2013

Visual Inspection of Casting Surfaces

Frank E. Peters  
*Iowa State University*, fpeters@iastate.edu

Richard T. Stone  
*Iowa State University*, rstone@iastate.edu

Kristopher Patrick Watts  
*Iowa State University*, watts.kristopher@gmail.com

Peihan Zhong  
*Iowa State University*, pei@iastate.edu

Alexander John Clemons  
*Iowa State University*, ajclemons@live.com

Follow this and additional works at: [http://lib.dr.iastate.edu/imse_pubs](http://lib.dr.iastate.edu/imse_pubs)  
Part of the [Industrial Engineering Commons](http://lib.dr.iastate.edu/imse_pubs), and the [Systems Engineering Commons](http://lib.dr.iastate.edu/imse_pubs)

The complete bibliographic information for this item can be found at [http://lib.dr.iastate.edu/imse_pubs/44](http://lib.dr.iastate.edu/imse_pubs/44). For information on how to cite this item, please visit [http://lib.dr.iastate.edu/howtocite.html](http://lib.dr.iastate.edu/howtocite.html).
Visual Inspection of Casting Surfaces

Abstract
Visual inspection of casting surfaces is a critical processing step in the metalcasting process. Measurement error can be significant, causing overprocessing or poor quality received by the customer. Many factors affect the human operator’s ability to effectively inspect castings. Cognitive ability is one such factor, which can be identified through a simple test. Rastering training was shown to improve the percentage of casting surface which gets inspected. Discrimination between acceptable and unacceptable surfaces continues to be problematic.

Disciplines
Industrial Engineering | Systems Engineering

Comments
This article is from *AFS Transactions* 121 (2013): 45. Posted with permission.
Visual Inspection of Casting Surfaces

F. Peters, R. Stone, K. Watts, P. Zhon, A. Clemons
Iowa State University, Ames, Iowa

Copyright 2013 American Foundry Society

ABSTRACT

Visual inspection of casting surfaces is a critical processing step in the metalcasting process. Measurement error can be significant, causing overprocessing or poor quality received by the customer. Many factors affect the human operator’s ability to effectively inspect castings. Cognitive ability is one such factor, which can be identified through a simple test. Rastering training was shown to improve the percentage of casting surface which gets inspected. Discrimination between acceptable and unacceptable surfaces continues to be problematic.

INTRODUCTION

The visual evaluation of casting surfaces is conducted several times during the production of metalcastings. The most obvious persons conducting these visual evaluations are the inspectors, who markup unacceptable areas for remediation or justification to be scrapped. However, evaluations are also done by operators while they are processing the castings, by management, by sales, and ultimately by the customer. Previous work has shown that there is significant measurement error, both repeatability and reproducibility. Undetected surface defects will result in additional rework cycles or worse, returns from the customer. Marking acceptable anomalies as defects will result in excessive processing. Operators must be provided with appropriate training and inspection aids to help reduce the measurement error.

KNOWN FACTORS THAT EFFECT VISUAL INSPECTION

The literature provides information on the factors that affect visual inspection performance which will be reviewed for the application to metalcasting tasks.

ENVIRONMENTAL FACTORS

The environment can play a major role in the outcome of visual inspection tasks. This is important because before assumptions can be made on individual or task factors, it must be first determined if the environment itself represents a confounding variable to our findings.

Temperature is an obvious but often overlooked factor in visual inspection environments. Temperature and humidity together affect the ability to achieve optimal thermal exchange with the environment, too hot and the body will focus on cooling, too cold and it must focus on the conservation of heat. Both of these actions affect the human’s ability to achieve optimal cognitive performance. It has been found that cooling sensations (e.g. from a fan) activate the brain and excite the nervous system controlling thermoregulation, and that this activation of the sympathetic nervous system elevates mental alertness and therefore increasing general attention capabilities. Interestingly, a slight (<1 degree) increase in body temperature can be an indication of heightened awareness and performance. This increased body temperature may be a result of the extra alertness of the operator. A higher environment temperature is not needed to achieve this result. Extremes in a work environment cause issues in the perception of moderate exertion and cooling can increase an inspector’s long term capabilities. Ideally, it is suggested to utilize work stations that have localized cooling sources (e.g. a fan or personal cooling unit) to keep an environment near 70°F (21C) with a relative humidity between 60 and 70%.

