# Proofreaders show a generalized ability to allocate attention to detect changes

**Article** *in* PSYCHOLOGIA · May 2010

Impact Factor: 0.09 · DOI: 10.1167/8.6.462

*Psychologia*, 2008, **51**, 126–141

# PROOFREADERS SHOW A GENERALIZED ABILITY TO ALLOCATE SPATIAL ATTENTION TO DETECT CHANGES

**Michiko ASANO[1)2)], Shoko KANAYA[1)], and Kazuhiko YOKOSAWA[1)]**

*[1)]The University of Tokyo, Japan, [2)]JSPS Research Fellow, Japan*

In proofreading, attention must be allocated to the entire text in order to detect anomalous words within the sentence context. Because professional proofreaders have superior control of spatial attentional allocation, it is hypothesized that their skill may generalize to certain other tasks. To test this hypothesis, professional proofreaders were compared to novices using change blindness and visual search tasks. Results of the change blindness task showed that proofreaders detected more targets (non-predesignated scene changes) than novices by allocating attention to the entire scene. By contrast, novices' attention tended to be attracted to the upper part of visual scenes. Results of the visual search task, which involved counting predesignated targets, showed no performance differences between proofreaders and novices. This implies that proofreaders have a highly developed ability for spatial attentional allocation that is generally applicable to attentional demanding search situations where both detection targets and distractors are not predesignated.

**Key words:** attentional allocation, expertise, proofreading

Human performance can sometimes be extended through experience and deliberate practice. People who have attained high levels of performance in a domain are called experts. Experimental approaches to expertise have examined skill acquisition across a wide range of different domains including medicine, music, sports and games (Ericsson & Smith, 1991; Ericsson & Williams, 2007). These studies provide interesting insights into the structure and acquisition of human skills and abilities.

A common observation is that expertise is highly domain specific, and therefore unlikely to transfer to other domains (Ericsson & Smith, 1991; Ericsson & Kintsh, 1995). For example, although masters in the game of chess showed a recall advantage for chess positions when presented with well structured (game-like) chess arrays, this advantage vanished when these arrays consisted of randomly placed pieces (Chase & Simon, 1973). Similarly expert scrabble players showed superior performance on certain kinds of verbal and visuospatial tasks which tap into abilities that are specific to scrabble. These abilities involve rapid memory retrieval of appropriate words as well use of knowledge about the spatial layout of words and letters on particular squares on the board which determine a play's point value (Halpern & Wai, 2007).

This study was partially supported by Grant-in-Aid for Scientific Research awarded to K. Y. and by Grant-in-Aid for Japan Society for the Promotion of Science (JSPS) Fellows awarded to M. A., both from JSPS.

Correspondence concerning this article should be addressed to: Michiko Asano, Graduate School of Humanities and Sociology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan (e-mail: asano@l.u-tokyo.ac.jp).

Interestingly, few studies have considered the involvement of visual attention in expertise. This is relevant because visual attention is typically considered to be a domain free system. Efficient execution of many, especially higher-order, cognitive tasks requires effective allocation of visual attention. Consequently, it may be predicted that, in addition to domain specific abilities, experts in cognitive tasks that are especially attention-demanding should have highly developed domain free abilities, which involve the allocation of visual attention. Consistent with this prediction, a study on visual attention allocation, using a variant of the Navon task, found that soccer players were quicker in switching attention from local to global information than non-athletes (Pesce, Tessitore, Casella, Pirritano, & Capranica, 2007). Although the experimental task in this case differed significantly from a real soccer situation, the two situations shared a dominant factor: both required visuo-spatial processing. Accordingly, further investigations of role of domain-free attentional allocation in expertise may require identifying tasks that share dominant factors, but which superficially appear to bear little relationship to the expert's particular domain.

*Professional Proofreaders*

In the present study, we focused on expert abilities of professional proofreaders. Proofreading refers to the process of reading written documents for errors; high precision in this process is necessary for insuring coherently constructed communication. Thus, good proofreading requires that one's attentional focus be systematically allocated throughout the entire text in order to detect all anomalous words that do not fit into relevant sentence contexts. Because professional proofreaders proofread articles in their daily work, it is reasonable to assume that they possess excellent control of attentional allocation. Here, we further assume that this attentional skill may generalize to different tasks, which share common underlying factors with proofreading.

The task of proofreading is well suited to the purpose of investigating the contribution of attentional allocation in expertise. Asano and Yokosawa (2007) have examined characteristics of professional proofreaders in various tasks; specifically, these include anomalous word detection, the Stroop task, and change blindness task. In the anomalous word detection task, professional proofreaders were better than novices in detecting anomalous words and this difference could not be attributed to differences in lexical knowledge because estimated vocabularies of the two groups were approximately equivalent. In the Stroop color-word task, interference with color naming is typically caused by incongruency between semantic and color information of stimulus in color naming. However, surprisingly when proofreaders performed this task, they showed no Stroop interference, whereas novices, with no proofreading skills, did show the predicted interference. In the change blindness task, a characteristic feature is "the striking failure to see large changes that normally would be noticed easily" (Simons & Rensink, 2005). People are unable to detect relatively large changes between two images during saccades, blinks, or flicker. Rensink (2002) has claimed that focused attention is required to detect an image change. This suggests that the change blindness task can be used as a measure of one's skill in allocating attention to different spatial locations (Simons & Rensink, 2005).

It is relevant, therefore, that Asano & Yokosawa (2007) found that proofreaders were better than novices at detecting scene changes in this task. In this case, experimenters presented a short video [the same stimulus used in Experiment 1 of Levin & Simons (1997)] in which some changes occurred across video cuts. In their review of these and related findings, Asano and Yokosawa attributed the lack of interference in the Stroop test as well as the superior performance in detection tasks (anomalous words, change detection) to proofreaders' superior ability to focus attention and, to their related ability to inhibit interference from potentially destructive information.

