Applied Ergonomics 41 (2010) 150–160

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Quality control in the optical industry: From a work analysis of lens inspection to a training programme, an experimental case study

Maryline Rebsamen [a], Jean-Michel Boucheix [a,*], Michel Fayol [b]

[a] Université de Bourgogne, Laboratoire d'Etude de l'Apprentissage et du Développement (LEAD-CNRS, UMR 5022), Pôle AAFE, Esplanade Erasme, 21065 Dijon, BP 26513, France
[b] Université Blaise Pascal, Laboratoire de Psychologie Cognitive et Sociale, (LAPSCO-CNRS, UMR 6024), 34 Avenue Carnot, 63037, Clermont-Ferrand, France

A R T I C L E I N F O

Article history:
Received 27 April 2007
Accepted 2 July 2009

Keywords:
Control
Expertise
Training

A B S T R A C T

A cognitive work analysis of quality inspection in the optical industry has been carried out in order to devise a training programme. The task concerned the inspection of high quality human eyeglass lenses. We conducted an experimental investigation of defect detection and acceptability decision-making tasks in 18 experts and novice inspectors. Detection and decision-making were investigated together and separately in two experimental sessions. We showed the effect of expertise on reaction times and errors, and we described the cognitive processes of novice inspectors. On the basis of the processing differences between the two groups, a training programme for new inspectors was devised and described. Finally, training effects were tested.

1. Introduction

The purpose of the applied quality inspection case study presented in this paper was to describe the cognitive processes involved in a difficult visual inspection task utilised by experts and novices, in order to build a training programme. Such a programme was considered a necessity for both professional development of trainees and quality enhancement by a major international optical industry factory. The material to be inspected consisted of corrective optical lenses for peoples' eyeglasses. Several key inspection stumbling blocks arose from the material properties, the work situation, and the lack of structured training sessions for trainee workers.

Paradoxically, as quality constraints have grown in most working situations in many European countries, ergonomic researchers' interest in inspection studies has dropped dramatically. For 25 years, very few studies were published on this topic. In other countries, such as the USA, for example, research into inspection (and training) has continued, particularly in aircraft inspection and maintenance, essential for safety (Gramopadhye et al., 1997a, b) in using technology properties (Gramopadhye et al., 1998, 2000) or more recently, in virtual reality (Vora et al., 2002).

The present case study had a more modest approach in terms of applied ergonomics. We used a quality inspection model (Drury, 1992; Drury and Chi, 1995) in an experimental investigation of defect categorisation (detection) and decision-making by experts and novices, inside the factory and outside the laboratory.

Quality inspection tasks in industry have evolved over time. Initially, a specialist department situated at the end of the production process performed a quality inspection (Stephaneck, 1966). Subsequent constraints placed by the "just-in-time" organisational approach contributed to a reorganisation of quality inspection. This task is now integrated with other tasks carried out in productions lines (François, 1989; Liévin and François, 1997).

It would be trivial to assume that quality inspection merely consists of searching for and recognising defects, and making a decision in respect of its acceptability within quality limitations (Drury, 1992). However, it does not consist of simple separation of non-defective products from defective products. Rather, it attempts to ensure that these products conform to a specific tolerance threshold, as defined by the quality inspection department. The inspector's task is complicated because tolerance thresholds have to be taken into account; post-detection; this consists of evaluating the defect in relation to specific standards and of making a decision based upon these standards. And sometimes conformity standards vary. These judgment and decision factors necessarily involve the "inspectors' subjectivity", which sometimes leads to insecurity and stress. Because inspection requires a high level of concentration, sustained attention, and prior training, operators frequently consider the task to be difficult.

2. Inspection in the optical industry

In the optical sector under study, lenses are produced by injecting polymer into special moulds. The inspection activity is

* Corresponding author. Tel.: +33 80 39 57 65; fax: +33 80 39 57 67.
E-mail address: jean-michel.boucheix@u-bourgogne.fr (J.-M. Boucheix).

0003-6870/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apergo.2009.07.004

part of an assembly – inspection task according to the chronology indicated in Fig. 1.

Lens inspection requires special physical conditions, particularly in terms of lighting. In the real working situation, each inspected lens is brought into the inspector's field of vision, as he/she sits facing a "black box", containing a fluorescent tube in front of which the lens is examined (Fig. 2). The lenses are round and transparent; the defect to be inspected could be located on the external surface of the lenses or inside. The size of the lenses may vary from 65 to 80 mm in diameter and from 10 to 16 mm thickness. There were a number of different defect categories.

Every lens is scrutinised. In order to clear the lenses of dust, material projections, and other particles, on their surface, the workstations are equipped with air blow guns (compressed air) and lenses are systematically dusted with special fabric cloths (imitation suede). The inspection is carried out alongside this procedure. This task appeared difficult and often a source of "slip-ups" for several apparent reasons, beginning with statistical studies, by the Quality Department and the operators.

The product is itself a source of uncertainty. A lens presents a certain thickness and a certain curvature, both of which vary. At times, lenses provide the same perceptive result as a magnifying glass, and the defects are all the more difficult, not only to track down, but also to locate in the mass and in the area of concern (core) of the lens. Another source of difficulty stems from the defects sought. Indeed, if some defects are easily detectable because they are very noticeable, the majority are very small, some tenths of millimetres. In addition, they are extremely diverse ("micro-grooves", free jet, white fleck, pollution, batch stone, cotton, etc…) and can assume various forms. Lastly, with regard to the standards established by the Quality Department, it is necessary to decide whether or not the defect is acceptable.

The decision depends on combining three parameters. The first of these is the default intensity. The Quality Department calibrated four levels, the sizes of which vary from a few tenths of a millimetre and all are less than a millimetre. The intensity is measured only for one or two types of defects (these are the most frequent), the other defects being unacceptable, irrespective of their intensity. The second parameter is the part of the lens containing the defect. Lenses have been "virtually" divided into three concentric zones ranging from zone 1 to zone 3 (Fig. 3). Zone 1, tolerates fewer defects than the other zones and those which are less serious. Zone 3, tolerates the most defects in terms of number and intensity, because the lens is systematically re-cut by the optician. Zone 2 is the intermediate zone between the two other zones. The third parameter is the limited number of acceptable defects in each of the zones and in general, on the lens' surface. Finally, two constraints are added: one is the monitoring time, limited by the production targets (a few seconds per lens, 5 s being standard); the other is the variation of the criteria for the rigorousness of the inspection standards. While some types of lenses are required to be completely free from defects or permit extremely minor defects, exclusively in zone three, other types of lenses may comprise more marked defects, even in zone two.

2.1. Why were we asked by the factory to carry out the study presented in this paper? And what was the problem to be solved?

In response to the first question, there were three main factors, with combined effects. Firstly, immediately prior to commencing the study, the factory in which this analysis took place became an international pilot site where 400 workers were employed. Thus a great number of inspection trainees had to be trained within a short period of time. The factory was recently involved in obtaining a quality certification ("ISO 9002") which required employees, particularly inspectors, to reach a higher level of skill. At the same time, a new automated process of industrial mass production led to an important increase in the factory's output. Secondly, the initial task of the employees at this stage of lens production was the dismantling and re-assembly of the mould containing the lens. Thus, the task of inspection was added to the initial task. According to the inspectors' management staff, these changes did not affect experienced or expert inspectors performance, but did change to a considerable degree the performance of less experienced employees or novices. Thirdly, and this is the main reason, many inspection errors were observed, even after a long training period (more than several weeks). Due to the new organisation, including the new automated process, the average rate of defects was neither known nor expected by inspectors. Defects were not rare, so inspectors were really likely to find many defects. Within the same category of defect some were acceptable and in contrast others were not; and, depending on their size and on the zone where the defects were situated, i.e. on (or within) the lens, (Fig. 3) inspectors were required to alter their decision criteria. As indicated above, this implied that inspectors had to make a decision about which zone on the lens contained the defect. Moreover, the tolerance threshold (for size and zone, as well as for the different categories of defects) varied across different types or different

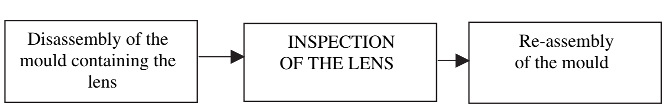


Fig. 1. Context of the inspection task.

