Misconceptions about NDT Workmanship Acceptance Criteria for Quality Control

Edward A. GINZEL

Materials Research Institute, Waterloo, Ontario Canada
e-mail: eginzel@mri.on.ca

Abstract
One of the final steps in any NDT inspection process is the disposition of the tested piece. The NDT operator compares any indications noted to the set of rules provided in the test procedure or applicable standard and is expected to pronounce judgment on the component. Suggesting that “we must make sure that harmful defects are detected and removed as part of a workmanship standard”, sounds like Orwellian Doublespeak. NDT operators are mostly of the opinion that the rules they are using are there to ensure the safe operation of the component and any misjudgement of the length or amplitude response of the indication could be critical to the serviceability of the part. Only in special cases where NDT is done in accordance with a qualified Procedure and when the acceptance criteria are calculated in accordance with fitness-for-purpose is this true. For the most part, acceptance criteria typically used in NDT are used for Quality Control purposes. This paper considers some of the factors that illustrate the disconnect between the perceived function of the workmanship criteria and the actual function of these acceptance criteria.

Keywords: NDT, Quality Control, acceptance criteria, fitness-for-purpose, workmanship

1. Introduction

In 1959 Robert McMaster [1] looked back over the previous two to three decades to note how non-destructive testing (NDT) had grown from a laboratory curiosity to an indispensable tool. From our vantage point in the early 21st century, it is difficult to remember NDT as a “laboratory curiosity”. Although there are now new methods being developed in the laboratories, the general concepts and purposes of NDT identified in the mid 20th century are still relevant.

NDT broadly speaking provides three functions;

1. Assurance of product reliability
2. Promoting safety by reducing failure rates
3. Providing a means of increased profit

It is fairly obvious that the first two points are related. In a production environment, where thousands of items are made, it may be expected that not all items were made the same. The material or the fabrication process may have been inconsistent and resulted in a characteristic that may cause the item not to function in the way it was intended. Alternatively, it may function as intended but not for as long as might be considered reasonable. In the second case, if the product fails while in service, the result could be inconvenient. In other situations a failure could be dangerous. For example, a failed diode in an automotive alternator will result in the vehicle not starting. For the most part, this is an annoying inconvenience to the driver. However, a corroded tube used as the brake-line in the same vehicle could fail when the driver pushes hard on the brake pedal and the vehicle fail to stop. Not only is that incident inconvenient, it could be dangerous.

How NDT provides a means of increased profit has not always been as obvious as the reliability and safety function. There is an obvious connection with profit and the reliability aspect. Customers that use reliable products are satisfied customers and they will recommend the product to others. NDT also improves profit by controlling the material quality.
Blemishes in the raw material that could get exposed during machining and render the product useless are best discovered prior to the added cost of machining. NDT also provides a potential for cost saving when its findings can be incorporated into design. Many of today’s aircraft are made with thinner materials than they might have been in the past because NDT inspection has increased the confidence in the design. Often, structures are over-designed in order to allow for uncertainties in material properties. An obvious advantage is had in aircraft designs when the same safe performance can be assured with a slightly thinner material. It can mean lower costs in material and potential for increased payloads or reduced fuel consumption.

NDT also provides a mechanism to regulate the processes in manufacturing. Typically a purchaser of a product will have a specification for certain parameters that they consider important for the service of their product. Hardness limits, dimensional limits, concentricity of tubular products, or even alloy content are examples of parameters that NDT might monitor. The quality control process may be applied to all parts or it may be applied to just a few in each batch to ensure that any checks carried out by the purchaser will not find products outside of the required specification. Failure to conform to the specification could result in the order being returned and hence a significant cost to the producer.

2. Acceptance Criteria

2.1 Workmanship versus Fitness-for-Service

NDT has been used as a Quality Control tool since NDT was recognised as a part of industry in the mid 20th century. As a Quality Control tool, NDT is aimed at checking if an acceptable workmanship level is met or not.

Forli [2] points out, that if the workmanship level is not met there is a chance, in the long run, to have unacceptable defects introduced into the object(s) in question.

This is quite a bit different than saying that all unacceptable indications identified in a workmanship quality control inspection are critical to safety and functionality. Yet that is what seems to be the most common interpretation of Quality Control (workmanship) acceptance criteria.

Even in 1959, McMaster acknowledged that a “practical quality level” should be considered. No product is “perfect” and total perfection can never be achieved. Setting practical quality levels is needed to ensure that all relevant concerns are addressed whether they are safety, reliability or profitability.

