNZ.org) and from the Internet (Figure 2). All the images were validated by three expert dermatologists to ensure that they were 'classical presentations' of the lesion; that is, their appearance was typical of the lesion type. The chosen images were scaled to fit within a 470 by 470 pixel frame and cropped to remove any potentially diagnostic body part information.

Melanoma detection task

The MDT is a measure of one's ability to discriminate melanoma and benign lesions. In the MDT, six images of each of the four types of melanoma and benign lesions were included in the test (48 trials in total).[1] In each trial, the participant saw one image of a skin lesion and was asked to decide whether the lesion was 'melanoma' or 'benign' by pressing the 'm' or 'b' keys, respectively. The test recorded both accuracy and response time. None of the images used in the MDT were used in the training.

Training

In the training, participants were trained to categorize 10 exemplars in each of the four types of melanoma and benign lesions over four consecutive sessions. In each session, participants were trained to discriminate one new type of melanoma from one new type of benign lesion, and then tested on all of the previously learned melanoma and benign lesions (if applicable) and all of the new ones learned in that particular session. The four training sessions were completed in four consecutive days. Participants learned to discriminate between seborrheic keratoses and superficial spreading melanoma on day 1, melanocytic nevi and acral lentiginous melanoma on day 2, lentigo and lentigo maligna melanoma on day 3 and blue nevi and nodular melanoma on day 4.

There are four phases in each training session. The training session began with a review phase in which all the skin lesions learned in previous sessions (if applicable), were reviewed sequentially, for 1 s each, with the label (i.e., melanoma, benign) presented above the image. The presentation was blocked by lesion type (i.e., melanoma, benign) and by training sessions. For example, in the review phase of session 3, participants were first presented with all the melanoma lesions learned in session 1, then with all the benign lesions learned in session 1, then with all the melanoma lesions learned in session 2, and finally with all the benign lesions learned in session 2. The review session was followed by the introduction phase, in which the lesions to be learned in the current session were introduced to the participant. In this phase, 10 new exemplars of melanoma were presented one after another, for 1 s each, with the label of 'melanoma' above it. This was followed by the same presentation of 10 new exemplars of benign lesions with the label of 'benign' above it. The introduction phase was followed by the training phase, in which participants were trained to correctly categorize 10 exemplars of melanoma and 10 exemplars of the benign lesion, in an accumulative way across multiple rounds. To be specific, participants were presented one untrained exemplar of each type of lesion with the category label (i.e., benign or melanoma) at the beginning of each round, and were then asked to perform a melanoma vs. benign categorization task with those two exemplars mixed with all the exemplars that have been trained in the previous rounds (if applicable). For example, participants received training in round 1 and 2 with melanoma exemplar 1, benign exemplar 1, melanoma exemplar 2 and benign exemplar 2.