OPERATOR FACTORS

Many individual factors have been shown to be of particular importance in the area of visual inspection. Some of these factors are obvious and include visual acuity, color vision, and depth perceptions. As an eye becomes less accurate so will its ability to accurately interpret defects. Experience on the job can be a very positive individual factor. In fact, inspectors who are experienced in performing a specific visual inspection task can concentrate longer and perform with a higher degree of accuracy than those with less experience. Most of these benefits, however, were found in search tasks that have a controlled amount of variation, unlike typical casting surface inspection tasks.

General intellectual aptitude is another known factor in determining the capabilities of a visual inspector. As one would expect, higher intellectual aptitude correlated positively with higher levels of visual inspection success. This appears to be due to an increased ability to adapt strategies and utilize memory techniques that aid in the visual inspection process. The cognitive inspection style of individuals also appears to be a factor in visual inspection capability. Specifically, the Matching Familiar Figures Test (MFFT) has been shown to have potential in determining individual potential in performing visual inspection tasks. The use of this test will be discussed subsequently.

TASK FACTORS

The task itself can greatly affect an operator’s ability to perform a visual inspection. Depending on the kind of
visual inspection being performed the posture and physical strain can impact visual inspection capabilities. Several authors have shown that select visual inspection jobs suffer from decline in visual inspection performance as a result of physical strain.\textsuperscript{4,9,10} Studies like this have shown that as time on task increases, participants tended to lean forward, change postures more frequently, report more discomfort and take more time performing visual inspection. However, this increased time does not lead to increased accuracy, rather, error rates go up. Issues related to postural discomfort, physical fatigue and biomechanical strain resulting in visual inspection decrements can largely be quantified using known evaluation methods. These factors could be significant given the typical components being inspected and the work areas of metalcasting inspection operations.

Another task factor that has been shown to impact the performance capability of visual inspectors is the level of documentation provided for on the job comparisons. It is much easier to perform a relative judgment than it is to make an absolute one. It can be very beneficial to provide aids, such as comparator plates that allow visual inspectors to compare a possible flaw to a known category of flaw. Several studies have shown that the visual inspectors that are given an effective comparison aid consistently outperformed those that were forced to make absolute judgments alone.\textsuperscript{6,11} Therefore, simply giving the visual inspector a visual reference aid, such as a comparator plate, can greatly increase their basic capabilities.

**EFFECT OF COGNITIVE STYLE**

Several studies have been proposed to classify and select inspectors for visual inspection.\textsuperscript{7,8,12-15} Schwabish and Drury designed an experiment to evaluate the influence of the reflective-impulsive cognitive style on visual inspection.\textsuperscript{7} They utilized the Matching Familiar Figures Test (MFFT), which asks the participant to choose from six figures the one that is the same as the one on top of the screen. The time until the correct match is made and how many errors are made before getting the correct match are the dependent variables that are recorded. Figure 1 is an example used in the MFFT. According to the accuracy and response time of subjects taking the MFFT, subjects were classified into four different cognitive styles: fast-accurates (spend shorter times, make fewer errors); reflectives (spend longer times, make fewer errors); impulsives (spend shorter times, make more errors) and slow-inaccurates (spend longer times, make more errors). Their results indicated that there was a grouping phenomenon based on the accuracy dimension. The accurate group (i.e., reflective and fast-accurate) was faster than the inaccurate group (i.e., impulsive and slow-accurate). Research in this field indicates that those classified as ‘fast accurates’ will have the greatest inspection capability followed by ‘reflectives.’

‘Impulsives’ and lastly those classified as ‘slow inaccurate.’