*Working Memory and/or Attentional Allocation*

We have suggested that one explanation for superior performance of proofreaders in certain tasks involves their attentional allocation abilities. However, other explanations are possible. Thus, it has been argued that working memory plays a role in the control of selective attention (de Fockert, Rees, Frith, & Lavie, 2001; Engle, 2002). Working memory capacity has also been implicated in the Stroop effect. Thus, Kane and & Engle (2003) reported that people with high working memory spans exhibit a reduced Stroop effect. In other words, professional proofreaders may have acquired higher working memory spans than novices and this factor accounts for their superior performance in certain tasks. Arguably, proofreading poses significant demands on working memory, and this may be a critical factor in explaining the superiority of proofreaders in the change blindness task as well as the Stroop task (Asano & Yokosawa, 2007). In short, a working memory explanation may account for differences in performance of proofreaders and novices in the change blindness task.

Another potential explanation for proofreaders' superiority in certain tasks involves search strategy. It remains unclear whether the superior performance exhibited by professional proofreaders in change detection tasks was due to an enhanced ability to engage in a simple exhaustive search strategy or the result of more efficient allocation of spatial attention than non-experts. Although Asano and Yokosawa (2007) attributed the superior performance of proofreaders in the change blindness task to their ability to allocate attention exhaustively, this can be separated into two aspects: one is the ability to literally search "exhaustively", that is, to search all items without exception, and the other is the ability to conduct a difficult, attention demanding, search using excellent spatial control of attentional allocation. Both of these aspects might be involved in proofreading. That is, proofreading certainly entails an exhaustive search for contextually anomalous word targets since the number of such targets is typically unknown; but, in addition, proofreading is undeniably an attentionally demanding task, since targets and distractors are not predesignated, namely they are not defined in advance by a distinctive visual feature (e.g., color). Therefore, on the face of it, in the change blindness task of Asano and Yokosawa (2007), where multiple (number unknown to participants), non-predesignated, changes occurred, it is difficult to ascertain whether the telling expertise of proofreaders in this experiment reflected proofreaders' ability to conduct a simple exhaustive search or whether (as the authors surmised) it reflected proofreaders' enhanced ability to allocate attention in space. Nevertheless, these two aspects of the proofreading task can be experimentally dissociated by employing two different tasks each of which draws, respectively, on one of these two aspects. Thus, a change blindness task in which one non-predesignated target (e.g., the flicker paradigm used in Yokosawa & Mitsumatsu, 2003) can isolate attention demanding aspects of performance, whereas a visual search task, in which several predesignated targets appear, where the number of targets is not specified to participants, can isolate defining features of an exhaustive search.

One goal of the present research was to test the hypothesis that expertise reflects a highly developed ability for attentional control. Proofreaders should have excellent control of spatial attention because successful proofreading requires a uniform allocation of attention throughout the entire text in order to detect all anomalous words. Furthermore, to the extent that this attentional skill is a general, domain free skill, then professional proofreaders should perform better than novices not only in anomalous word detection tasks, but also in other tasks that share major features of the anomalous word detection tasks. Accordingly, in Study 1 we first examined performance of professionals and novices in an anomalous word detection task to verify their skill differences. In Study 2, we addressed the generality of the proficiency differences observed in Study 1 using two new tasks, the change blindness and visual search tasks.

## STUDY 1

The goal of Study 1 was to establish differences in anomalous word detection performance between professional proofreaders and novices. We sought to confirm that the professional proofreaders in our experiment were really professionals. In addition, vocabulary estimation and working memory tests were conducted to ascertain whether any proficiency differences in anomalous word detection ability that emerged between these two participant groups could be attributed to vocabulary size and/ or working memory capacity.

## METHOD

### *Participants*

A total of 34 participants (professional proofreaders and novices) served in Study 1. All were native Japanese speakers. The two groups were matched on age, sex, and educational background (most graduated from a university or a college).

*Professional proofreaders*. Seventeen of the 34 participants were professional proofreaders (9 men and 8 women). The mean age for this group was 44.2 years (between 26 and 64 years old, $SD$ = 10.1). They were proofreaders and editors who work for publishers and newspaper companies or as freelance, and who proofread documents as their daily work. They had, on average, 17.3 years of experience of professional proofreading (between 0.5 and 44 years, $SD$ = 11.5).

*Novices*. The group of novices consisted of 17 people (9 men and 8 women) who had no experience of professional proofreading. The mean age for this group was 45.1 years (between 24 and 62 years old, $SD$ = 9.2).

*Vocabulary Estimation Test*

A vocabulary estimation test, provided by NTT Communication Science Laboratories, was used to measure participants' vocabulary size. In this test, participants marked words for which they could identify the meaning and could use in a list of 50 Japanese words. Participants' vocabulary size was estimated based on this data.

*Verbal Working Memory Test*

The Japanese reading span test (the Japanese RST; Osaka & Osaka, 1994) was used to measure participants' verbal working memory capacity. In this test, participants read sets of Japanese sentences aloud while remembering the underlined word in each sentence. At the end of the each set, participants were asked to recall the underlined word in each sentence orally. The number of sentences in a set was increased from two to five; span value was evaluated as the highest level (set size) at which a participant could achieve a correct recall rate that satisfies a criterion. Full details are available in Osaka (1998).