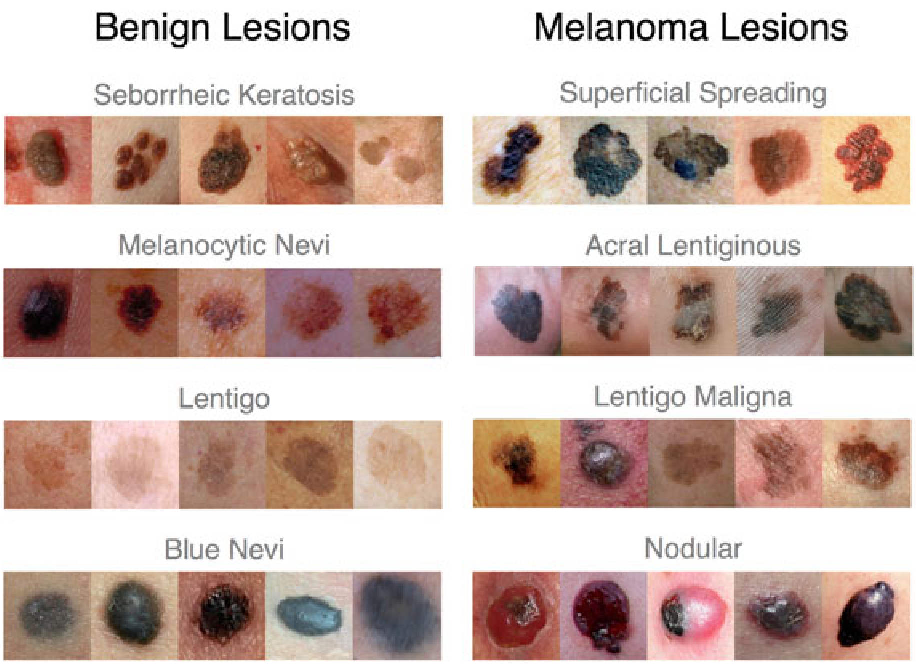

Figure 2. Examples of the four types of benign lesions (i.e., seborrheic keratosis, melanocytic nevi, lentigo and blue nevi) and the four types of melanoma lesions (i.e., superficial spreading, acral lentiginous, lentigo maligna and nodular)

In the third round, participants were first presented with melanoma exemplar 3 with the label of 'melanoma' and benign exemplar 3 with the label of 'benign', and were then trained with melanoma exemplar 1, 2 and 3 and benign exemplar 1, 2 and 3 mixed together and in a randomized order. Immediate feedback on the correctness of the response was provided in each trial. Only when the performance in the current round of training reached the 95% accuracy, would the new round of training be started; otherwise, the same round of training was repeated with all the incorrectly categorized exemplars until the criterion was met. When the participant was able to successfully categorize all 10 benign and 10 melanoma lesions, the testing phase was launched. In the testing phase, the participants were asked to categorize the 10 melanoma exemplars and 10 benign exemplars learned in the current session, mixed with all the melanoma and benign lesions learned in previous sessions, if applicable. Unless the participant's accuracy reached 95%, the testing phase was repeated with all the exemplars until the criterion was met. By the fourth and final day of training, all participants who had successfully completed the training were able to achieve 95% accuracy on all four melanoma types and all four benign types. None of the images seen during training set were used in the MDT.

PROCEDURE

On the first day, all participants were given background information by reviewing the ABCDE rules (Figure 1) for 5 min. Then, all participants took the MDT, as the pre test. After the pre test, the control group participants were sent home and were told to come back in a week. The training group remained in the laboratory to finish session 1 of the training and came back to the laboratory on days 2, 3 and 4 to finish the rest of the training sessions. Both the MDT and the training program were computerized, and ran on Dell Precision desktop computers with a 17-inch monitor. All participants finished each session of the training individually and independently. The length of the training sessions ranged from 10 min for some participants to 40 min for others.

On day 4, the control group participants returned to the laboratory. Both groups of participants completed the MDT again, as the first post test. One week later, the training group participants completed the MDT one more time as the second post-test. None of them reported reviewing any materials related to melanoma during the week. The stimuli used in the MDT were the same in pre test, post test 1 and post test 2.

RESULTS

Using correct detection of melanoma as a hit, and categorizing benign lesion as melanoma as false alarm, sensitivity (d') of the pre and post-test was calculated for both the training and control group as the difference between the Z transforms of the hit rate and the Z transforms of the false alarm rate (i.e., $d' = Z_{hit} - Z_{false\ alarm}$). A 2×2 ANOVA was conducted, with test (pre vs. post) as within-subject variable, and group (training vs. control) as between-subject variable. The main effect of test ($F_{(1,25)} = 7.16$, $p < 0.05$, $\eta_p^2 = 0.22$) was significant. More importantly, the interaction between test and group was significant ($F_{(1,25)} = 20.99$, $p < 0.001$, $\eta_p^2 = 0.46$). This interaction was driven by the significant improvement of performance in the training group ($p < 0.001$), but not in the control group ($p = 0.18$) (Figure 3).

A further investigation of the hit rate and the false alarm rate (Table 1) showed that the interaction between test and group was only significant for the false alarm rate ($F_{(1,25)} = 1.41$, $p = 0.25$, $\eta_p^2 = 0.05$), but not hit rate ($F_{(1,25)} = 13.26$, $p < 0.01$, $\eta_p^2 = 0.35$). In order to interpret this interaction in false alarm rate, multiple comparisons showed that the false alarm rate did not change between pre and post tests for the control group ($p = 0.19$), but decreased significantly after training in the training group ($p < 0.01$).