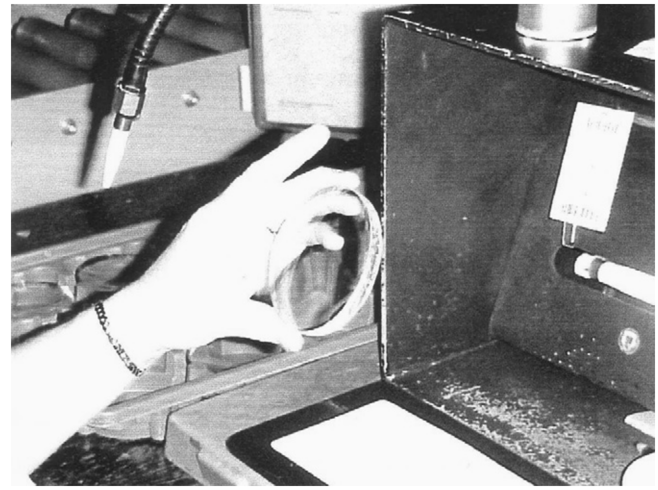


Fig. 2. Lens inspection at the workstation in the optical industry.

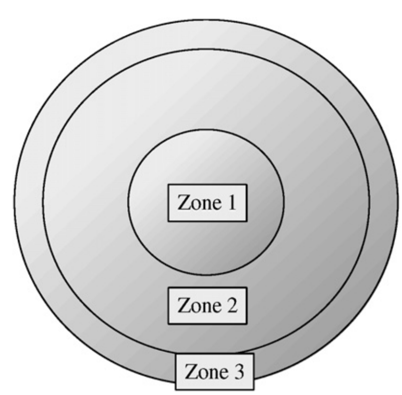


Fig. 3. Virtual lens areas.

series of lenses (mainly depending on their quality). The types of defects were not equally visible (depending on the type of lens): cotton was probably easier to detect than a white fleck.

2.2. How were inspectors trained before we were asked to design their training sessions?

Preceding our experiments, several interviews were held, both with beginners and experienced inspectors and with supervisory inspectors for orientation purposes. These preliminary investigations revealed that beginners recently appointed to the new inspection workplace complained about not having sufficient time to carry out the inspection. They explained they were often overloaded and made the inspection without sufficient knowledge of the accuracy of their decision. This uncertainty led to bad feeling, and we may assume that inspectors stopped searching, given the time normally allocated to this task on the production line. Interviews with experienced inspectors and “experts” also revealed interesting data regarding the traditional training method, and with respect to the environmental and material aspects of inspection task. Training was mainly based on “learning on the job” without explicit instructions concerning the visual criteria for inspection and the defects. Explicit instructions concerned more general conditions of the inspection: the best way to hold the lens in front of the fluorescent tube, the best physical position to adopt during the task (with respect to eye vision performance), the best way to dust the lenses with a blow gun, general rules for accepting a defect. This kind of information is easily verbalised; it is declarative knowledge (Anderson, 1983). In contrast, more cognitive aspects of control were never described: the way to explore the lens, how to detect the defect, or how to correctly assess its severity. This second type of information is not easy to verbalise because much of it is concerned with procedural knowledge: for example, experienced inspectors said that “they did not search for the defect, because the defect sticks out a mile”. This idea of a “stimulus that sticks out a mile” has been described by Treisman and Gelade (1980). Stimulus discrimination does not require searching or concentration of attention when the target is perceived as very different from the background on which this target appears. Expertise could produce an effect in which, for the inspector, the defect stands out more than the non-relevant stimuli on the lens (superficial dust, reflection, acceptable category of defect). Briefly, traditional training provided more information about the general purpose of the inspection task than instructions and specific practical exercises concerning how to go about the task of inspection. In this way, we may hypothesise that differences between experts and novices could not be explained by self-selection effect alone, due the difficulty of the task. Finally, the actual workstations were not confined to a contained space but were situated in large high-ceilinged, airy workrooms.

Thus, for the factory, the goal was to optimise training sessions as well as the learning content thereof, all within a short time scale. Therefore, in consideration of the context presented above, this goal was rather crucial.

3. Studies and models relating to quality inspection

Studies on quality inspection led by Drury and his collaborators since the 1970s have greatly contributed to providing an awareness of the fact that the human operator, even if he/she is not infallible in inspection tasks, is nevertheless more effective in decision-making than most automatic systems (Drury and Fox, 1975; Drury and Sinclair, 1983) and can improve his/her performance with specific training (Czaja and Drury, 1981; Duchowski et al., 2000; Gallwey and Drury, 1986; Gramopadhye et al., 1997a, b, 1998, 2000; Vora et al., 2002).

The training of inspectors has taken on a more important role in new industrial organisations, when it rapidly became clear that there was a lack of experienced inspectors. Cognitive processes related to inspection were modelled by Drury (1975) and Spitz and Drury (1978) in a series of studies on a glass inspection task. These authors broke the inspection activity down into two steps: detection, which implies categorisation, and decision-making. The increased demand for inspectors required the implementation of a specific type of training that would allow trainees to rapidly acquire sufficient skills to attain a level of inspection performance close to that of experts trained “on the job”.

In order to study lens inspection, the inspection model developed firstly by Drury (1975), Spitz and Drury (1978) can be used (Fig. 4). This model is based on two processing stages: visual search activity for detection and decision-making. However, it is possible that there is some to-ing and fro-ing between the two stages.

The detection stage involves a categorisation process based on the activation of high quality representations of defect categories. In the optical industry, fragments of material or dust frequently stick to the lens, which is noise, giving the appearance of a defect. If the detected signal is not a defect, a new scanning of the lens is often initiated. When no signal is detected, the acceptance decision is probably made progressively in relation to the amount of time spent examining the lens. For this reason there is an extra box in the model (Fig. 4) named “is available time used up?” which can either return to the visual scanning phase or go down to acceptance. The detection and decision-making steps are thereby more or less combined up until the final decision, particularly in the case of experts, who use functional, goal and task related constraints and representations. The signal must then be assessed in the light of the acceptance standards in order to determine if any lens remains within the acceptance boundaries, or if they are located outside these boundaries. The underlying first model of inspection has its origin in signal detection theory (Tanner and Swets, 1954; Sylla and Drury, 1995).

4. What is expertise in lens inspection?

Our goal was to study the cognitive processes involved in lens inspection, in respect of the degree of operator expertise, in order to use the results in the development of training programmes (Boucheix, 2003, 2004). Most studies on expertise simply emphasise the fact that experts succeed better and faster than trainees in the execution of the task in which they are experts. We should therefore start with a definition of expertise. However, there are many different approaches to understanding expertise, depending on the field in question, and many requirements for applying this comprehension. As a result it is difficult to provide a general and final definition. Reviews of psychological characteristics and experts' strategies drew attention to a series of features; some of them seem relevant to inspection (Glaser and Chi, 1988; Shanteau, 1992; Cellier et al., 1997; Farrington-Darbi and Wilson, 2006). Experts have extensive and up to date knowledge. They have a highly developed sense of perception and ability for concentration. Experts also have a highly developed a sense of what is relevant when making decisions. They are able to simplify complex problems, and act better in the face of adversity. Experts are better at identifying and adapting to anomalies and exceptions and they are also capable of adapting their strategies to changing task conditions. They show a strong sense of responsibility. Experts make small errors; they are used to avoiding significant errors. Experts excel mainly in their own field. They perceive significant meaningful patterns in their field, and they have strong self-monitoring skills. Experts have greater anticipation skills, with a more global functional view of the situation and take a wider range of data into account in diagnosis. Experts encode new information more quickly and completely, and are deemed to have a greater repertoire of strategies.