*Anomalous Word Detection Task*

*Apparatus.* A stimulus text was displayed on the 12.1 inch LCD monitor of a tablet laptop PC (HP Compaq tc4400 Tablet PC). The monitor was laid flat in horizontal orientation. The tablet PC allowed participants to write on the monitor using a digital pen, so that participants could circle anomalies in the stimulus text displayed on the PC monitor as if they were proofreading (with a pen) a text printed on papers flush with a desktop. The stimulus text consisted of three pages where pages were flipped by a key press. Total performance time was recorded from onset of a stimulus text to the end of the task using a digital millisecond timer.

*Stimuli.* A Japanese text of 1,250 characters was used to present stimulus words. This text was displayed in three pages, written vertically in black on a white display, and double-spaced. All the characters were displayed in MS Mincho (Japanese fixed-width font) and the font size was 16 point (approximately 0.5 degrees of arc).

The stimulus text included 24 two-Kanji-character anomalies (12 non-words and 12 contextually anomalous words). Kanji is a logographic script used in Japanese, and two-Kanji-character compound words and two-Kanji-word errors are common in Japanese texts. Four conditions involved modifications of selected base words; these modifications were designed to vary the degree of effective similarity between the resulting anomalous word and its original (base) word, which is its contextually consistent counterpart. Three conditions had anomalies that were: phonologically identical, orthographically similar, and semantically similar; a fourth, and control condition had anomalies that were phonologically, orthographically, and semantically dissimilar from the base word. All base words were two-Kanji-character compound nouns. Anomalies of base words were created by replacing one of the two characters of each base word with another character that was either phonologically identical, orthographically similar, semantically similar, or generally dissimilar (control) with the original character, so that it became either a non-word or a contextually anomalous word. All base words had word familiarity ratings over 5.0 on a scale that ranged from 1 to 7. Word familiarity was based on the Japanese vocabulary database "Nihongo-no Goi Tokusei" (Amano & Kondo, 1999). Congruencies of all anomalies within the stimulus text were assessed by a questionnaire completed by 18 judges (who were not participants); they were asked to select the contextually correct words from word pairs comprising anomalies and base words. All of the anomalies were rejected as contextually incorrect with high confidence (above 6 on a seven-point scale of confidence).

*Procedure.* The main task was to detect target anomalies in the stimulus text. Participants were asked to circle each anomaly with a digital pen; they were told not to correct the anomaly. Participants were free to re-read the text. Before the experiment started, all participants were instructed to complete the task within 5 minutes, however, the actual time limit allowed for about half of the participants (9 participants in each group) was 8 minutes. After 5 minutes had passed, these longer-time-limit participants were told that they could have another 3 minutes. This generated two time-limit conditions: 5 minutes and 8 minutes. The rationale for this involved a concern that a strict 5 minute limit, while adequate for normal reading of this text, may be unduly restrictive for the detection task. Eight minutes was considered sufficient time. We decided to test these two time-limit conditions by randomly assigning participants to one of each of the two conditions while retaining constraints on matching age and sex of the professionals and novices within each condition.

Table 1. Mean Percentage of Anomalous Non-words/Words Detected in Anomalous Word Detection Task in Study 1 (%)

| | Non-Word | | | | Word | | | |
|---|---|---|---|---|---|---|---|---|
| | P | O | S | C | P | O | S | C |
| 5 Minutes[a] | | | | | | | | |
| Professionals | 86.3 | 62.7 | 74.5 | 84.3 | 84.3 | 33.3 | 41.2 | 62.7 |
| | (4.8) | (6.7) | (7.6) | (4.9) | (4.9) | (7.3) | (4.4) | (6.7) |
| Novices | 78.4 | 39.2 | 64.7 | 74.5 | 58.8 | 19.6 | 29.4 | 56.9 |
| | (6.2) | (5.7) | (9.0) | (7.1) | (6.5) | (6.2) | (6.1) | (8.2) |
| 8 Minutes[b] | | | | | | | | |
| Professionals | 96.3 | 66.7 | 81.5 | 96.3 | 96.3 | 40.7 | 48.1 | 66.7 |
| | (3.5) | (7.4) | (9.2) | (3.5) | (3.5) | (10.2) | (5.5) | (9.1) |
| Novices | 85.2 | 48.1 | 77.8 | 77.8 | 48.1 | 11.1 | 22.2 | 48.1 |
| | (5.5) | (5.5) | (7.4) | (7.4) | (5.5) | (5.2) | (7.4) | (5.5) |

*Notes.* P = phonogically identical. O = orthographically similar. S = semantically similar. C = control (phonologically, orthographically, and semantically dissimilar). Values in parenthesis represent standard errors. [a] N = 17 in each of the two groups. [b] N = 9 in each of the two groups.

## Results

### *Vocabulary Estimation Test*

The mean estimated vocabulary size was 50,671 words ($SE$ = 1182) for professionals and 48,729 words ($SE$ = 2,546) for novices. A $t$ test revealed no significant differences between the estimated vocabulary sizes for these groups, $t(32) = .67$, n.s.

### *Verbal Working memory Test*

The mean reading span was 2.35 ($SE$ = .09) for professionals and 2.18 ($SE$ = .07) for novices, and the differences between them were not significant, $t(32) = 1.48$, n.s.

### *Anomalous Word Detection Task*

Data from both of the time-limit conditions for the first five minutes were pooled to permit analysis of performance of all 34 participants. In addition, detection performance of the 18 participants who met the 8 minute limit was analyzed separately. Mean detection rates of anomalies in 5 minute and 8 minute conditions appear in Table 1. The data in 8 minute condition represent the cumulative detection rates from the beginning of the task. The detection rates in each of the two time limit condition were analyzed using a three factor mixed factorial ANOVAs with participant group (professionals, novices) serving as the single between-participants factor. Lexicality of the anomalies (non-word, word) and similarity type of the anomalies (phonology, orthography, semantic, control) were the two repeated measures variables.