When used simply as a quality control tool, the result of failing to meet the quality requirements generally leads to corrective actions. This could mean that, in production runs, parts are scrapped. In a fabrication environment this corrective action is often just a punishment that forces repair or replacement.

More recently it has been recognised that “punishment” is not always an effective or appropriate tool. In some cases the repair process can be more damaging than leaving the flaw in place. Repairs could result in the original material properties being reduced or a new flaw could be introduced that is subsequently un-detected. Recognition of these concerns has
led to the development of an alternative approach to workmanship acceptance criteria; fitness-for-purpose acceptance criteria.

Welding fabrication codes sometimes specify maximum tolerable flaw sizes and minimum tolerable Charpy energy, based on good workmanship, i.e. what can reasonably be expected within normal working practices [3]. Arguably these factors are somewhat arbitrary. If welding does not achieve them it does not necessarily mean that the structure is at risk of failure. When sufficient information is known about the part an Engineering Critical Assessment (ECA) can be carried out. ECAs is an analysis, based on fracture mechanics principles, of whether or not a flaw of known size and orientation can be safely tolerated under conditions of brittle fracture, fatigue, creep or plastic collapse under specified loading conditions. The ECA concept is also termed fitness-for-purpose analysis, and provides the alternative to simple punishment for not achieving quality requirements.

The concepts of fitness-for-purpose (FFP), now popular in engineering, have placed greater demands on NDT to provide quantitative results. Acceptance criteria presented to the ultrasonic operator are sometimes difficult to comply with as they seem not to consider the limitations and tolerances intrinsic in the techniques used [4]. When FFP concepts are used the traditional approach to NDT is not adequate. The NDT technique needs to demonstrate not only the ability to reliably detect potentially critical flaws, it usually needs to also demonstrate its ability to size and position the flaws within the component.

Under some situations, failure to meet the workmanship standards could lead to a fitness-for-purpose assessment and this could avoid damaging repairs to be done.

However, the added costs of a qualified NDT technique with the commensurate probability of detection (PoD) and sizing requirements added to the extra costs of an ECA are usually not justified and the fabricator is left with the punishment of replacement or repair.

2.2 Workmanship

NDT quality levels have been developed and are now seen in essentially all standards relating to NDT. However, contrary to common belief, they are empirically based and have no quantitative foundation in engineering calculations as would be the case for FFP acceptance criteria.

When one hears an experienced NDT user suggesting that “we must make sure that harmful defects are detected and removed as part of a workmanship standard”, it sounds like Orwellian Doublespeak. Such an attitude illustrates that the concept of workmanship acceptance criteria is not understood. Indeed, some flaws (large or small) may be considered harmful under some conditions. However, any determination of what constitutes “harmful” places the judgement in the hands of fracture mechanics engineers.

Instead, Workmanship rules are supposed to be used to indicate if good fabrication practices, e.g. the welding procedure specification (WPS), have been followed. Using NDT observations we might be able to aid the quality control department to identify problems in the process, such as; poor arc-control, inadequate grinding or back-gouging, improper material storage or preheat, etc. However, amplitude and length of an “indication” is not adequate information to decide if the structure is fit for service.
Quality levels are generally expressed by defect severity parameters such as size, type and location. As an example, for welded joints these concepts are standardised in ISO 5817 [5] and the quality levels in that document provide nothing more than categories of sizes of the flaws typically found in welding. ISO 5817 specifically notes “This International Standard is directly applicable to visual testing of welds and does not include details of recommended methods of detection or sizing by non-destructive means.”

The choice of quality level to be used for a particular application is to be made by the application standard or the designer in conjunction with the manufacturer, user and/or other parties concerned. ISO 5817 recommends that “The choice of quality level for any application should take account of design considerations, subsequent processing (e.g. surfacing), mode of stressing (e.g. static, dynamic), service conditions (e.g. temperature, environment) and consequences of failure. Economic factors are also important and should include not only the cost of welding but also of inspection, test and repair.”

Since there are no limitations noted as to the components tested (e.g. the rules could as easily apply to a welded sign support in a parking-lot as to a critical component in an electric generating station), the Levels in ISO 5817 appear to be both arbitrary and severe. Efforts to apply NDT methods other than visual inspection would be hard-pressed to meet all aspects in that standard so one is left with trying to interpret the statement made on the connection to other NDT methods. ISO 17635 [6] is then the link between ISO 5817 and the NDT standards to be applied depending on the product to be tested.

When we examine the specifics of some of the NDT Standards that incorporate workmanship acceptance criteria, we can see no rationale to connect the criteria to engineering concerns such as might be seen in FFP derived criteria. Instead, we are essentially stuck with an arbitrary decision made by those that assembled the Standard.