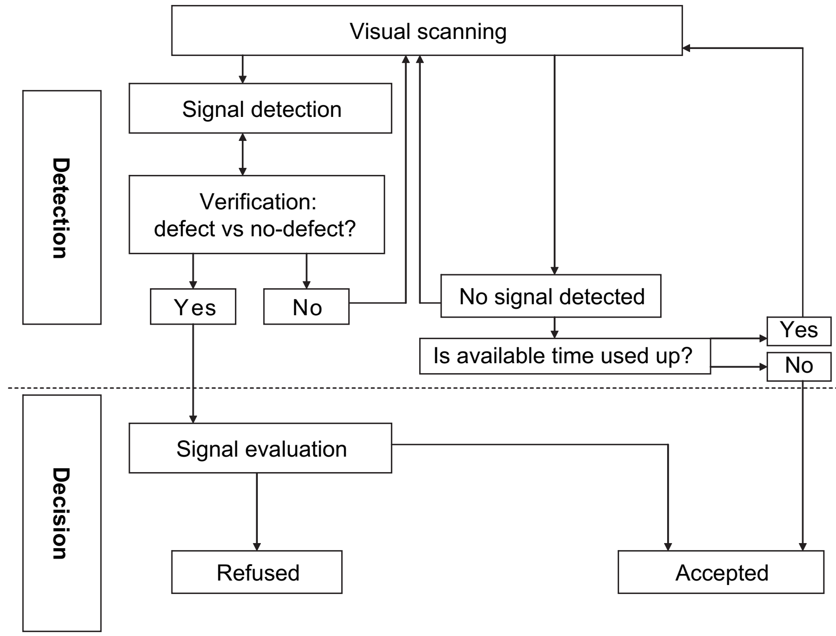


Fig. 4. Visual inspection model for lenses inspired by Drury and co-workers (1975, 1992, 1995).

Ericsson and Lehaman (1986) have highlighted the fact that expertise consists of maximum adaptation to the constraints of the task in hand. Vicente and Wang (1998) demonstrated the constraints attunement hypothesis in experts. For example, Biederman and Shiffar (1987) compared the performances of novices and experts in a task consisting of determining the sex of chickens. The observed configurations were very poorly distinguishable. However, experts were able to classify 1000 chickens per hour with a 98% rate of accuracy.

At a perceptual level we can bring closer inspection tasks and studies about medical diagnosis. Myles-Worsley et al. (1988) compared different levels of expertise in a study on the recognition of chest X-rays. Participants had to distinguish between normal and abnormal images after an initial sensitisation phase. Novices showed a lower recognition rate and failed to distinguish between normal and abnormal films. For the first-year professionals, the recognition rate for the two types of film was improved compared to the novices. For the expert group, the rate of recognition for the abnormal films increased in proportion to the experience of the subjects, from 24% after 4 years of experience, to 35% after 22 years. However, this tendency was reversed for normal films. Experts may have lost the ability to recognise normal attributes. In fact, it is likely that they have perfectly integrated the notion of normal attributes, and as a consequence, they no longer process this information. The very brief time period given to participants to memorise the attributes (500 ms) suggests that the experts only processed the relevant stimuli, meaning the abnormal films. So expertise seems to depend upon a deep understanding of normal and abnormal attributes, since processing of the former was automated (which explains why they were not memorised). We were inspired by this study and by the researches of Drury and colleagues to develop the experiment reported below.

5. Method

5.1. Experimental overview

Two groups of participants who differed in their level of expertise were tested. The study took place in the factory. We reproduced the workstation (as in a full scale simulation situation) in the same area as the true working stations. All the same real physical conditions were maintained, including the same material. Fig. 2 illustrates what the inspector did, in the experiment, from the moment he first looked at the glass to the moment of making a final decision in respect of that particular lens (see also the model Fig. 4). We gathered and indexed a panel of lenses that presented variable characteristics depending on presence or absence of defects.

Our panel was a representative sample of the real defect frequency, made up of lenses having either a single defect or no defect at all. In the real working situation, there the appearances of different kinds of defect were interdependent. Then, a single type of defect could be seen (involving a specific type shining point in the lens) the location of which on the lens and its intensity were subject to variation. According to this location and intensity of the defect, the lens is deemed to be either acceptable or non-acceptable. Our panel included four degrees of lens acceptability: Good lenses: without any Defect (GWD); Good lenses: with an Acceptable Defect (GAD); Bad lenses: with non-acceptable Limited Defect (very close to acceptability) (BLD); Bad lenses: with a non-acceptable Non-limited Defect (BND). Experts (quality inspectors) validated our panel. In order to increase the relevant features of real work lens inspection, the degrees described above included some disparities, which is to say that, no lens, even among those having a "non-limited" defect, displayed a very significant defect. The hypothesis tested was that the "limited" lenses (GAD and BLD) would be more difficult to process than the "non-limited" lenses (GWD and BND).

In order to study detection and decision-making separately, we used a dual time experimental procedure and varied the specific instructions between the two time periods. In the first time period of the experiment – detection task, the participants were instructed to carry out a simple detection – "*to indicate whether or not the lenses contained a defect, without having to decide upon the acceptability of the lenses*". In the second time period – inspection task, taking place just after the first task and which included detection + decision, the instructions were to carry out a traditional inspection, that is, – "*to decide whether or not the lenses were in* *conformance with the acceptance standards"*. Therefore, in this detection + decision task, it was not possible to make a decision without carrying out the detection stage beforehand.

Thus, we wished to store data during the detection stage alone, as well as on the detection and decision-making stages. The differences in performance between the two tests could provide information about the judgment/decision process, particularly with respect to its duration. If the two processing stages succeed each other, the inspection time period (detection + decision) should have been significantly greater than the detection time period, and the time necessary for the decision should be added to that of the detection.

Three categories of performances measures were used. The two first were response time (reaction time) for each lens and errors (false alarms and omissions). Errors in judgment could have four possible sources: errors in the evaluation of the area where the defect is located, errors in the evaluation of the extent of the defect, errors in applying the inspection standard (poor extent/area matching) and finally errors in not seeing a defect and therefore reporting the lens as being good. We also performed a Signal Detection Theory (SDT) analysis, calculating for each group values of d' and β.

5.2. Participants

Eighteen operators (all female) working in visual inspection stations agreed to participate in the study. Nine of them were expert inspectors and nine were non-expert operators ("the least experienced among available subjects", Alwood, 1986; Cellier et al., 1997). Experts were inspection supervisors (they had more than 15 years of experience), non-experts had been "in-line" operators for a short time only (having less than one year's of experience, but without really being complete learners). The participants we labelled as experts in this experiment were experienced inspectors with extensive practice of inspection. Because of their inclusion by the factory in the definition of the categories of defects, of the standards, as well as in the design of the quality process and in the training sessions, these inspection supervisors appeared to be closer to a group of experts rather than an experienced only group.

5.3. Materials

The experimental material was composed of a panel of 50 lenses including 20 lenses without defect (GWD), 10 lenses with an acceptable defect (GAD), 10 lenses with a non-acceptable limited defect (BLD), and 10 lenses with a non-acceptable non-limited defect (BND). These lenses were randomly selected and numbered from 1 to 50. The material also included a visual inspection station with light conditions in conformance with current standards (Fig. 2), a bottle of acetone and suede cloth for the detection test, a zonal chart and a grease pencil for the inspection test. Data sheets were required to collect data for each lens. Opposite each lens number, the experimenter timed (in seconds) the response and recorded each item of data on the sheet booklet. A precise stop-watch was started when the inspector moved the lens before the lamp and stopped when he began his response.

5.4. Procedure

Each participant took both tests under the same conditions and at the same station. The same panel of lenses was presented to them with a week-long interval between each test. The detection test was taken along with the inspection test (detection + decision tests). For the detection test, the instructions given to participants were to say, after examining each lens, if it contained a defect or not. We asked for a *"yes or no"* response for each lens. For the inspection test (detection + decision tests), the same participants were seen again. This time, the instructions were to carry out the inspection, that is to say, to decide if each lens was acceptable or not, after examining it. We asked for a *"good versus bad"* answer.