*Detection performance in 5 minutes.* Main effects of participant group and lexicality were significant with $F(1, 32) = 4.55, p < .05$ and $F(1, 32) = 60.34, p < .01$), respectively. These indicate that professionals detected more anomalies than did novices, and non-word anomalies were more likely to be detected than word anomalies. In addition, a main effect of anomality similarity was present, $F(3, 96) = 37.94, p < .01$. Following up on the latter outcome, Tukey's HSD tests showed that detection rates for the phonologically identical and the control anomalies were higher than those for both the semantically and the orthographically similar anomalies with the orthographically similar anomalies least likely to be detected (all three $ps < .01$). An interaction between lexicality and similarity type of the anomalies was also significant, $F(3, 96) = 3.84, p < .05$. Tukey's HSD tests showed that non-word anomalies were more likely to be detected than word anomalies in all the similarity types with the exception of the phonologically identical condition (all $ps < .01$). Tukey's HSD tests also revealed that when the anomalies were non-words, orthographically similar anomalies were more likely to be overlooked than anomalies of the other three similarity types. On the other hand, when the anomalies were words, orthographically and semantically similar anomalies were less likely to be detected than phonologically identical and control anomalies (all $ps < .01$).

*Detection performance in 8 minutes.* Main effects of participant group and lexicality, were significant with $F(1, 16) = 16.39$ , $p < .01$, and $F(1, 16) = 74.51$, $p < .01$, respectively. These indicate that professionals detected more anomalies than did novices, and that non-word anomalies were more likely to be detected than word anomalies. In addition, anomaly similarity type also proved significant, $F(3, 48) = 33.73$, $p < .01$. Consequent Tukey's HSD tests showed that (again) the detection rates for the phonologically identical and the control anomalies were higher than those for the semantically and the orthographically similar anomalies with the latter (again) least likely to be detected (all three $ps < .01$). The interaction between participant group and lexicality of the anomalies was also significant, $F(1, 16) = 5.99, p < .05$. Tukey's HSD tests revealed that the professionals detected more anomalies than did novices only when the anomalies were words ($p < .01$). The interaction between lexicality and similarity type of the anomalies was marginally significant, $F(3, 48) = 2.77, p = .05$.

## Discussion

Professional proofreaders and novices did not differ in estimated vocabulary size or in working memory capacity measured by the Japanese RST. This means that the greater ability of proofreaders to detect anomalous words, confirmed by the anomalous word detection task, is not due either to greater lexical knowledge or to superior working memory capacity.

Professionals detected more anomalies than did novices in the anomalous word detection task, both in 5 and 8 minutes time limit conditions. In addition, the results from the 8 minutes condition (relative to the 5 minute condition) suggest that professionals displayed their ability when the anomalies were words provided they were insured sufficient to complete the task.

It is reasonable to expect that the more salient an anomalous word/non-word "target" is within the text, the more likely it would be detected. In turn, this leads to the expectation that anomalies that are similar to their contextually correct counterparts should be relatively hard to detect. Overall the manipulation of anomaly similarity turned out to be effective. Regardless of differences in participant expertise and time limit conditions, good detection rates emerged for two anomaly similarity conditions, the phonologically identical and the control condition, whereas relatively poor detection rates were found for two other anomaly conditions, the orthographically and semantically similar conditions. Poor detection performance in the latter two conditions is consistent with the results of Morita & Tamaoka (2001), Shimomura & Yokosawa (1995), and Morita & Tamaoka (2002). In these studies, systematic effects of both orthographical and semantic information in reading Japanese sentences were revealed using proofreading tasks. However, one finding does not conform to the expectations regarding anomaly similarity mentioned above; the finding that phonologically-identical anomalies, which is by some criteria highly similar to original base word, nonetheless elicits relatively good performance. The good performance for phonologically identical anomalies contradicts the prediction noted above and some previous proofreading studies in which short sentences were used as stimuli (e.g., Jared, Levy, & Rayner, 1999), however, is consistent with the results reported in the anomalous word detection task in Asano & Yokosawa (2007). One reason for the detection advantage with phonologically-identical anomalies is that such anomalies are quite common in Japanese languages, in which many homophones exist. Therefore, this may promote a strong attentional set for such anomalies in all participants.

We also found that word anomalies were detected with more difficulty than non-word anomalies. Presumably this is because the detection of word anomalies requires attentional allocation. A non-word can be judged as an anomaly by processing only the target item, whereas correct detection of a word anomaly will depend upon processing of the sentence context surrounding it. This is because a word anomaly is *defined* as a word that does not fit the sentence context. Sentence context processing may degrade the recognition of words that do not fit the sentence context (Potter, Stiefbold, & Moryadas, 1998), therefore attention allocation is necessary for the detection of word anomalies. The ability to allocate attention to specific words within a sentence may be a component of the skill that leads to superior performance levels for proofreaders in this task.

## STUDY 2

The main goal of Study 2 was to assess the degree to which a refined attentional allocation skill, found in professional proofreaders, generalizes to other, superficially different, tasks which share important properties of the proofreading task. To accomplish this we examined the performance of professional proofreaders and novices in two new tasks: a change blindness task and a visual search task. Our aim is to examine whether the attentional demands in either of these two tasks might selectively favor the acquired attentional skills of professional proofreaders.

Both the change blindness and the visual search (with multiple targets) tasks require a serial search for target(s) amidst distractors, as noted in the introduction. However, the change blindness task shares with the anomalous word detection task significant demands on the allocation of spatial attention. During the search in a change blindness task, participants must detect a single non-predesignated scene change. By contrast, the visual search task merely requires that a viewer detect occurrences of a number of targets that possess a predesignated feature. In the latter, to insure an exhaustive rather than a self-terminating visual search, multiple (predesignated) targets were embedded in search arrays and participants had to report the number of targets they detected. Although this search task may require some attention, it can be argued that the change blindness task is more demanding in that it requires an excellent control of attentional allocation in much the same fashion as do anomalous word detection tasks.