Standard NDT methods used to detect indications do not generally produce the same results, e.g. radiography and ultrasonic methods are both acceptable methods of volumetric flaw detection however they do not always find the same flaws, and when flaws are detected the parameters for rejection are not based on the same principles so one may accept an indication that the other rejects. This further illustrates that the acceptance or rejection of a component based on workmanship is somewhat arbitrary.

Forli [2] notes the random nature of this process and observes that “a quality or workmanship level is something that should be met on an ‘average’ in the long run”. He concludes that the requirements of the applied NDT and acceptance criteria are therefore not more stringent than to reject defects corresponding to the quality level on an average, i.e. in 50% of the cases. This is far from detecting all of the defects all of the time.

3. Misconceptions

In 2002 W. Visser [7] compiled a report on the probability of detection findings of several studies using a large variety of NDT methods. In Table 2 of that report there is a summary overview of the acceptance criteria seen for surface methods (magnetic particle inspection - MPI and dye penetrant - DP) and the volumetric methods (ultrasonic - UT and radiographic testing -RT) in seven different standards from around the world. This table is reproduced here as Table 1. The codes listed are a mix of structural steel and pressure vessel standards.
Table 1  A summary of selected acceptance criteria (from W. Visser [7])

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MPE/DP</td>
<td>surface flaws</td>
<td>not acceptable</td>
<td>not accepted</td>
<td>free of relevant linear indications</td>
<td>free of relevant linear indications</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>isolated porosity</td>
<td>t&lt;sub&gt;4&lt;/sub&gt; and 6mm</td>
<td>t&lt;sub&gt;5&lt;/sub&gt; and 4mm</td>
<td>long length: t&lt;sub&gt;3&lt;/sub&gt; and 20mm round: t&lt;sub&gt;4&lt;/sub&gt;</td>
<td>t&lt;sub&gt;4&lt;/sub&gt; and 4mm</td>
<td>t&lt;sub&gt;3&lt;/sub&gt; and 6mm</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cluster porosity</td>
<td>3mm</td>
<td>2mm</td>
<td>4mm</td>
<td>t and 12.5mm</td>
<td>t&lt;sub&gt;4&lt;/sub&gt; and 5mm</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scattered porosity</td>
<td>20mm</td>
<td>20mm</td>
<td>t&lt;sub&gt;5&lt;/sub&gt; and 20mm</td>
<td>2% of as isolated pores</td>
<td>2% of as isolated pores</td>
<td></td>
</tr>
<tr>
<td></td>
<td>slag inclusion</td>
<td>width: t&lt;sub&gt;4&lt;/sub&gt; and 6mm length: 2t and 50mm</td>
<td>width: t&lt;sub&gt;5&lt;/sub&gt; and 4mm length: 2t</td>
<td>long length: t&lt;sub&gt;3&lt;/sub&gt; and 20mm round: 4mm</td>
<td>long length: t&lt;sub&gt;3&lt;/sub&gt; and 20mm</td>
<td>round: 4mm and 4mm</td>
<td></td>
</tr>
<tr>
<td></td>
<td>incomplete penetration</td>
<td>length: t and 25mm</td>
<td>not acceptable for full penetration welds</td>
<td>not acceptable any size</td>
<td>not acceptable any size</td>
<td>not permitted</td>
<td></td>
</tr>
<tr>
<td></td>
<td>lack of fusion</td>
<td>length: t and 25mm</td>
<td>not acceptable</td>
<td>not acceptable any size</td>
<td>not acceptable any size</td>
<td>not permitted</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cracks</td>
<td>not acceptable</td>
<td>not acceptable</td>
<td>not acceptable any size</td>
<td>not acceptable any size</td>
<td>not permitted</td>
<td></td>
</tr>
<tr>
<td>UT</td>
<td>general</td>
<td>if maximum and length &gt; 40mm and -100% dense see note on NORSOK</td>
<td>if maximum and length &gt; 40mm and -100% dense see note on NORSOK</td>
<td>&gt;100% not acceptable imperfections which produce a response -20% shall be investigated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>porosity</td>
<td>repair if it masks other defects at &gt;50% of ref level length: t&lt;sub&gt;5&lt;/sub&gt; and 10mm</td>
<td>at &gt;100% of ref level length: t&lt;sub&gt;2&lt;/sub&gt; and 10mm</td>
<td>at &gt;100% of ref level width: t&lt;sub&gt;10&lt;/sub&gt; or length: t&lt;sub&gt;2&lt;/sub&gt;</td>
<td>at &gt;100% of ref level width: t&lt;sub&gt;2&lt;/sub&gt; and 20mm at 50-100% if h &gt; 3 then l = 5mm long: 20-50% if h &gt; 3 then 1 = 1 l</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>slag inclusion</td>
<td>if &lt;100%2t and 50mm</td>
<td>depending on characterisation: see cracks</td>
<td>depending on characterisation: see cracks</td>
<td>at &gt;100% unacceptable at 50-100% unacceptable t&lt;sub&gt;3&lt;/sub&gt; and 20mm long: 20-50% if h &gt; 3 then l = 5mm long: 20-50% if h &gt; 3 then 1 = 1 l</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>lack of fusion or</td>
<td>at &lt;100%: t and 25mm</td>
<td>depending on characterisation: see cracks</td>
<td>that imperfections are acceptable in no case at all (original text)</td>
<td>at &gt;100% unacceptable at 50-100% unacceptable t&lt;sub&gt;3&lt;/sub&gt; and 20mm long: 20-50% if h &gt; 3 then l = 5mm long: 20-50% if h &gt; 3 then 1 = 1 l</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>incomplete penetration</td>
<td>at &gt;100%: t and 25mm</td>
<td>characteristic: see cracks</td>
<td>that imperfections are acceptable in no case at all (original text)</td>
<td>at &gt;100% unacceptable at 50-100% unacceptable t&lt;sub&gt;3&lt;/sub&gt; and 20mm long: 20-50% if h &gt; 3 then l = 5mm long: 20-50% if h &gt; 3 then 1 = 1 l</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cracks</td>
<td>unacceptable regardless of size and amplitude</td>
<td>not acceptable at 20% of ref level</td>
<td>not acceptable at 20% of ref level</td>
<td>not acceptable, regardless of amplitude t &gt; 1 = t&lt;sub&gt;3&lt;/sub&gt; and 20mm unacceptable regardless of size and amplitude</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>root defects in single sided welding</td>
<td>echo exceeds ref. curve t and 25mm</td>
<td>--</td>
<td>--</td>
<td>not acceptable</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>