Given the success of Schwabish and Drury in using MFFT for evaluating the potential of individuals to be aircraft inspectors, it appears that MFFT could also be effective in the metalcasting industry. To evaluate the effectiveness of the MFFT test for casting operations, data was collected on people inside and outside of the metalcasting industry. Three groups of people have been evaluated using MFFT for the current work. They include undergraduate engineering students, members of the United Association of Plumbers and Pipefitters Union undergoing initial welding certification and workers in the metalcasting field that have visual inspection as part of their work tasks. The metalcasting personnel were from three different companies. The groups are referred to as the student, pipefitter and casting groups, respectively in this paper. All of the participants were given the MFFT as well as a computerized visual inspection task. The MFFT data set included 58 students, 16 pipefitters and 33 casting personnel. Since the population groups were not equal, the analysis randomly selected an equal number (12) from all three groups. This was repeated five times, and the average values are reported in Table 1. For the time to complete the MFFT tasks, the person was determined to be ‘fast’ or ‘slow’ based on whether they completed the tasks in less than or more than the median time for the group of thirty six subjects. Similarly, they were labeled ‘accurate’ or ‘inaccurate’ based on their score relative to the median score for accuracy. Of the students, 22% were labeled as fast accurate while 17% of the pipefitters and 9% of the casting personnel achieved this classification. For the slow inaccurate category, the students and pipefitters each had 3% in this category while 39% of the casting personnel fell in this category based on their MFFT performance. While the sampling to create equal populations eliminates the bias of having unequal population sizes, these three groups of people are still not representative of the general population.

Regardless, the results do provide insight on the potential of these populations to do the cognitive tasks needed for visual inspection and the need to make them more effective.
Table 1. MFFT Results by Population

<table>
<thead>
<tr>
<th>Population Group</th>
<th>Fast Accurate</th>
<th>Reflective (Slow Accurate)</th>
<th>Impulsive (Fast Inaccurate)</th>
<th>Slow Inaccurate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students</td>
<td>22</td>
<td>31</td>
<td>43</td>
<td>3</td>
</tr>
<tr>
<td>Pipefitters</td>
<td>17</td>
<td>47</td>
<td>33</td>
<td>3</td>
</tr>
<tr>
<td>Metalcasting</td>
<td>9</td>
<td>25</td>
<td>28</td>
<td>39</td>
</tr>
</tbody>
</table>

The participants also completed a computerized visual inspection task; an example of which can be seen in Fig. 2. In this task, the time and accuracy was recorded. When the cognitive style was considered with respect to this visual search task, the ‘fast accurate group’ had the best performance, in terms of speed and accuracy. There was a statistically significant difference in the speed of the task between the ‘fast accurate’ and both the ‘reflective’ and ‘impulsive groups.’ Interestingly, the ‘fast accurate’ did not have significantly different time compared with the ‘slow inaccurate’ subjects. This is attributed to the ‘slow inaccrances’ speeding up to just get done with the task. This conclusion is supported by the fact that in terms of accuracy of the decision-making, the ‘fast accurate’ were superior to the other three groups.

In conclusion, the MFFT was shown to be useful in determining the cognitive ability of personnel and there was a correlation between cognitive style and their performance on a visual search task. This test could be used as a screening tool for new operators, or to identify the need for additional training and inspection aids among existing and future operators. It should be noted that the underlying cognitive ability cannot be improved through training, but the strategies in which a person inspects a part can be improved to compensate for their lower abilities. The cognitive style was considered as a factor on the experiments involving casting inspection tasks described subsequently, but unfortunately, the sample sizes were not large enough to develop meaningful conclusions.

FIELD STUDY RESULTS

Field studies were conducted at both steel and iron foundries producing castings from 1 to 10,000 lb (.5 to 4,500 kg) from shell, investment, chemically bonded sand and green sand. The goal of these field studies was to identify factors influencing visual inspection that could be an opportunity for improvement. An eye tracker was developed so studies of the search strategies being utilized during the inspection process could be determined (Fig. 3). This tool allows analysis of the search pattern as well as determining the percent of the casting surface that is being inspected. It is mounted on a face shield and goggles to provide the eye protection. This is often used in the shop floor environment. The eye tracker has three cameras, two recording the world view in front of the operator and the other records the location of the pupil within the eye. Two cameras on the world view are necessary to collect data on operators that are inspecting larger castings. The video from the two world view cameras is ‘stitched’ together. After this, the world view and the pupil view are compiled together which resulted in a video of the world view with crosshairs to show where the user was looking at any given time. This compiled video can be analyzed, frame by frame, to determine the operator’s search patterns.