The particular change blindness paradigm used in Study 2 was the flicker paradigm (Rensink, O'Regan, & Clark, 1997), which involves static pictures of common scenes (e.g., parks, shops, and streets). In this paradigm two (slightly different) static scene images repeatedly alternate, separated by brief blanks, until a participant reports detecting the single scene change. This technique allows control over spatial loci of changes, and this is an important consideration in any investigation of attentional allocation. Unlike a movie paradigm, in which objects move continuously (e.g., see Asano & Yokosawa, 2007), in the flicker paradigm, the use of static images allows us to more reliably assess a viewer's identification of the spatial location of a change within a scene.

Typically, picture scenes, such as those used in the change blindness task, contain more informative objects, such as people and buildings, in the upper portions of the scene as compared to the lower area, where the ground can be found. Therefore, the average viewer (e.g., a novice) tends to allocate attention to upper scene locations. However, one hypothesis regarding spatial attentional control in this task holds that if professional proofreaders have good control of attentional allocation and if this control generalizes to related tasks, then unlike novices, professional proofreaders should be able to efficiently allocate attention to the entire scene.

By contrast, in the visual search task of Study 2, the attentional allocation hypothesis does not predict differences in performance between proofreaders and novices. Although this task requires a serial search, as does the change blindness task, the search strategy required does not depend upon precision allocation of spatial attention. In this task, arrays of digits and letters are presented, with features of the targets (digits) specified in advance. If skilled proofreading rests on efficient attentional allocation and not upon enhanced proficiency in conducting an exhaustive serial search, then, unlike performance in a change blindness task, performance in this visual search task should not differ as function of levels of expertise (i.e. professionals versus novices).

## METHOD

*Participants*

The same participants as in Study 1 participated in Study 2. They participated in Study 2 on the same day as they participated in Study 1. Adequate intermissions were given between these two studies, and between change blindness and visual search tasks. The task order was fixed; change blindness task was followed by visual search task.

*Apparatus*

The same laptop PC used in Study 1 (for anomalous word detections) was used in Study 2. The PC monitor was placed vertically and could be tilted so that participants could view it comfortably. The viewing distance was approximately 50 cm. Responses in the change blindness task were made and recorded through both PC keyboard and a computer mouse. In the visual search task, only PC keyboard responses were made and recorded.

*Change Blindness Task*

*Stimuli.* Twelve sets of color photographs of common scenes served as stimuli. These were a subset of materials used by Yokosawa & Mitsumatsu (2003). Each of the 12 sets comprised two images, an original and a modified image, where the modified image involved either a single color change or the disappearance of a single object in the original image scene. Locations of these changes within the spatial layout of a scene were balanced and distributed across the whole image. Over the course of a session, equal numbers of changes appeared in the upper and lower portions of images. In both cases, the extent of a change was limited so as not to exceed one spatial section; a section was determined by dividing the total area of an image into 24 equal-area sections ($4 \times 6$ matrix). On the monitor, the size of each stimulus image was 24.5 cm (approximately 28 degrees of arc) wide × 18.5 cm (approximately 21 degrees of arc) high.

*Procedure.* Each trial began with a fixation cross (+) presented for 300 ms in the center of the monitor screen. Then the original image (250 ms) repeatedly alternated with a modified image (250 ms), with a brief blank (250 ms) inserted between the images. The blank stimulus was black. The participants' task was to press the space key as soon as they saw the change. Participants had to respond within a time limit of 1 minute. When the response was made or 1 minute had passed from the beginning of the trial, the flickering stopped and the original image was displayed with black lines drawn along the boundary regions of the 24 sections.

A total of 12 trials comprised a test session; half the trials contained changes in the upper sections of a scene, the remainder contained changes in lower sections. Five practice trials were given before the test session began. On each trial, following a key press (detection response), the participant used the PC mouse to point to a section that included the changed region. The following trial started 80 ms after the participant had answered the changed location. No feedback was given on the correctness of the answer.

*Visual Search Task*

*Stimuli.* A stimulus array consisted of a total of 16 items (letters, digits), with the number of the digits in a given array ranging from 1 to 4. The letters (distractors) were chosen randomly, without duplication, from all uppercase alphabetical letters except for I, J, and O. The digits (targets) were chosen randomly from 1 to 9 without duplication. The sixteen search items were randomly plotted in the cells of a $6 \times 6$ virtual matrix to form a search array. The overall size of the matrix was 20 cm (approximately 23 degrees of arc) wide × 20 cm high on the monitor. The size of an item was 1 cm (approximately 1.1 degrees of arc) × 1 cm. Items were displayed in white in Tahoma font on a black background.

*Procedure.* Stimuli were displayed at the center of the PC monitor. Each trial began with the presentation of a fixation cross (+), which appeared for 1000 ms. Immediately following the offset of the fixation cross, a search array was presented. Participants were required to answer the number of digits (from 1 to 4) in the array as quickly as possible, by pressing one of four keys each of which were assigned to 1, 2, 3, and 4, respectively. No feedback was given. Response time was measured from the onset of the search array to the key press. The search array disappeared after the response was made. The next trial began 1000 ms later.

A test session consisted of 72 trials, in which each of the four target number conditions were presented randomly and with equal probability. There were short intermissions after every one fourth of the test session was finished. Ten practice trials were given before the test session began.