Response time of correct trials of the pre- and post-test was calculated for both the training and control group. A 2×2 ANOVA was conducted, with test (pre vs. post) as within-subject variable, and group (training vs. control) as between-subject variable. The main effect of test ($F_{(1,25)} = 17.71$, $p < 0.001$, $\eta_p^2 = 0.41$) was significant. The interaction between test and group was not significant ($F_{(1,25)} = 0.22$, $p = 0.64$, $\eta_p^2 = 0.01$), indicating that the response time in the post test is significantly faster than the pre-test in both groups ($p < 0.05$) (Figure 4).

Response criterion (c) of the pre- and post-tests was calculated for both the training and control group. A 2×2 ANOVA was conducted, with test (pre vs. post) as within-subject variable, and group (training vs. control) as between-subject variable. Both the main effect of test ($F_{(1,25)} = 2.88$, $p = 0.10$, $\eta_p^2 = 0.10$) and the interaction between test and group ($F_{(1,25)} = 2.70$, $p = 0.11$, $\eta_p^2 = 0.10$) were close to be significant. Although the two-way interaction between test and group was marginal, visual inspection suggested that the response criterion did not change in the control group, but changed in the training group. Planned multiple comparisons indicated that the response criterion did not shift in the control group ($p = 0.97$), but shifted significantly in the training group from liberal (biased to categorize lesions as melanoma) to neutral ($p < 0.05$) (Figure 5).

Sensitivity, response time of correct trials and response criterion in the first post-test and the second post-test for

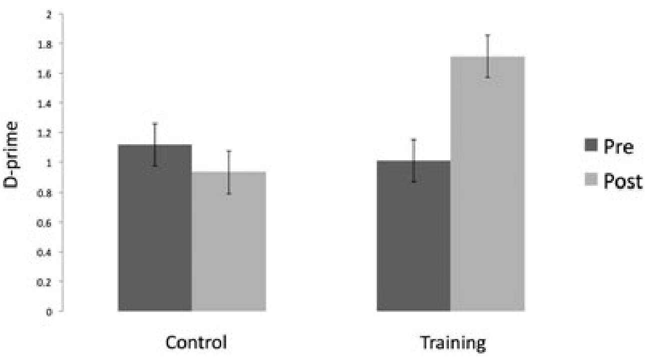

Figure 3. Sensitivity (d') of the pre and post test in the training and control group. The error bars refer to 95% confidence interval. The dark bars refer to the sensitivity in the pre test and the light bars refer to the sensitivity in the post test

Table 1. The averaged hit and false alarm rates in the pre and post test for the control and training groups. Numbers in the parentheses refer to the standard deviation

	Hit rate			False alarm rate		
	Pre-test	Post-test	*p* Value	Pre-test	Post-test	*p* Value
Control	0.83 (0.11)	0.81 (0.15)	0.673	0.45 (0.16)	0.52 (0.16)	0.193
Training	0.75 (0.17)	0.80 (0.07)	0.226	0.42 (0.13)	0.23 (0.14)	0.001

the training group were compared to test whether the train effect was retained after a week. *T*-tests showed that the training effect did not differ between the first and second post-test in the measurement of sensitivity, response time and response criterion (Table 2).

DISCUSSION

The current study provided compelling evidence, for the first time, that participants can be trained in the laboratory in a relatively short period of time to perform melanoma detection using the perceptual expertise training method. Participants were trained with exemplars of four types of melanoma lesions and four types of benign lesions over 4 days. The training effectiveness was evaluated by the pre- and post-tests using MDT. As compared to the control group participants who received no training, participants in the training group who received perceptual expertise training showed significant improvement in sensitivity to detect melanoma. More importantly, the change was retained, as shown in the second post-test one week later.

Although previous training programs have been shown to improve melanoma detection, the training effects were either measured using a small set of the same images in the pre- and post-tests, or were measured immediately after the training (Rourke et al., 2015). In contrast, the current study showed that the trained melanoma detection skills were transferred to the categorization of novel, untrained stimuli. In real life scenarios, transferability from the training is vital because the person who detects melanomas (e.g., primary care physicians) has to make decisions each time with new exemplars of pigmented lesions on a patient.