6. Results

6.1. Response times

Mean and standard deviations according to the two tests and to the categories of lenses are provided Table 1. Data was analysed using a 2 × 2 × 4 ANOVA (*group* by *test* and by *lens*). Analyses were only related to the times associated with correct responses.

Firstly, Expert group ($m = 6.43$ s) was faster than Non-Expert group ($m = 9.02$ s), $F(1, 16) = 5.63$, $p = 0.03$.

Secondly, the task factor (detection vs detection + decision) had a significant effect on response time. Inspection (detection + decision) took longer time than detection, ($F(1, 16) = 26.23$, $p < 0.001$). However, there was not much difference between detection and inspection (detection + decision) for GWD lenses, but levels were higher for GAD (Non-Expert group, $t(17) = 2{,}9$ $p < 0.02$; Expert group, $t(17) = 3.55$, $p < 0.01$); for BLD (Non-Expert group, $t(17) = 4.6$, $p < 0.01$; Expert group, $t(17) = 3.27$, $p < 0.01$) and for BND (Non-Expert group, $t(17) = 3.39$, $p < 0.01$; Expert group, $t(17) = 2.71$, $p < 0.02$).

Thirdly, the highest mean times were those involving good lenses with no defects; times then progressively decrease for the lenses displaying a "limited" defect. The shorter mean times were those observed for the lenses with a "non-limited" defect, $F(3,48) = 21.08$, $p < 0.001$. This result is in accordance with a prediction from Drury's model, which hypothesises that the inspector will keep inspecting for defects until available time is used up (Fig. 4). In contrast, the detection of a defect will terminate inspection immediately.

Interactions between the group factor and the other two factors task type and lenses, did not show any significant effect. However, interaction between the task (detection and inspection) and lens

Table 1
Mean response time (seconds) according to each category of lenses and to the task type (detection versus inspection – detection + decision) for experts and non-experts inspectors.

		Categories of lenses				
		Good without defect	Good acceptable defect	Bad limited defect	Bad non-limited defect	Total
Detection	Experts	7.18 (4.14)	4.46 (1.68)	4.65 (2.21)	3.2 (1.74)	4.87
	Non-experts	10.41 (3.68)	7.39 (1.58)	5.3 (1.38)	4.13 (1.40)	6.81
	Total	8.79	5.93	4.98	3.67	5.84
Inspection (det + dec)	Experts	7.39 (2.18)	9.64 (4.03)	8.73 (3.02)	6.17 (2.77)	7.98
	Non-experts	10.98 (4.63)	12.42 (5.56)	13.86 (5.41)	7.65 (2.78)	11.23
	Total	9.18	11.03	11.30	6.91	9.60
Total		8.99	8.48	8.14	5.29	7.72

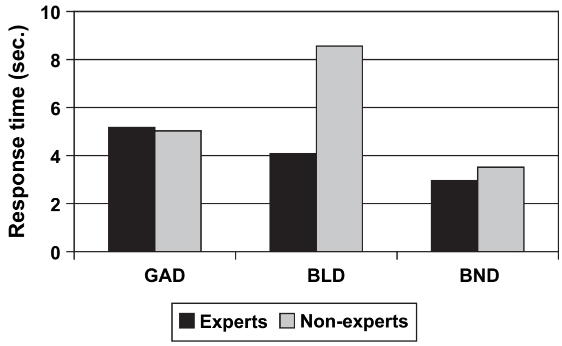


Fig. 5. Mean time responses differences (seconds) between the detection task and the inspection task (detection + decision) according to the category of defect and expertise.

factors was significant ($F(3,48) = 16.29$, $p < 0.001$). This effect highlighted the fact that inspection times (detection + decision tests) were longer than detection times for GAD, BLD, and BND lenses. It should be noted that detection times were all the more significant when the defect was less extensive, since BND lenses were scanned for the least amount of time and GAD lenses were scanned for the greatest amount of time. On the other hand, "limited" lenses (GAD and BLD) required the most time during the inspection test (detection + decision). The least amount of time was spent on the BND lenses during both detection and inspection (detection + decision); these lenses gave rise to less uncertainty. Expertise matters little when there is only a small amount of uncertainty.

Three-way interaction (group, test, lenses) was marginally significant ($F(3,48) = 2.78$, $p = 0.051$). Then, we performed statistics on the differences between the two groups for each type of lens and for each test. This outcome indicated that experts were able to detect less extensive defects (GAD) quicker than novices ($t(17) = 3.82$ $p < 0.001$). They marginally responded quicker when the lens did not have any defect (GWB), ($t(17) = 1.75$ $p = 0.09$). But, there was no significant difference between the two groups for the bad lenses BLD and BND. The results from the inspection test (detection + decision) times illustrated that the expert's inspectors tended to give a quicker response for GWB ($t(17) = 2.11$, $p = 0.051$) and for BLD ($t(17) = 2.48$, $p = 0.024$), but not for GAD ($t(17) = 1.21$, $p = 0.24$) and BND ($t(17) = 1.12$, $p = 0.27$). Bad lenses with limited defect seemed to create the most problems for non-expert inspectors.

Finally, our second test (inspection test: detection + decision) did not itself allow us to obtain information on the decision-making activity alone. But, since in each test we were dealing with the same panel of lenses and the same participants, we subtracted the mean response times for the first test (detection) from those of the second (inspection: detection + decision). This subtraction was performed in order to obtain an "estimate" for the time required for judgment alone when dealing with the GAD, BLD, and BND lenses. Fig. 5 illustrates the results of these subtractions. Decision times were not different between experts and non-experts except for BLD lenses: the mean decision time required for lenses with defects within non-acceptable limits for the non-expert group had a mean value 4.48 s longer than that of the expert group (Post-Hoc test, $p = 0.023$). In the case of these lenses, which were borderline acceptable, experts decided to reject them more quickly.

6.2. Errors

Means and standard deviations of errors (expressed in percentages) according to the two tests (detection and inspection) and to the categories of lenses are provided in Table 2. Data were again analysed using a $2 \times 2 \times 4$ ANOVA.

Expert group made more errors than non-expert group (respectively, M Expert = 19.44% and M Non-Expert = 14.58%), ($F(1,16) = 5.59$, $p < 0.05$). Test factor (detection vs inspection) had no effect on the error rate–mean error rates for the two tests were fully equivalent. However, breakdown of the errors was not the same depending on test and lens type. Total error rate was greater for "limited" lens containing either an acceptable defect (GAD) or a non-acceptable defect (BLD) than for non-limited defects (GWB; BND, ($F(3,48) = 15.06$, $p < 0.001$)).

Interaction between test and lens factors revealed significant effect ($F(3,48) = 16.62$, $p < 0.001$). GAD lenses were difficult to detect but once detected easier to judge, whereas BLD were easy to detect but more difficult to judge as non-acceptable. Interaction between test (detection vs inspection) and group was not significant, $F(2,48) = 2.12$, $p = 0.11$.

At this point, we distinguished false alarms and omissions. In the detection test, errors relating to lenses with no defect (GWD) were false alarms and errors relating to three other lens types (GAD, BLD, and BND) were omissions. In contrast, in the inspection test GWD and GAD lenses were acceptable, so errors relating to these lenses (rejects) were false alarms. BLD and BND lenses should have been rejected and errors relating to these lenses (accept) were omissions. The rates of false alarms and omissions for each group are presented in Fig. 6.

Omissions rates were always higher than false alarms rates, ANOVA ($F(1,16) = 14.2$, $p < 0.001$) and a significant interaction between error types and levels of expertise was found, $F(1,16) = 7.19$, $p < 0.02$. In both tests, the rate of false alarms was higher for Non-Experts than for Experts (error rates for GWD lenses were respectively 17.7% for Non-Experts and 10.5% for experts in the detection test; respectively 4.44% and 1.7% in the inspection – detection + decision test). The expert group was more efficient than the non-expert group in recognising a good lens.