## RESULTS

*Change Blindness Task*

A trial was counted as correct when a response was made within a one minute time limit and it correctly identified the location (spatial section) of a scene change. One participant in the professional group showed high false alarm rate (i.e., incorrect identification responses made within the time limit) that exceeded 80%, therefore, data of this participant were eliminated from the following analysis. Mean false alarm rates for professionals and novices were 1.0% ($SE = 0.7$) and 2.0% ($SE = 1.1$), respectively; the difference between these two groups was not significant, $t(31) = .68$, n.s.

*Correct detection rate.* The mean rates of correct change detections for upper- and lower-half trials (which occurred equally often) are shown in Figure 1 (left) for both professionals and novices. We performed a two-factor mixed factorial ANOVA on correct detection rates in which groups (professional, novice) functioned as the between-participants variable and change location (upper, lower) was the repeated measures variable. Significant main effects were found for both variables; for groups, the $F(1, 31) = 7.73$, $p < .01$, and for change location, the $F(1, 31) = 16.02$, $p < .01$. Overall, professionals were more accurate than novices; in addition accuracy was greater for changes located in the upper (versus lower) sections of a scene (Figure 1). However, of most interest is the finding of a significant interaction between these two factors, $F(1, 31) = 5.13$, $p < .05$. Tukey's HSD tests showed that novices detected significantly fewer changes when these occurred in lower half of stimulus images ($p < .01$), whereas there was no reliable effect of change location for professionals. Consistent with this finding, professionals detected significantly more changes in the lower half of the stimulus image than did novices ($p < .05$); this difference also appears in Figure 1 (left).

*Reaction time.* Only the data from correct trials were used in the analyses of reaction time. One novice failed to detect any changes in lower-half trials, therefore, the reaction time in the lower-half trials for this participant was treated as a missing data point. Different patterns of results were again obtained for professionals and novices (Figure 1,

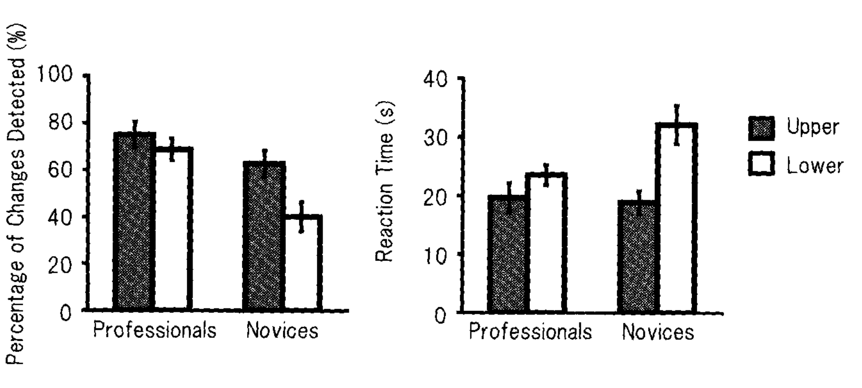


**Fig. 1.** Mean percentage of changes detected (left) and mean reaction time (right) in change blindness task in Study 2. Error bars represent standard errors.

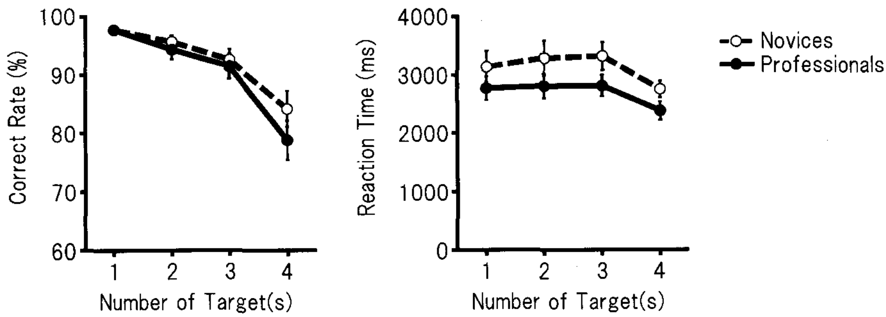


**Fig. 2.** Mean correct rate (left) and mean reaction time (right) as a function of number of target(s) in visual search task in Study 2. Error bars represent standard errors.

right). A two-factor mixed factorial ANOVA with groups (professional, novice) and change location (upper, lower) revealed a significant main effect of location of change, $F(1, 30) = 13.78$, $p < .01$, and a significant interaction, $F(1, 30) = 4.43$, $p < .05$, between participant group and location. Following up on this interaction, Tukey's HSD tests showed that novices took significantly longer to detect changes in the lower half of the stimulus images than those in the upper half ($p < .01$), while no effect of change location was found for professionals.

*Visual Search Task*

*Correct rate.* The mean correct rates for both participant groups were generally high as shown in Figure 2 (left). We performed a two-factor mixed factorial ANOVA on correct response rates with groups (professional, novice) as a between-participants factor and target number (1–4) as a repeated measures factor. Only the main effect of target number was significant, $F(3, 96) = 29.06$, $p < .01$. Tukey's HSD tests revealed that search accuracy was significantly lower when an array contained 4 targets than when less than 4 targets were present (all $p$s $< .01$), and significantly poorer when 3 targets were present than that for arrays containing only one target ($p < .05$).

*Reaction Time.* Only the data from correct trials were used in the analyses of reaction time. Mean reaction times for the visual search task are shown in Figure 2 (right). A two-factor mixed factorial ANOVA, again with factors of groups and number of targets was performed on reaction time data. This revealed a single significant main effect, due to target number, $F(3, 96) = 13.08$, $p < .01$. Tukey's HSD tests revealed that the reaction time was significantly shorter when a search array contained 4 targets than in other conditions (all $p$s $< .01$).