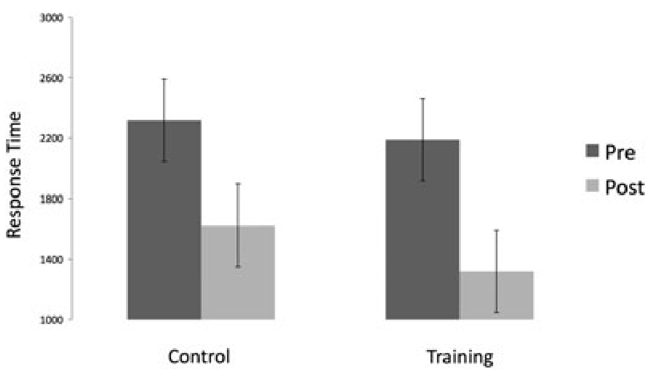

Figure 4. Response time of the pre and post test in the training and control group. The error bars refer to 95% confidence interval. The dark bars refer to the response time in the pre test and the light bars refer to the response time in the post test

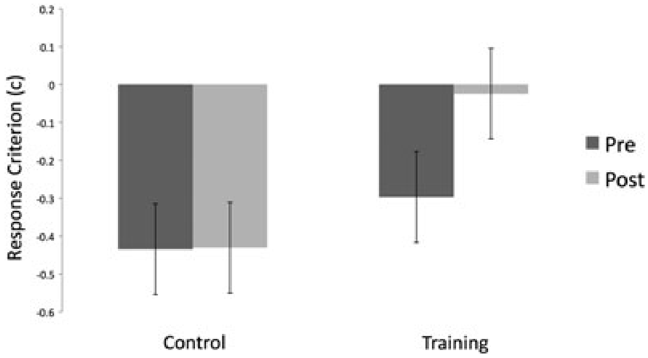

Figure 5. Response criterion (*c*) of the pre and post test in the training and control group. The error bars refer to 95% confidence interval. The dark bars refer to the response criterion in the pre test and the light bars refer to the response criterion in the post test

Critically, the effects of perceptual expertise training for melanoma detection are not short-lived. The retention of perceptual skill is an important hallmark of perceptual expertise training that has been demonstrated for the long-term effect in discrimination of faces (Heptonstall et al., 2013) and cars (Scott et al., 2008). Despite the absence of continued training in the current study, the gains of the perceptual expertise training group were present a week after the training, which is consistent with the results with other perceptual expertise training studies (Scott et al., 2008). However, retention is not always achieved in traditional medical training. As for the training in melanoma detection, none of the customary training programs measured the retention of the training effect (Rourke et al., 2015).

Moreover, the results suggested that after the training, participants became less liberal in categorization (i.e., biased to categorize a lesion as melanoma). It should be noted that, although the change in the response criterion (c) of the training group participants was significant, the ANOVA with test and group in response criterion only yielded marginal interaction effect ($p = 0.11$), suggesting that the size of the change of response criterion in the training group did not significantly differ from that of the control group. Therefore, the robustness of this marginal effect should be validated in further studies with larger sample sizes. Critically, in the ANOVA with false alarm rates, a significant interaction ($p < 0.01$) was found between group and test, suggesting that perceptual training significantly lowered the participants' false alarm rate, as compared to the control group. Importantly, the change of the false alarm rate was not a result of compromising the hit rate, as suggested by the non-significant interaction between test and group in the ANOVA of hit rates ($p = 0.25$). There are important real-world implications for this finding. In real life situations, high false alarm rate in melanoma detection is not optimum. If a benign lesion is misdiagnosed as melanoma, then the patients will have extra financial costs, unnecessary emotional stress and physical and cosmetic changes from the biopsy. Moreover, the unnecessary referral by the primary care physicians to the dermatologist because of the false alarm of melanoma also creates additional burden on medical resources. In the current study, the training group participants decreased the false alarm rate without compromising the hit rate. Therefore, the perceptual expertise training method

Table 2. Results of the first and second post tests for the training group. Numbers in the parentheses refer to the standard deviation

Measure	Post test 1	Post test 2	*p* Value
Sensitivity (*d′*)	1.71 (0.53)	1.57 (0.21)	0.32
Response time	1319 (490)	1258 (377)	0.38
Response criterion (*c*)	−0.02 (0.29)	−0.05 (0.39)	0.71

was useful to address the issues resulting from the false alarms in melanoma detection.