For GAD lenses, errors rates were higher in the expert group than in the non-expert for detection test (omissions), 47.8% versus 30%; but on the other hand, for inspection test (detection + decision) error rates (false alarms) were higher in the non-expert group (17%) than in the expert group (10%). This result was close to the findings of Myles-

Table 2
Mean percent of errors according to each category of lenses and to the task type (detection versus inspection – detection + decision) for experts and non-experts inspectors.

		Categories of lenses				
		Good without defect	Good acceptable defect	Bad limited defect	Bad non-limited defect	Total
Detection	Experts	10.55 (1.04)	47.8 (1.2)	15.55 (1.23)	5.55 (1.01)	19.86
	Non-experts	17.8 (1.6)	30 (2.55)	5.55 (1)	3.33 (0.5)	14.17
	Total	14.17	38.9	10.55	4.44	17.01
Inspection (det + dec)	Experts	1.66 (0.35)	10 (1.73)	43.33 (2.6)	21.11 (1.9)	19.02
	Non-experts	4.44 (0.5)	16.66 (1.5)	31.11 (1.9)	7.8 (0.97)	15.01
	Total	3.05	13.33	37.22	14.45	17.02
Total		8.61	26.11	23.88	9.44	17.01

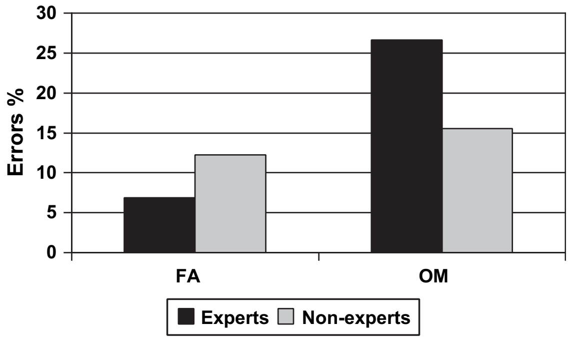


Fig. 6. Mean rates of false alarms and omissions according to the level of expertise.

Worsley et al. (1988) showing that experts did not process normal displays (detection) and confined cognitive processing to the most relevant features of the task (inspection = detection + decision).

For BLD lenses, fewer errors were produced in the detection test (10%) than in the inspection test (detection + decision) (37%). This result obtained for ambiguous lenses, showed that detection could be easier than full judgment, which implied precise updated knowledge and strict application of acceptability standards. Similar patterns of responses were found for BND lenses, with fewer errors.

In order to shed further light on the nature of the difference between experts and novices, a signal detection theory analysis was carried out.

6.3. Signal detection theory analysis

Given that we collected "omissions" and "false alarms" for our subjects, it was quite possible to calculate values d' and β. These values are statistically derived from errors and good answers (especially false alarms and omissions) as may be seen in the model indicated in Table 3. Indicator d' measures the discriminability of the subjects from noise (their ability to detect a flaw). It is a sensibility criterion[1]. Indicator β measures their criterion value – their willingness to decide the flaw is a defect. It is a response or a decision criterion[2]. Differences in d' between the experts and the novices would imply either that perhaps experts are self-selected, or that they have indeed improved their ability to search lenses through practice. Differences in β would indicate that novices are not yet good at decisions – for example, that they are too cautious.

We performed d' and β analysis for each test studied in the experiment: detection only and inspection (detection + decision). However, in the detection test, there was no decision on the acceptability of the default, thus in this task d' only was the most relevant criterion. In contrast, in the inspection task which included detection and decision about acceptability, d' as well as β appeared to be relevant criteria. However, because the detection part of the task involved for the second time the same lenses as the first detection only task, the d' measure for the inspection task appeared less reliable than for the first detection task. In contrast, in the inspection task, β was the most relevant criterion.

In the detection test, we have to remember participants had only to detect the presence or absence of a defect for each lens irrespective of their acceptability. The Participant was asked for a "*yes or no*" response for each lens. In this test the "no" answers concern GWD lenses only and "yes" answers were relative to the other types of lenses: GAD, BLD and BND: d' and β were performed on the basis of this opposition (GWD vs the other type of lenses). Results showed respectively d' value of 4.03 for novices and 7.58 for experts; and β value of 4.64 for novices and 8.58 for experts. Then, in a "pure" defect recognition test (GWD vs GAD, BLD, BND) expert seemed to outperform novices for the two criteria: discriminability and decision. However, here, the similarity of the results for the two criteria is not surprising, because in the detection test only, there was no decision on the acceptability of the defect (in contrast to the detection + decision task). So, in this detection only task, we measured discriminabilty of the subjects from noise.

In the inspection test (detection + decision), we need to remember participants had to carry out the complete inspection and to decide if each lens was acceptable or not, after examining it: "*good versus bad*". In this second test "good" answers concerned GWD as well as GAD lenses and "bad" answers concerned BLD and BND lenses: d' and β were performed on the basis of this opposition (GWD and GAD vs BLD and BND).

Results revealed respectively d' value of 8.29 for novice group and 10.65 for expert group; and β value of 8.43 for novice and 16.24 for experts. In this second test d' values seemed different between expert and novice but not too distant while β values appeared more different between the two groups. However we have to remember that d' values were less reliable in the inspection task than in the detection task only. Here β was the most relevant information.

7. Conclusion: from experimental work analysis to training programmes

In this concluding section, we will firstly summarise and discuss the main results of this research; and secondly we will describe how we built the training programme which followed this applied case study. We will also briefly summarise the main results of the training programme.

Using the model employed by Drury (1975) and Spitz and Drury (1978) we checked that, for this case study with lens materials, the two well-known processing stages during the inspection activity were sequential. Since the mean inspection time was 9.6 s and the mean detection time 5.84 s, it may thus be deduced that a mean value of 3.76 s is necessary to make a decision after detection of a defect. This result confirms a sequential model.

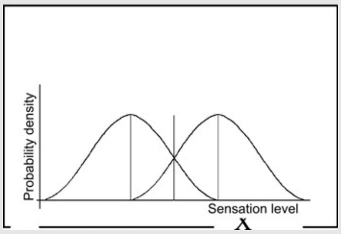
Results from response time revealed that expert inspectors took less time to provide an exact response than the less experienced inspectors. From good and bad answers, it appeared that expert inspectors have a precise memory representation of what constitutes a normal lens configuration, with the end result that they are no longer able to process good lenses, even if they are explicitly asked to do so. No doubt this ability to restrict information processing to pertinent information, merely for the sake of the task's efficiency, is based on the superiority of the experts' performance. Thus, it is a matter of maximum adaptation to the task constraints, since inspectors are subject to time constraints in actual working conditions. So, it seems that what is acquired with experience is the ability to recognise normal and abnormal configurations and to process only the latter, without losing any time on the former. In the case studied here, even when defects were uncovered, decision-making, which implied using precise updated knowledge about contextualized acceptance standards, could remain very difficult for both groups.

Results from Signal Detection Theory analysis brought interesting data about the difference between experts and non-experts regarding the tasks of detection and decision-making. We found differences in d' between the experts and the novices. This result

[1] Sylla and Drury (1995) showed that $d = \text{mean}\ (\text{fsn}(X) - \text{mean}\ (\text{fn}(X)/\sigma$; or in terms of probability: $d' = p(s/s) - p(s/n)/\sigma$. Variances of the two distributions have to be the same.

[2] Sylla and Drury (1995) and Tanner and Swets (1954) showed that $\beta = p$ (acceptance$/p'$) or $\beta = p(s/s)/p(s/n)$.

Table 3
Payoff matrix from attribute inspection with four possible answers in the SDT and distributions of the evidence variable X in the payoff matrix, under condition of noise and signal, from Sylla and Drury (1995) and Tanner and Swets (1954).