To understand the results of the eye tracker, the useful field of view (UFOV) needs to be defined. UFOV is the area where humans can see the most detail in their vision due to the greater concentrations of rods at that particular location. For the purpose of this study a reasonable value of 6 degrees is assumed. Assuming that the parts were 22 in. (56 cm) from the operator’s eye, the useful field was calculated to be 2.25 in. (5.7 cm) wide. It should be noted that the UFOV used in this study was generous and assumed the best case scenario for viewable area (this was done to avoid biasing results in favor of experimental solutions). This was used to determine the percent of casting surface which was being inspected during the field studies. Figure 4 shows an example of a casting with the search path line and the useful field of view overlaid on the part.

![Fig. 2 This is an example screen of the visual inspection task with the 'defect' circled.](image-url)
For data collected on castings up to approximately 15 in. (38 cm), the eye track revealed that on average, 59% of the casting surface was not being viewed sufficiently enough for the inspector to make meaningful decisions. Furthermore, the search patterns used were quite random since 71% of the time the search pattern was not consistent with the one used for the immediate preceding casting of the same geometry.

Figure 5 shows the results from another field study at a foundry producing large castings (greater than 1000 lb [450 kg]). The initial results from the eye tracker showed a random search pattern used to inspect the castings, which led to only 84% coverage of the casting surface. After some basic rastering training, where the operators were instructed to use a consistent left to right search pattern, the eye tracking results were again collected. This time, 96% of the casting surface was adequately covered (Fig. 6).

LABORATORY EXPERIMENTS

The field studies showed that operators were not using a consistent search pattern nor was there adequate coverage of the casting by the operator. To help alleviate these problems, laboratory studies were conducted on methods to improve the search pattern as well as to study the ability to discriminate between acceptable and unacceptable surfaces. It should be noted that the laboratory experiments, described subsequently, predominately utilized college students of various backgrounds. As such, the average age for each experiment was in the low 20s. The vision for each participant was tested and determined to have 20/20 vision, either naturally or corrected. All tests were scheduled when the participants were normally awake and at a time convenient to them. These conditions limited the impact on the UFOV of the participants on the laboratory results.

Fig. 3. The eye tracker is used to determine where the inspector was looking.

Fig. 4. An example of a casting with the search pattern and the UFOV highlighted is shown.

Fig. 5. This is a larger casting with the search pattern and UFOV indicated.

Fig. 6. This is a larger casting with the search pattern indicated after the operator received rastering training.

RASTERING TRAINING WITH OVERLAYS

In the first experiment, three conditions were tested related to the training and inspection aids provided to the subjects. The first group was only given instructions on what surface indications for which they were to look, but training as well as using a left to right search strategy was not instructed on how to inspect the castings (basic training group). The second group received the basic
called rastering (raster group), and they were instructed to raster the entire time. The third group was provided the same training as the raster group but was also provided green static lines projected onto the part with a projector (overlay group). There were 24 participants in this study.

Specifically, the raster training included instructions to first follow lines on a sheet of paper as if they were reading (left to right, top to bottom). They were encouraged to continue this process and increase their speed through multiple runs. The participants were encouraged to raster on an image, which had points of interest that the participants were to point out. Once this had been completed the participants repeated this procedure with an actual casting.

Once training had been completed, the participants were asked to inspect 90 parts from five differing casting types ranging from 2 to 55 lb (1 to 25 kg). The laboratory experiment was set up to closely resemble a casting visual inspection process. They would inspect the part for surface anomalies. Each anomaly that met the criteria was to be circled using a piece of chalk. Then, the part was placed back onto the conveyance system and the process repeated for each part.

The variables being measured in this experiment are the performance measures related to signal detection (d', hit rate and false alarm rate). Hits are the correct identification of a surface area as unacceptable. Conversely, a false alarm is when a participant designates an area to contain an unacceptable indication when it was actually acceptable. The surface acceptance criterion was set to ASTM A802 SCRATA plate standard at the B4 and C1 level. This was done independently by three individuals familiar with the standard. For each part, the number of hits and false alarms were determined for each subject.