Although the current study has provided important evidence that the perceptual expertise training approach can improve melanoma detection, it is still unknown whether this approach is superior to the conventional training approach currently used in medical training, such as the rule-based ABCDE method. Tschandl, Kittler, Schmid, Zalaudek, and Argenziano (2015) compared the effectiveness of two approaches to teach diagnosis of pigmented skin tumors, namely the 'verbal-based analytic' approach and the 'visual–global heuristic' approach. The verbal-based analytic approach employs the semantic criteria and conducts detailed analysis of the individual features of a lesion, which is similar to the rule-based ABCDE method. On the other hand, the visual–global heuristic approach focuses on the recognition of the overall pattern of a lesion, which is close to the perceptual strategies used by dermatologists (Gauchon, Beaulieu, & Sei, 2005). Tschandl et al. found that after 1 h of training, both groups showed similar size of improvement in diagnostic accuracy, indicating that the heuristic and analytic methods were equally effective in short training programs. However, there are important differences between the visual global heuristic approach used in the Tschandl et al. study and the method used in the current study. While the training program in the current study was computer-based, adaptive, self-paced and spaced across 4 days, the heuristic approach training in the Tschandl et al. study was delivered in the form of a lecture in the classroom and lasted for 1 h. Therefore, because of those differences, the findings from Tschandl et al. should not imply that the perceptual training method used in the current study was not superior to the rule-based method. Because of the lack of standardized and rigorous evaluations of the existing training programs in melanoma or skin cancer detection training (see Goulart et al., 2011, for a review), a direct comparison of the training effectiveness between the current study and existing teaching approaches was not possible. However, future studies should test the efficacy of the perceptual expertise training method to current dermatology teaching practices (e.g., the rule-based ABCDE method).

It should be noted that some limitations exist in the current training program that can be improved in the future. First, the improvement in melanoma detection performance after training was majorly driven by the decreased false alarm rate while maintaining the hit rate. Although decreasing the false alarm rate has important practical implications, as discussed in the previous paragraph, it is important to also increase the hit rate. Future training program should explore how to effectively both improve the hit rate and decrease the false alarm rate. Second, images in the training and testing were selected to represent the classical presentation of each type of lesion, which set limitations on the transferability of the training effect to real life situations. Future training program can be designed to include more exemplars that deviate from the classical presentations of the lesion. Finally, same number of melanoma and benign lesions were used in the training program. However, in the real world practice, the appearances of benign lesions usually outnumber that of the melanoma lesions (Carli et al., 2004; Terushkin et al., 2010). Future training programs should adjust the melanoma to benign ratio to more objectively reflect the base rate of melanoma in real life.

In summary, results from the current study provide preliminary but promising evidence that the perceptual expertise training is an effective method for teaching melanoma detection. Following the expertise protocol, the melanoma detection skill can be acquired in a relatively short period of time (i.e., 4 sessions of 20 to 25 min), transfers to novel lesions and is retained over a period of seven days without retraining. Because the primary care physicians lack sufficient skill in melanoma detection (Chen et al., 2001), the perceptual expertise training program can help them to improve their ability to detect melanoma within a feasible time period. Moreover, the training program can be made available to the public, so that people at high risk of melanoma can receive the training. Ultimately, the program can help decrease the delay between lesion development, detection and diagnosis, and increase treatment to reduce mortality rates of melanoma. Future studies will focus on comparing the effectiveness of the perceptual expertise training method to the other conventional training method (e.g., the rule-based ABCDE method) and whether the perceptual expertise training method is also effective in the education and training in other domains that require visual medical expertise (e.g., radiology, cardiology).

NOTES

Endnotes

1. During data analysis, one stimulus was found to be used both as the exemplar of lentigo maligna and melanocytic nevi category. Therefore, it was removed from both categories in the analysis. As a result, 23 exemplars remained in both the melanoma category and the benign category. All the data analysis was based on the performance of categorizing the remaining 46 stimuli.

REFERENCES

American Cancer Society (2015). *Cancer facts & figures 2015*. Atlanta: American Cancer Society.

Aldridge, R. B., Maxwell, S. S., & Rees, J. L. (2012). Dermatology undergraduate skin cancer training. *BMC Medical Education*, *12*, 27.

Ashby, F., & O'Brien, J. (2005). Category learning and multiple memory systems. *Trends in Cognitive Sciences*, *9*(2), 83–89.