	True state of item	
	Good (noise: n), Fn(X)	Faulty (signal + noise: $(s+n)$), Fsn(X)
Accept	Correct accept, probability $P(n/n)$	Omission, probability $P(n/s)$
Reject	False alarm, probability $P(s/n)$	Correct reject, probability $P(s/s)$

implies that the experts have indeed improved their ability to search lenses through practice. However, we cannot exclude the fact that experts might have been self-selected. If we consider the former alternative as the most parsimonious[3], then we have a good basis for training their search of defects on lenses. *This search and detection training in novices could be improved by first acquiring more specific representations of the categories of defects, and precise memory representation of what constitutes a normal lens configuration.*

We found also an important difference in β between experts and novices, in the inspection task; which indicate that novices are not yet good at decisions. The reason is that they seem too cautious in comparison with experts.

This result also fits well with the fact that error rates for GAD lenses, in the inspection, were higher in the non-expert group than in the expert group. For BLD lenses, fewer errors were produced in the detection test than in the inspection test. This result obtained for ambiguous lenses, showed that detection could be easier than full judgment, which implied precise updated knowledge and strict application of acceptability standards. The main feature of acceptability standards is the intensity of the defect. This feature seems to be more critical that the zone where the defect is situated, because a zonal chart is always (and quickly) available in the workplace. There is no such tool for intensity. *Decision training in novices could be improved by updating knowledge of acceptability standards. This implies learning intensity of defects.*

On the basis of these conclusions, deriving from the experimental work analysis, we designed a training method and programme. The first goal of inspection training, more directly related to the detection phase, should be to improve the acquisition of precise representations of the categories of defects and of what constitutes a normal lens. The second goal, which is more related to decision phase, should be to provide specific knowledge about acceptability standards and particularly intensities of defects.

The training programme which is presented below consisted of two experimental learning stages in a simulated situation (with a real, full scale, workstation) and one control-training stage in a real working situation. The first two stages were experiments designed to check respectively, the effect of the acquisition of representations of categories on detection and the effect of learning intensity standards on decisions. The last stage was real work training for new trainee inspectors, attending for real working place situation.

All three phases were also conducted as applied experiments from which we registered quantitative (reaction times and errors) and qualitative data. In the following section, we will present the main issues and results of the learning programme.

7.1. Learning categories of defects

The goal of the first experiment concerned defect categories learning, in order to build in long-term memory, precise and stable representations of defects, forms and categories. We based this defect category-learning phase on two factors: the distance from the prototype for each defect category (derived from the theory of categorization of Rosch, 1978) and the presence of immediate or delayed feedback on the categorisation response. Previous research on inspection, and perceptual tasks close to inspection, showed that feedback on learners' categorisation answers enhanced trainee performances (Czaja and Drury, 1981; Gramopadhye et al., 1997a, b). In the experiment, learners had to process multiple series of trials, with lenses containing defects close or far from the prototype of their category.

Participants were 30 new recruits (21 females and 9 males) without any inspection experience. The experiment was carried out in the factory.

The material comprised a complete individual workplace station, two samples of lenses including a limited series of 5 categories of defects: one prototype sample (5 categories X2 prototypes = 10 lenses, selected by expert inspectors) designed for learning; and one test sample (20 lenses), including exemplars more or less distant from the prototype with four distance levels (selected with expert inspectors), for each of the five defect categories. The five defect categories were: "micro-grooves", "free jet", "white fleck", "pollution" and "batch stone".

We performed the following three-step procedure in individual learning session (mean duration: 45 min.): (1) main inspection gestures learning; (2) explicit categories learning (with prototypes); (3) test phase (identification of exemplars identical or distant from the prototype for each category). For the test phase, four randomised trial blocks were performed. Half the group ($N = 15$) were given immediate feed-back after each participant answer. Half the group ($N = 15$) were given delayed feed-back (after each block). Response times and answers were recorded. Results on good answers in the test phase are indicated in Fig. 8.

Significant effects were found for feed-back ($F(1,28) = 5.37$, $p = .027$) in favour of immediate feed-back; for learning blocks ($F(3,84) = 10.05$, $p < 0.001$); and for distance from prototypes

[3] At the beginning of this research, we first proposed to both novices and experts to perform visual discrimination aptitude tests. Results of these first investigations showed no differences between experts and novices regarding pure visual search and discrimination ability.

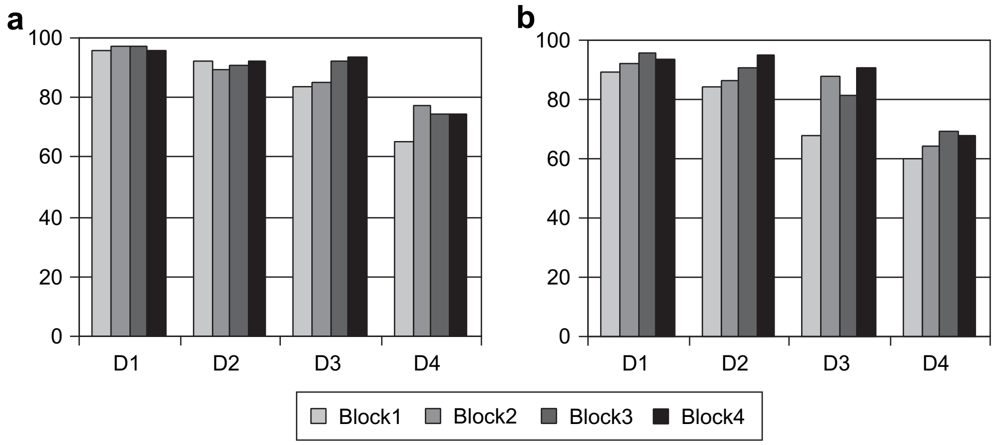


Fig. 7. Learning Categories of Defects. (a) Good answers (%) with immediate feedback; (b) good answers (%) with deleted feedback.

($F(3,84) = 47.08$, $p < 0.001$). However, an interaction between feedback group and blocks showed that the effect of the feedback was limited to the first learning block. Similar results were found with time responses. To summarise, learning precise and various exemplars of categories (based on the prototype theory) was effective in recognising defects, with more than 85% of good answers in one "short" session.

7.2. Learning intensity

The goal of the second learning phase was to build precise mental criteria for acceptability standards. During the initial task analysis carried out at the beginning of our investigation in the factory, we often observed that intensity evaluation was the skill which took the longest time to be acquired. So, the experiment was devoted to acceptable or non-acceptable defect intensity learning. How were inspectors trained before our intervention for the evaluation of the intensity of defects? There was no real training. Inspectors were told about the general rules of intensity evaluation with a short exposure to an extensive external scale of intensity. We created a new yardstick of intensity classes, easy to handle, (Fig. 7) which fixed the boundaries of the four intensity levels.

As we had noted that experts did not use external scales of intensity, the goal of the experiment was to compare two training methods: training involving intensity judgment by systematic comparison of the defect with the new external yardstick (yardstick group) versus training involving judgment by systematic memorisation of the intensity classes (memory group).

Participants were 30 new recruits (different from those in the above experiment, including 22 males and 8 females) without inspection experience. The experiment was carried out in the factory. The material was composed of a complete individual workplace station and two samples of lenses showing defects within four possible intensity classes. The first sample was the training sample, including 24 lenses, each containing one defect (6 lenses $\times$ 4 intensities). The second sample was a test sample, including 20 lenses (different from the first sample) containing one defect (5 lenses $\times$ 4 intensities). In order to obtain optimal control of the type of defect, with the help of experts and technicians we ourselves created the defects on lenses (with a great level of realism).

We used a three-step procedure: the training session was immediately followed by a test session 1; and a delayed test session 2 took place after 48 h. In the yardstick group ($N = 15$); inspectors were trained to systematically use the external yardstick tool to compare the defect of the lenses with the grades of intensity indicated on the yardstick. Feedback was immediately provided for each lens. In the memorisation group ($N = 15$); inspectors were first shown visually the four grades of intensity using the yardstick. Immediately after this first explicit learning of the intensity grades they were told to answer "intuitively" by memory during the rest of the training session. Feedback was also immediately given for each lens. Presentation of the lenses was randomised. Response times and answers were recorded for test session 1 as well as for test session 2. Results of the good answers for the two test phases are indicated in Fig. 9.