An important metric to consider with signal detection theory is d'. This is a single measure that combines hits (correct rejections) and misses (incorrect acceptance) for a particular operator. Specifically, it is the ratio of how many hits to how many misses the participant had, therefore larger values are preferred.

The analysis of the variance showed that there was a significant difference in the d' values for the subjects with different training conditions. A plot of the d' score by training condition with box plots can be seen in Fig. 7. In order to determine which training condition was significant from another, a Tukey-Kramer HSD test was done. This showed that the raster training group is superior to the basic training group. However there is no statically significant difference between the overlay group and either other group.

For the false alarm rates, there was a significant difference between the training. A plot of the false alarm rate by training condition with box plots can be seen in Fig. 8. The basic training group was significantly different from both the overlay and the raster training group. The lack of search pattern training resulted in more false alarms.

The analysis of variance showed that there was not a significant difference in hit rates among the different training groups. The raster training group was superior to the basic training group with regards to overall signal detection (d'). However, despite the fact the overlay group received raster training they were not superior to the basic overlay group. Due to previous literature it would be expected that the overlay group would have the same level of performance (d') as the raster training group if not increased performance.16-17

The likely reason for the overlay group not performing superior to the others is that the overlay lines were distracting. They are very salient and caused the inspectors' eyes to be drawn to those areas and not to the surface areas to be inspected. This is a common effect with humans.18-23

The likely reason subjects in previous literature responded well to a visual overlay is that they were presented in a
different media. Most were inspecting on a computer screen where the lines were static and did not move. In this experiment, the static overlay lines would move relative to the casting. If they were holding the part in their hands during the inspection, the lines could fluctuate and move. Therefore, the alignment issues may have affected an individual’s ability to effectively adapt to edge prediction (a factor in the development of systematic search pattern training). This would lead to slow downs and eventual skipping of raster lines. Another factor to consider was that the participants would have to utilize more mental effort to confirm line existence, therefore slowing the process. This time loss had to be compensated by reducing peripheral inspection activities. These conclusions were confirmed with the results from the eye tracker used on these subjects.

The raster training group had the highest percentage of the part viewed and were able to make the best decision about what an indication was, leading to the higher d’ values. The overlay group had the second highest d’ values, however, the distraction that the lines provided likely caused their performance to drop. Finally, the basic training group had the lowest training percentage of the part viewed. The participants would view only a certain percentage of the part before making a decision, which would generally be a bad one. However, it was found that both groups with rastering training did adopt a more optimal (rastering) search pattern as opposed to the basic training group, whose search patterns were more random. The systematic search pattern would have likely helped with both rastering groups, if not for the distracting effect of the overlay.

VIDEO RASTERING TRAINING
Since it proved to be impractical to have an overlay projected onto the casting surface, another experiment was conducted to investigate methods of improving rastering. A video training was created to train subjects to move their eye in a systematic search pattern by following a prompt on a computer screen, at first across a blank screen and then over casting images. The field studies showed a very strong tendency for operators’ view to reach the edge of the casting and then follow the edge and miss the center area of the casting. Therefore, the video training specifically had the search path run over the edges of the castings.

Three conditions were tested: a control group with basic training on what they were supposed to inspect; those that had the basic training and the video training. One group was instructed on a left-right, left-right raster pattern and the other had left-right-left-right. The former pattern is the same as used when reading, the latter was included because it would be more efficient since the eye did not need to move back to the left edge without an inspection task. After the video training, the eighteen subjects were asked to inspect about 100 castings and look for surface areas which were not acceptable and then mark those areas, similar to the previous experiment.

This research shows that the video rastering training increased the percentage of unacceptable areas, which were correctly identified or an increased ‘hit rate’. The training also caused more areas, which were acceptable, to be marked as unacceptable, or increased ‘false alarm’ rate. This is logical in since the operators are getting the subjects to systematically look at more of the surface. These results in more ‘hits’ but also more ‘false alarms’. Subjects are ‘seeing’ more of the surface, but have the same difficulty in discriminating between acceptable/unacceptable surfaces. The rastering video training will be packaged onto a website for use by industry.