Chen, S. C., Bravata, D. M., Weil, E., & Olkin, I. (2001). A comparison of dermatologists' and primary care physicians' accuracy in diagnosing melanoma: A systematic review. *Archives of Dermatology*, *137*(12), 1627–1634.

Carli, P., De Giorgi, V., Crocetti, E., Mannone, F., Massi, D., Chiarugi, A., & Giannotti, B. (2004). Improvement of malignant/benign ratio in excised melanocytic lesions in the 'dermoscopy era': A retrospective study 1997–2001. *British Journal of Dermatology*, *150*(4), 687–692.

Gachon, J., Beaulieu, P., Sei, J. F., Gouvernet, J., Claudel, J. P., Lemaitre, M., … Grob, J. J. (2005). First prospective study of the recognition process of melanoma in dermatological practice. *Archives of Dermatology*, *141*(4), 434–438.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*, 191–197.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. W. (1998). Training "Greeble" experts: A framework for studying expert object recognition processes. *Vision Research*, *38*, 2401–2428.

Goulart, J. M., Quigley, E. A., Dusza, S., Jewell, S. T., Alexander, G., Asgari, M. M., & INFORMED (INternet curriculum FOR Melanoma Early Detection) Group (2011). Skin cancer education for primary care physicians: A systematic review of published evaluated interventions. *Journal of General Internal Medicine*, *26*(9), 1027–1035.

Hagen, S., Vuong, Q. C., Scott, L. S., Curran, T., & Tanaka, J. W. (2014). The role of color in expert object recognition. *Journal of Vision*, *14*(9), 9.

Hagen, S., Vuong, Q. C., Scott, L. S., Curran, T., & Tanaka, J. W. (2015). The role of spatial frequency in expert object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(3), 413–422.

Heptonstall, B., Tanaka, J. W., & Hoven, M. (2013). The fate of holistic face representations in long-term memory. *Journal of Vision*, *13*(9), 102.

Howlander, N., Noone, A. M., Krapcho, M., Neyman, N., Aminou, R., & Alterkuse, S. F. (2012). *SEER cancer statistics review, 1975–2009 (vintage 2009 populations)*. Bethesda, MD: National Cancer Institute.

Johnson, K., & Mervis, C. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General*, *126*, 248–277.

Meyer, L. J., Piepkorn, M., Goldgar, D. E., Lewis, C. M., Cannon-Albright, L. A., Zone, J. J., & Skolnick, M. H. (1996). Interobserver concordance in discriminating clinical atypia of melanocytic nevi, and correlations with histologic atypia. *Journal of the American Academy of Dermatology*, *34*(4), 618–625.

Oliveria, S., Chau, D., Christos, P., Mushlin, A., & Halpern, A. (2004). Diagnostic accuracy of patients in performing skin self-examination and the impact of photography. *Archives of Dermatology*, *140*(1), 57–62.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal Experimental Psychology*, *77*, 353–363.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning taking memory tests improves long-term retention. *Psychological Science*, *17* (3), 249–255.

Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, *4*(3), 328–350.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605.

Rourke, L., Oberholtzer, S., Chatterley, T., & Brassard, A. (2015). Learning to detect, categorize, and identify skin lesions: A meta-analysis. *JAMA dermatology*, *151*(3), 293–301.

Scott, L., Tanaka, J., Sheinberg, D., & Curran, T. (2008). The role of category learning in the acquisition and retention of perceptual expertise: A behavioral and neurophysiological study. *Brain Research*, *1210*, 204–15.

Terushkin, V., Warycha, M., Levy, M., Kopf, A. W., Cohen, D. E., & Polsky, D. (2010). Analysis of the benign to malignant ratio of lesions biopsied by a general dermatologist before and after the adoption of dermoscopy. *Archives of Dermatology*, *146*(3), 343–344.

Tschandl, P., Kittler, H., Schmid, K., Zalaudek, I., & Argenziano, G. (2015). Teaching dermatoscopy of pigmented skin tumours to novices: Comparison of analytic vs. heuristic approach. *Journal of the European Academy of Dermatology and Venereology*, *29*(6), 1198–1204.

Tanaka, J. W., Curran, T., & Sheinberg, D. L. (2005). The training and transfer of real-world perceptual expertise. *Psychological Science*, *16* (2), 145–51.

Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(1), 122–31.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, *23*, 457–482.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338–353.