Although the memory group performed 70% and the yardstick group 65%, the difference was not significant ($F(1,28) = 2.65$, $p = 0.11$). However, significant effects were found for the test session (test session 1 < test session 2; $F(1,28) = 17.42$, $p < 0.001$); and for intensity ($F(3,84) = 54.03$, $p < 0.001$), in favour of the extreme classes (i1 and i4). Significant interaction between the group and test session ($F(1,28) = 7.16$, $p < 0.05$) showed that the memory group was better than the yardstick group; however, this effect concerned only the first test session ($F(1,28) = 15.13$, $p < 0.001$). Marginal interaction between intensity and group ($F(3,84) = 2.47$, $p = .067$) suggested the memory group judged slightly better than the yardstick group i1 and i2 intensities. To summarise, this experiment revealed that the access to precise representations of the boundaries of levels of intensity might have a considerable effect on a correct acceptance decision in inspections.

7.3. Integrated training sessions

The last part of the programme was aimed at integrating the two first steps of the learning phase in a new and real inspection

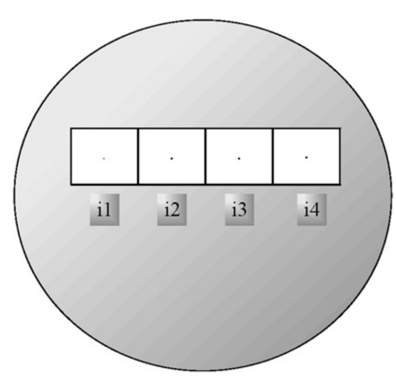


Fig. 8. A yardstick used to learn intensity.

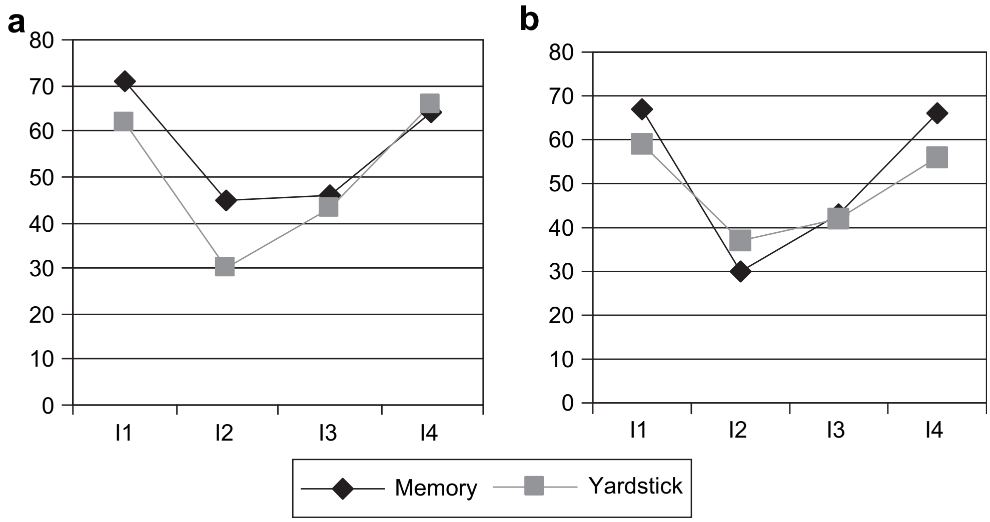


Fig. 9. Learning defect intensity. (a) Good answers (%) for the evaluation of four intensities in memory and yardstick condition test session 1; (b) good answers (%) for the evaluation of four intensities in memory and yardstick condition test session 2 (after 48 h).

training session. Design of the training schedule was based on the two previous studies and on the recommendations of Czaja and Drury (1981). Thus four design principles were used: (1) presentation of an enormous variety of defects on the basis of prototypes (and non-prototypes) for each defect; and of bad as well as good lenses; (2) continuous progression in learning sessions regarding the complexity of the successive samples of defects; (3) active participation of inspectors engaged in the training sessions in real workstation exercises; (4) systematic immediate feed-back on answers about detection as well as decision.

Fifteen new inspectors (all volunteers, and different from those taking part in the two previous studies) without control experience, but all experienced employees of the factory in other posts, were trained in order to become full-time inspectors. The training period was staggered over 12 months, with regular learning sessions, inside the factory. Training periods included two main phases. The first phase was about the acquisition of the basic rules of inspection (including detection and decision). This phase (20 h) was mainly declarative (Anderson, 1983). It was followed by the training session on real workplace phase. This second phase was mainly procedural, including progressively real specific ranges of lenses and real time constraints. This second phase was run by quality inspectors and took place alternately with a progressive hold on of real work productions stations. At the beginning, the second phase also included paper-based exercises with paper drawings of lenses with their defects (categories, intensities, zones, and numbers of defects).

During training sessions, ten different ranges of lenses with, and without, defects were used. The samples of lenses were progressive in terms of acceptance difficulty and type of defect. Each range of lenses was composed of a series of 28 lenses. Also, in the first ranges of lenses, the same categories of defects were grouped together; in contrast, in the last ranges they were randomly mixed. At the outset of training, defects were cued on the lenses, later this cueing disappeared. Ranges of lenses could be used several times until each participant performed a high proportion of good answers (more than 90%).

The training programme was subjected to quantitative assessment at different stages of the training session. We designed this assessment as an experimental investigation. Thus, in experimental conditions, each trainee had to inspect a series of new lenses, containing defects (or no defects) similar to those in real workstations. This experimental assessment was carried out four times during the training programme. The trainees' performance was compared to the performance of a group of experts (with a minimum of 10 years of experience as inspectors). Participants were fifteen trainees and seven experts. The material of the experimental assessment was composed of four similar samples of 26 lenses (14 lenses acceptable – 7 Good Without Defect, GWD, and 7 Good Acceptable Defect, GAD, 12 lenses non-acceptable, Bad Non-acceptable Defect, BND). Each trainee was provided with a sample of 26 lenses to be inspected four times: (1) in the middle of the first training phase (after the phase including the basic and declarative part of the programme of 20 h duration) and after learners have already been trained on the first fifth ranges of training lenses; (2) at the end of this first training period; (3) after two months of alternating between training sessions and production sessions at real workplace; (4) and after six months of alternating between training sessions and the real workplace.

For each lens of each sample, trainees had to perform a complete inspection. In contrast to training sessions, no feedback was provided during the test sessions. In order to obtain information about detection processes as well as decision processes, for each inspected lens, trainees had to give two answers (as quickly as possible): the first in respect of detection (defect or non- defect) and following this as soon as possible the second in respect of decision (acceptance or not). Response times and answers (good answers as well as errors, false alarms and omissions) were recorded.

Main results about good answers in detection and decision are indicated in Fig. 10. Regarding detection, ANOVAs were carried out on good answers for good lenses and on good answers for lenses containing a defect. For Good lenses we found an effect of test sessions revealing an effect of training ($F(3,24) = 4.7$, $p < 0.01$); in the last test session, there was no difference between trainees and experts ($F < 1$, ns). For lenses containing a defect, we also found an effect of test sessions showing the powerful influence of training. However, at the fourth test session, experts still performed better than trainees ($F(1,14) = 6.01$, $p < 0.03$). It seems that after six month out of twelve, additional training time is still required for optimal performance in defect recognition. Regarding decision, ANOVAs were carried out on good answers and on correct rejects. For good answers we found an effect of test sessions indicating a strong influence of training ($F(3,24) = 9.60$, $p < 0.001$); and in the last test