DISCRIMINATION EXPERIMENT
The purpose of this experiment was to test the ability of subjects to discriminate between acceptable and unacceptable surfaces. The participants were asked to make a binary decision, acceptable or unacceptable, for twenty five casting surface samples. The casting surface replicas were made to avoid the distraction of handling castings and searching the whole casting for areas of interest. Using rubber mold making and epoxy model materials, actual casting surfaces were replicated in hard grey colored epoxy samples that were four inch squares. These samples had the texture, color and finish very similar to ferrous casting surfaces.

There were forty participants, of these, thirty-two returned after a month for the second session. At the time of the first session, the participants were not aware that they would be invited back for the second session.

Three comparator plates with surface anomalies, similar in type to the test surfaces, were used in this experiment. The participants were instructed to consider the comparator plate surfaces to be the threshold between acceptable and unacceptable. During each session, the participants conducted the experiment twice, once with and once without the comparator plate. In the first session, it was randomly chosen whether they had the comparator plates for the first or second trial. For the second session, the first inspection task was without the plates to test their recall on what was acceptable or not.

The hit rate was calculated as the percentage of correct rejections over total number of defects. The false alarm rate was calculated as the percentage of false rejections over total number of acceptable samples.

Before each test, training and a practice trial was provided. In the training, the standard for acceptance was illustrated by three comparator plates, followed by a practice trial with 5 casting samples.
The results for hit and false alarm are presented in Table 2. While there is an increase in hit rate when the comparator plates were used in the second session, given the sample size the results are not significant. While anecdotal evidence from industry indicates that comparator plates are useful, unfortunately, it could not be proven from this experiment. It should be noted that even for this most basic visual inspection task, a subject’s ability to correctly reject a sample was at best 78% and the lowest false alarm rate was 25%.

REPRODUCIBILITY
The discrimination experiment was also used to extract information on reproducibility or the agreement between inspectors. As described previously, each of forty participants made a binary acceptable/unacceptable decision for each of twenty five samples. The range of agreement among subjects for the samples was 58 to 100%. The average agreement was 80% and the median value was 78%. It should be noted that this was a very easy inspection task compared with the typical casting surface inspection. The specimen was a square sample with essentially one area of interest centered on it. Despite this, the subjects could only agree among themselves 80% of the time.

Table 2. Hit and False Alarm Rates

<table>
<thead>
<tr>
<th>Condition</th>
<th>Hit Rate</th>
<th>False Alarm Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Session – With Comparator Plate</td>
<td>64%</td>
<td>25%</td>
</tr>
<tr>
<td>1st Session – Without Comparator Plate</td>
<td>66%</td>
<td>29%</td>
</tr>
<tr>
<td>2nd Session – With Comparator Plate</td>
<td>78%</td>
<td>30%</td>
</tr>
<tr>
<td>2nd Session – Without Comparator Plate</td>
<td>67%</td>
<td>26%</td>
</tr>
</tbody>
</table>

CONCLUSION
There are many considerations for improving the visual inspection process. The Matching Familiar Figures Test is effective in identifying the cognitive ability of the inspector and cognitive ability is correlated to performance on a visual inspection task. MFIT can be used to help identify new inspectors and also to recognize the training needs for those of lower cognitive ability. Field studies showed that large amounts of the casting surface are not being adequately seen by the operators for them to make a useful judgment. Experiments that provide an overlay on the casting to encourage the operators to inspect the entire part proved to be distracting and therefore ineffective. Random training did prove to be effective and can be easily implemented to help train operators have a systematic and thorough search pattern. A person’s ability to discriminate between acceptable and unacceptable remains problematic and requires diligence to ensure that inspectors continue to make appropriate decisions.

ACKNOWLEDGMENTS
Research was sponsored by the US Army Contracting Command Joint Munitions and Lethality Contracting Center, Picatinny, NJ on behalf of Benet Laboratories, and was accomplished under Cooperative Agreement Number W15QKN-10-2-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of Benet Laboratories or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

REFERENCES