Notes:
- "a and b" implies "not exceeding a and not exceeding b"
- the table is limited to thicknesses t &gt; 20mm
- for NORSOK UT type of defect to be decided by supplementary NDT
- "at &gt;100%" implies 'exceeding the reference curve'
- "at 50-100%" (or at 20-50%) implies between 50% and 100% (or between 20% and 50%) of the ref. curve
- 'long' stands for longitudinal planar defect
- 'perp' stands for perpendicular planar defect
- 'round' stands for rounded defects
The acceptance criteria in the tabulated standards are based on essentially the same approach of NDT, regardless of the Standard. In all cases radiography is applied using X-ray or gamma radiation on Class I or II film. Ultrasonic pulse-echo testing is applied by manual probe movement with standard refracted angles using sensitivities set on side drilled holes (typically 2.5mm to 3mm diameter). Colour-contrast techniques are standard for dye penetrant, and magnetic particle testing is typically accomplished using either wet fluorescent or colour contrast dry particles with an AC yoke.

Similar acceptance criteria can be read in other common standards such as American Welding Society D1.1, American Petroleum Institute API 1104, ISO 11666 (UT), 10675 (RT), ISO 23277 (PT), ISO 23278 (MT).

All such standards have the user identify a flaw condition (indication) and then assign it an allowable length which, if exceeded, constitutes a rejectable condition. Some add to the length criteria a requirement to characterise the origin of the flaw. This is seen in Table 1 where some of the codes indicate that cracks and lack of fusion and lack of penetration are unacceptable regardless of length.

There can be significant problems with these arbitrary requirements. Selection of the level of quality should consider the provisos that are so thoughtfully included in some of the standards (including the statement in ISO 5817 that “Economic factors are also important and should include not only the cost of welding but also of inspection, test and repair.”)