## DISCUSSION

The results of Study 2 indicate that proofreaders detected more scene changes than did novices. Moreover, the location of a change in an image had no impact on proofreaders, whereas it did have a significant impact on novices. Novices missed more changes in the lower (versus upper) half of the scene pictures, and their performance in detecting changes in the lower half of scenes was distinctly poorer than that of proofreaders. These findings suggest that professional proofreaders allocated spatial attention to the entire scene, in contrast to novices' whose attention tended to be attracted to the upper part of visual scenes.

By contrast, in the visual search task, no differences in search performance of the two groups were found. Although the overall reaction time for professionals was slightly faster than novices, this difference did not attain statistical significance ($p > .15$). Further research is needed to clarify apparent differences in the reaction time between professionals and novices. The main effect of target number was significant both for the reaction times and the correct rates. When there were 4 targets, reaction time decreased compared to reaction times on trials when fewer targets were present. This suggests that all participants conducted an exhaustive search without using self-terminating strategy, except for when there were 4 targets. It also implies that the exhaustive search strategy was equally effective for both professional proofreaders and novices. In light of this, it seems unlikely that the higher rate of change detection for professionals observed in the change blindness task can be explained in terms of differences in efficiency of an exhaustive search. With regard to correct response rates, increasing the number of targets also induced a tendency toward lower accuracy. This result may be interpreted as a reflection of task difficulty. Digit targets were not very salient among the letter distractors, and it became harder to detect all of the targets as the number of targets increased.

In sum, the results of Study 2 imply that professional proofreaders have a highly developed and general ability for spatial attentional allocation in tasks where both of detection targets and distractors are not predesignated. Moreover, proofreaders' superior performance cannot be explained in terms of greater proficiency in conducting an exhaustive search strategy.

## General Discussion

In the present research, professional proofreaders were compared to novices in various tasks such as change blindness and visual search tasks. Our aim was to test the hypothesis that expert proofreaders have developed excellent control of spatial attentional allocations, and that this attentional skill can be generalized to other tasks which share certain common factors with the proofreading task.

In Study 1, as expected, professional proofreaders performed much better than novices in an anomalous word detection task, a task which simulates important aspects of proofreading. It is noteworthy that in this study, the two participant groups did not differ in either estimated vocabulary scores or in reading span, which is a measure of verbal working memory capacity. These results permit us to rule out two explanations of superior proofreading performance, namely that it is due to proofreaders' greater lexical knowledge or that it is due to their greater working memory capacity.

In Study 2, one aim was to examine two additional explanations for the superior performance of proofreaders in Study 1. One, an attention allocation hypothesis, holds that the superior performance of proofreaders (relative to novices) depends upon a highly developed ability to allocate spatial attention in attention demanding search tasks. The other, an exhaustive search hypothesis, proposes that proofreaders are simply better (than novices) in conducting the exhaustive searches, such as required by an anomalous word task (as in Study 1).

To evaluate these hypotheses, Study 2 employed two different search tasks, a change blindness task and a visual search task. The change blindness task placed greater demands on attentional allocation than the visual search tasks because it required that participants detect scene changes (targets) where features of both targets and of distractors were not predesignated. To succeed in this task, participants had to efficiently allocate their spatial attention to all parts of a scene. The visual search task, by contrast, placed a lower burden on attentional allocation. To succeed in this task, participants merely had to engage in an exhaustive search to count all targets that exhibited a predesignated feature.

The results of Study 2 showed that proofreaders detected more changes than did novices in the change blindness task. In addition, in contrast to novices, proofreaders allocated attention to all parts of a visual scene in this task. Novices tended to focus attention on upper parts of a complex scene thereby missing changes in the lower half of the scene pictures. This was not the case with proofreaders, who showed no difference in change detection accuracy for upper and lower portions of scenes. By contrast, in the visual search task, which required only exhaustive search, we found no difference in performance levels of experts and novices. Taken together, these results imply that professional proofreaders have a highly developed ability for spatial attentional allocation and that this skill can be effectively applied to detect target changes in situations where neither the targets nor distractors are predesignated.

There remains yet another possible explanation for our results in Study 2 that should be mentioned. Perhaps proofreaders are simply better at perceptually segregating figure from background in complex scenes. Thus, changes in the change blindness task might be more apparent to proofreaders in this task whereas in the visual search task there is little cost to segregating figure from background since targets could be easily distinguished from distracters. Although, in principle this perceptual explanation is possible, it seems highly unlikely for the following reason: Neither visual figure-ground scenes nor target-distractor patterns are found in proofreading materials, meaning that there are no grounds for assuming that proofreaders might have acquired such a skill. Proofreaders are trained to respond to printed texts and to anomalous word targets, defined as words that do not fit into the sentence contexts. Superficially, these differ greatly from visual scenes and they do not have visual features that distinguish figure from background information. Therefore, it is unlikely that expert proofreaders have developed ability for figure-background segregation through their experience in proofreading.

The results of Study 2 also speak to the issue of generality of attentional skills. Attentional allocation is considered to reflect a domain free system. Yet expertise is often considered to be highly domain specific, and hence unlikely to transfer to other domains (Ericsson & Smith, 1991; Ericsson & Kintsh, 1995). However, our findings appear to qualify this claim. They suggest that expertise in attention demanding cognitive tasks, such as proofreading, depends upon a highly developed skill that involves attentional allocation. Furthermore, Study 2 reveals that this aspect of expertise, evident in the anomalous word detection task, transfers to another, rather different, task which requires detection of visual scene changes. In other words, the ability to allocate focused attention, which is highly developed in skilled proofreaders, reflects a skill that is applicable to other tasks such as change detection.