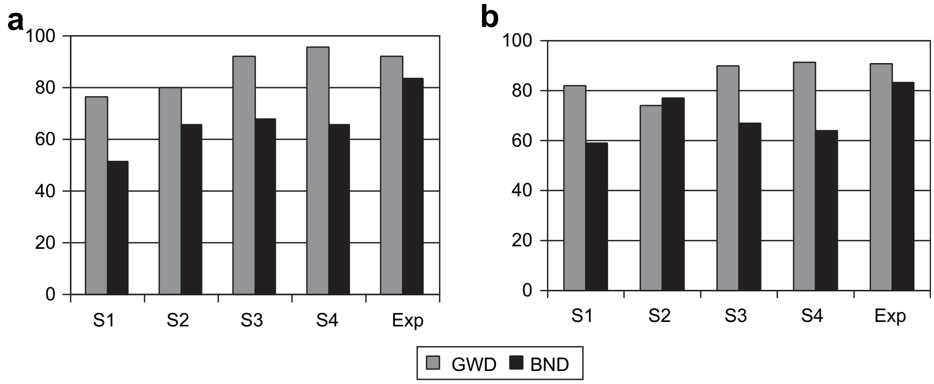


Fig. 10. Real workplace training. (a) Detection training – good answers (%), for each evaluation session of trainees training compared to experts; (b) Decision training – good answers and correct rejects (%) of lenses with defects for each evaluation session of trainees training compared to experts.

session there was no difference between trainees and experts ($F < 1$, ns). For correct rejects, we found an effect of test sessions ($F(2,28) = 5.32$, $p < 0.02$). However, univariate comparisons showed a significant progression only between the first test session and the others ($F(1,8) = 9.85$, $p < 0.02$). At the fourth session experts performed even better than trainees ($F(1,14) = 8.33$, $p < 0.02$). Thus, for decision it seems that after six month out of twelve, additional training is needed for optimal performance in rejection.

At the end, after 12 months of training alternating with real production activities in the workplace, the randomised samples taken (for statistics) by quality inspectors on production lines showed a mean percent of errors of 1.69 for experts and of 2.19 for the trainees.

Finally, we will conclude this article signalling some limitation to our case study of expertise in quality inspection which needs to be underlined. In particular, the experimental aspect of our investigation affects our results, notably by limiting the time constraints that are usually applied in real work conditions. Results regarding performance might have been different using the customary time constraints. However, let us make the hypothesis that the deep nature of the detection and decision cognitive processes has not been changed but only reduced in speed by limiting time constraints. Perhaps inspectors implicitly adopted their usual speed regime in achieving our task.

Acknowledgments

This research was funded by grants to M. Rebsamen from ESSILOR Development and Research Centre and from the Regional Council of Burgundy (Grant 2-001-21).

References

Alwood, C.M., 1986. Novices on a computer: a review of the literature. International Journal of Man-Machine Studies 25, 633–658.
Anderson, J.R., 1983. The Architecture of Cognition. Harvard University Press, Cambridge.
Biederman, I., Shiffar, M.M., 1987. Sexing day-old chicks: a case study and expert system analysis of a difficult perceptual-learning task. Journal of Experimental Psychology: Learning, Memory and Cognition 13 (4), 640–645.
Boucheix, J.M., 2003. Ergonomie et formation. Ergonomie cognitive des apprentissages en formation professionnelle. Psychologie Française 48 (2), 17–34.
Boucheix, J.M., 2004. Simulation and comprehension aid tool for complex technical documents. Advanced Learning Technology, ICALT, I.E.E.E., 898–900.
Cellier, J.M., Eyrolle, H., Mariné, C., 1997. Expertise in dynamics environments. Ergonomics 40 (1), 28–50.
Czaja, S.J., Drury, C.G., 1981. Training programmes for inspection. Human Factors 23 (4), 473–484.
Drury, C.G., 1975. Inspection of sheet materials – model and data. Human Factors 17 (3), 257–265.
Drury, C.G., 1992. Inspection performance. In: Salvendy, G. (Ed.), Handbook of Industrial Engineering, second ed. John Wiley & Sons, Inc., pp. 2283–2314.
Drury, C.G., Chi, C.F., 1995. A test of economic models of stopping policy in visual search. IIE Transactions 27, 382–393.
Drury, C.G., Fox, J.G., (Eds). (1975). The imperfect inspector (Introductory paper). *Human reliabilityb in quality control*. Taylor & Francis Ltd.London.
Drury, C.G., Sinclair, A.M., 1983. Human and machine performance in an inspection task. Human Factors 25 (4), 391–399.
Duchowski, A.T., Shivashanskaraiah, V., Rawls, T., Gramopadhye, A.K, Melloy, B., Kanki, B., November 6–8, 2000. Binocular eye tracking in virtual reality for inspection training. In: Proceedings of the Eye Tracking Research and Applications Symposium, Palm Beach Gardens, FL, pp. 89–96.
Ericsson, K.A., Lehaman, A.C., 1986. Expert and exceptional performance: evidence of maximal adaptation to task constraints. Annual Review of Psychology 47, 273–305.
Farrington-Darbi, T., Wilson, J.R., 2006. The nature of expertise: a review. Applied Ergonomics 37, 17–32.
François, M., 1989. Ergonomie des postes de contrôle-qualité. Cahiers de Notes Documentaires INRS 137 (1753), 595–606.
Gallwey, T.J., Drury, C.G., 1986. Task complexity in visual inspection. Human Factors 28 (5), 595–606.
Glaser, R., Chi, M.T.H., 1988. Overview. In: Chi, M.T.H., Glaser, R., Farr, M. (Eds.), The Nature of Expertise. N.J. Erlbaum, Hillsdale, pp. 6–21.
Gramopadhye, A.K., Drury, C.G., Sharit, J., 1997a. Feedback strategies for visual search in airframe structural inspection. International Journal of Industrial Ergonomics 19 (5), 333–344.
Gramopadhye, A.K., Drury, C.G., Prabhu, P.V., 1997b. Training for visual inspection. Human Factors and Ergonomics in Manufacturing 7 (3), 171–196.
Gramopadhye, A.K., Bhagawat, S., Kimbler, D., Greenstein, J., 1998. The use of advanced technology for visual inspection training. Applied Ergonomics 29 (5), 361–375.
Gramopadhye, A.K, Melloy, B.J., Chen, S., Bingham, J., August 2000. Use of computer based training for aircraft inspectors: findings and recommendations. In: Proceedings of HFES/IEA Annual Meeting, San Diego, CA.
Liévin, D., François, M., 1997. Intégration d'une tâche de contrôle dans l'activité d'opérateurs de fabrication: le cas d'un atelier d'électronique. Les notes scientifiques et techniques de l'INRS (NS0153).
Myles-Worsley, M., Johnson, A., Simons, M.A., 1988. The influence of expertise on X-Ray image processing. Journal of Experimental Psychology: Learning, Memory and Cognition 14 (3), 553–557.
Rosch, E., 1978. Principles of Categorization in Cognition and Categorization. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey.
Shanteau, J., 1992. The psychology of experts. An alternative view. In: Wright, G., Bolder, F. (Eds.), Expertise and Decision Support. Plenum, New York, pp. 11–23.
Spitz, G., Drury, C.G., 1978. Inspection of sheet materials, test of model predictions. Human Factors 20 (5), 521–528.
Stephaneck, P., 1966. Problèmes ergonomiques du contrôle de qualité. In: Faverge, J.M., Lolivier, M., Delahaut, J., Stephaneck, P., Falmagne, J.C. (Eds.), L'ergonomie des processus industriels. Université libre de Bruxelles, Edition de l'institut de sociologie.
Sylla, C., Drury, C.G., 1995. Signal detection for human error correction in quality control. Computers in Industry 26, 147–159.
Tanner, W.P., Swets, J.A., 1954. A decision-making theory of visual detection. Psychological Review 61 (6), 401–409.
Treisman, A.M., Gelade, G., 1980. A feature integration theory of attention. Cognitive Psychology 12, 97–136.
Vicente, K.J., Wang, J.H., 1998. An ecological theory of expertise effects in memory recall. Psychological Review 105, 33–57.
Vora, J., Nair, S., Gramopadhye, A.K., Duchowski, A.T., Melloy, B.J., Kanki, B., 2002. Using virtual reality technology for aircraft visual inspection training: presence and comparison studies. Applied Ergonomics 33 (2), 559–570.