Consider some of the issues:

- Radiographers identifying pore sizes
  - Rarely do radiographers consider the penumbra effect and oversizing is a common result in spite of a relatively low risk from pores in most situations

- Ultrasonic technicians assessment of flaw severity by amplitude
  - Contrary to common beliefs, pulse-echo ultrasonic testing does not provide a direct relationship to flaw severity with respect to amplitude. Figure 6.1 from Visser [7] is illustrated here as Figure 1. It indicates a totally random scatter of amplitude from planar flaws. Setting an arbitrary reject threshold at some level below 100% reference is not related to any degree of flaw criticality. Some UT standards provide multiple levels of amplitude evaluation (as for example does ISO 11666, AWS D1.1 and JIS 3060). When viewed with respect to Figure 1, we can see how this attempt to relate amplitude to flaw criticality is a misconception.
- Ultrasonic technicians characterising flaws by echo-dynamics
  - Although well-considered descriptions of how to use echo-dynamics to characterise flaws is given in training institutes and instructional documents (e.g. EN 1713), the fact remains that the process is not very effective. PANII 3 [8] results state that 32% of respondents correctly characterised the defects on the basis of echo dynamic. Relying on operators to correctly (and reliably) characterise flaws is another misconception in the workmanship standards.
- Dye penetrant technicians evaluating size based on bleedout after minimum dwell time
  - Bleedout is given as a minimum time without clear instruction as to a maximum (other than a caution against waiting too long and a risk of indication degeneration). The variability of perceived indication size in a period of 10-30 minutes could mean that indications viewed early in an area tested might be assessed as smaller (acceptable) compared to the same sized indications viewed 15-20 minutes later in the same tested area.

4. Variability

Provided all parties understand the variability of the “detection process” and the fact that the process as used with workmanship acceptance criteria is not intended to identify all critical flaws, NDT can provide a useful quality control on the fabrication process.

In spite of the supposed procedural controls to ensure reliable results, most general NDT is sufficiently qualitative that results are not likely to be identical if carried out by two different operators with two different kits. A client that decides to do their own double-check of an NDT test using the same NDT method, should not expect identical results.

When inspection results are to be used for quantitative assessment of serviceability or deterioration, it is unlikely that the quality control workmanship acceptance criteria will be suitable. For quantitative results the NDT procedure will need to be demonstrated to be repeatable and reliable and the acceptance criteria need to be more than a simple quality control indicator.
The difference between a general good-practice NDT inspection procedure versus a qualified NDT procedure designed for a particular task is considered in the SINTAP Task 3.4 Final Report [9]. The report concludes that if an NDT inspection is required to suit the needs of structural integrity it will almost always require some form of qualification to demonstrate detection and sizing capabilities.

However, the report also adds to the concerns that even Structural Integrity engineers must be reasonable in their demands. They use as an example, setting goals for reasonable targets (e.g. for stainless steel piping, fully safe defect detection and sentencing of 40% T instead of 10% T).

5. Conclusions

Perhaps the greatest misconception held by NDT technicians and engineers alike is the notion that the standard workmanship criteria found in codes are directly linked to component safety.

When users of workmanship acceptance criteria take the attitude that the test results “must make sure that harmful defects are detected and removed as part of a workmanship standard”, it illustrates that the concept of workmanship is not understood. Perhaps this sort of statement is well-intended; however, in order to define “harmful” the user needs to carry out an Engineering Critical Assessment to determine the size of flaw that might be harmful. Then the inspection technique will need to be qualified to ensure it can reliably detect that size of flaw and be able to reasonably size the indication to compare it to the calculated allowed flaw size.

When using workmanship acceptance criteria, users should keep in mind that its actual function is to simply assess “quality”. One should not hold to some false sense of safety being associated with it.

Another misconception about workmanship acceptance criteria could occur if applying the concepts of “workmanship” to in-service inspections. Whereas a quality level of workmanship might be applicable to the fabrication stage of a component, it is unreasonable to then apply the same concepts of fabrication quality to a product that has undergone the wear and tear of service conditions. At that point, the effects of service conditions rather than workmanship quality should be monitored if there is concern for the continued service of the component.

We see in the SINTAP report [9] that when acceptance criteria move beyond basic workmanship requirements, qualified NDT procedures should be used. But even these should still have reasonable expectations (targets) for detection and sizing. It would therefore seem equally prudent for those using and making workmanship acceptance criteria to consider reasonable levels of quality that consider the limitations and variability of unqualified NDT methods and techniques.

Dedication
I would like to dedicate this paper to Michael Moles who passed away during the preparation of this paper. Michael was a great friend and colleague who contributed much to our industry.
References

5. ISO -5817, Welding — Fusion-welded joints in steel, nickel, titanium and their alloys (beam welding excluded) — Quality levels for imperfections, Published by ISO, Switzerland, 2014
6. ISO 17635:2010, Non-destructive testing of welds -- General rules for metallic materials, Published by ISO, Switzerland, 2010
7. W. Visser, POD/POS curves for non-destructive examination, Prepared by Visser Consultancy Limited for the Health and Safety Executive, 2002