We have been able to eliminate a number of different explanations for proofreaders' superior ability in these tasks, all of which do not involve attentional allocation. One of these maintained that the performance difference between experts and novices (i.e., in the anomalous word detection task) reflects differences in lexical knowledge. Another posits differences in working memory span; for instance, a higher working memory span might enable proofreaders to exert more effective control over selective attention (de Fockert et al., 2001; Engle, 2002). Data in Study 1 casts doubt on these explanations. Another possibility, examined in Study 2, was that proofreaders' superior ability derives from exceptional skill in exhaustive search strategies. However, the fact that a performance advantage for proofreaders did not transfer to a visual search task that features exhaustive search allows us to eliminate this explanation. We also argued against a role for figure-ground segregation skills in this task. In the end, the most likely explanation for our findings in Studies 1 and 2 is that proofreaders are simply better than novices at allocating attention effectively, although neither the size of their attentional resource pool nor their search ability differs from those of novices.

We also collected post-experiment interviews with participants regarding their performance in the anomalous word detection task (Study 1). This revealed an interesting theme. Most professionals reported that in their proofreading of articles, they read an article several times changing the type of information to which they attended each time. For example, they pay attention to word surface features like forms on the first reading, then to semantic information of words and phrases on the second and third reading, and so on. The level of information involved and its ordering priority appeared to differ with individuals. Such comments were rarely heard from the novices. From this, it is possible to infer that professional proofreaders are good at screening the type of information to be attended.

Amount of experience should be an important factor, one worth discussing in expertise, however, is not discussed in this article. This is because the effects of years of experience tended to co-vary with age. Further study is needed to investigate the contribution of amount of experience in expertise.

Finally, this research provides a new insight into what is involved in expertise, the role of focused spatial attention in expertise and the generality of skills that rely upon allocation of attention. This finding might be also useful for exploring effective methods of skill acquisition.

## REFERENCES


Amano, S., & Kondo, T. 1999. Nihongo-no goitokusei [Lexical properties of Japanese]. *NTT Database Series* (1st ed., Vols. 1–6). Tokyo: Sanseido.

Asano, M., & Yokosawa, K. 2007. Proofreaders' characteristics of high-level vision. *The Japanese Journal of Psychonomic Science*, **26**, 29–37. [Manuscript in Japanese with English summary]

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. 2001. The role of working memory in visual selective attention. *Science*, **291**, 1803–1806.

Engle, R. W. 2002. Working memory capacity as executive attention. *Current Directions in Psychological Science*, **11**, 19–23.

Chase, W. G., & Simon, H. A. 1973. The mind's eye in chess. In W. G. Chase (Ed.), *Visual Information Processing* (pp. 215–281). New York: Academic Press.

Ericsson, K. A., & Kintsh, W. 1995. Long-term working memory. *Psychological Review*, **102**, 211–245.

Ericsson, K. A., & Smith, J. 1991. Prospects and limits of the empirical study of expertise: an introduction. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 1–38). Cambridge: Cambridge University Press.

Ericsson, K., A., & Williams, A., M. 2007. Capturing naturally occurring superior performance in the laboratory: translational research on expert performance. *Journal of Experimental Psychology: Applied*, **13**, 115–123.

Halpern, D., & Wai, J. 2007. The world of competitive scrabble: novice and expert differences in visuospatial and verbal abilities. *Journal of Experimental Psychology: Applied*, **13**, 79–94.

Jared, D., Levy, B. A., & Rayner, K. 1999. The role of phonology in the activation of word meanings during reading: evidence from proofreading and eye movements. *Journal of Experimental Psychology: General*, **128**, 219–264.

Kane, M. J., & Engle, R., W. 2003. Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, **132**, 47–70.

Levin, D. T., & Simons, D. J. 1997. Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review*, **4**, 501–506.

Morita, A., & Tamaoka, K. 2001. Effects of orthographic similarity in processing Japanese kanji. *Psychologia*, **44**, 237–249.

Morita, A., & Tamaoka, K. 2002. Semantic involvement in the lexical and sentence processing of Japanese Kanji. *Brain and Language*, **82**, 54–64.

NTT Communication Science Laboratories. 1999. *Goisu suitei tesuto* [Vocabulary estimation test]. Retrieved on 12/27/2007, from http://www.kecl.ntt.co.jp/icl/mtg/goitokusei/goi-test.html

Osaka, M. 1998. Yomi-to working memory [Reading and working memory]. In N. Osaka (Ed.), *Yomi: Nou-to kokoro-no Jouhousyori [Reading: Information processing of brain and mind]* (pp. 239–262). Tokyo: Asakura Shoten. [Manuscript in Japanese]

Osaka, M., & Osaka, N. 1994. Working memory capacity related to reading: measurement with the Japanese version of reading span test. *Japanese Journal of Psychology*, **65**, 339–345. [Manuscript in Japanese with English summary]

Pesce, C., Tessitore, A., Casella, R., Pirritano, M., & Capranica, L. 2007. Focusing of visual attention at rest and during physical exercise in soccer players. *Journal of Sports Sciences*, **25**, 1259–1270.

Potter, M. C., Stiefbold, D., & Moryadas, A. 1998. Word selection in reading sentences: preceding versus following contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24**, 68–100.

Rensink, R. A. 2002. Change detection. *Annual Review of Psychology*, **53**, 245–277.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. 1997. To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science*, **8**, 368–373.

Shimomura, M., & Yokosawa, K. 1995. The processing of constituent characters in Kanji words in proofreading Japanese sentences. *Psychological Research*, **58**, 51–60.

Simons, D. J., & Rensink, R. A. 2005. Change blindness: past, present, and future. *TRENDS in Cognitive Sciences*, **9**, 16–20.

Yokosawa, K., & Mitsumatsu, H. 2003. Does disruption of a scene impair change detection? *Journal of Vision*, **3**, 41–48.


(*Manuscript received December 29, 2007; Revision accepted April 7, 2008*)