A Rose By Any Other Name: Certification Seen As Process Rather Than Content

John R. Wilson

University of Nottingham

Introduction

It is perhaps indicative of perceptions and knowledge about certification that a small sample of colleagues, on hearing that I was writing this paper, all assumed it was to do with certification of human factors or ergonomics professionals. On the other hand, this may just illustrate the self-referential nature of professional groups and genuine concerns about professional certification, and these concerns are germane to the present argument since certification of systems may in the end be reduced to certification of professionals in the process (this is an issue returned to below). In fact, misunderstanding over the term "certification" may be in part responsible for misgivings and a lack of a wholehearted welcome for it, certainly amongst some ergonomists of the author's acquaintance. If it is seen as a formalisation and standardisation of their activities, then there is considerable opposition. When explained as a "design review and approvals" procedure, response seems more favourable.

Green (1990) believes that the two main factors safeguarding flying from human error are both related to certification and regulation. First is the increasingly proceduralised nature of flying whereby as much as possible is reduced to a rule-based activity. Second is the emphasis placed upon training and competency checking of aircrew in simulators and in the air, both generally and for all particular types of aircraft flown. This leaves, believes Green, other human factors that are relatively unaddressed as yet and which can give rise to human reliability problems. These include: hardware factors and especially the compatibility of control/display relationships and the way information is presented in relation to pilots' expectations; social factors and especially pilot/co-pilot relationships; and system factors including fatigue and cost/safety trade-offs. He also, importantly, identifies problems with the integration of the "electronic crew member" following increased automation. Human reliability failures with artificial intelligence and automation, due to over-reliance on the system fail-safe mechanisms, or to operator under-confidence in the integrity or self-regulating capacity of the system, or to out-of-loop effects, are widely accepted as being due to deficiencies in plant design, planning, management and maintenance more than to "operator error" – Reason's (1990) latent error or organisation pathogens argument. Reliability failures in complex systems are well enough documented to give cause for concern and at least promote a debate on the merits of a full certification programme.

Human Factors Certification of Advanced Aviation Technologies
Edited by J. A. Wise, V. D. Hopkin, and D. J. Garland

The purpose of this short paper is to seek out and explore what is valuable in certification, at the least to show that the benefits outweigh the disadvantages and at best to identify positive outcomes perhaps not obtainable in other ways. On both sides of the debate on certification there is general agreement on the *need* for a better human factors perspective and effort in complex aviation systems design.

What is at issue is how this is to be promoted. It is incumbent upon opponents of certification to say how else such promotion be enabled! This is an exploratory and philosophical review, not a focused and specific one, and it will draw upon much that is not firmly in the domain of complex aviation systems.

Parallels

There used to be an unwritten "law" in work study – or motion and time study in the USA – that no one was so enthusiastic about work measurement and standardisation as those whose own jobs – they felt – precluded any possibility of such a process being applied to them. Equally, no-one was quicker than management to oppose, utterly, any attempt to assess their own work when this was suggested, on the grounds that this was inappropriate, represented an unnecessary effort, and was regardless impossible for any analyst to understand what they really do.

I saw first hand another salutary lesson within the past year or two. One of my students was attached to work with a "high flier" in a major UK consultancy; they were charged with applying quality assurance procedures to the activities of the consultancy itself. This involved vetting all areas of their operation for compliance with BS5750 and ISO9000, the relevant service quality assurance standards. The student, and the formerly popular and high achieving consultant, quickly became the villains of the group, pariahs to be avoided, because they were seeking to propose some formality for the consultants' work, some prescription of how they should operate. And what was the core activity of this consultancy group? Advising industry on the need, processes and procedures for quality assurance!!

What all this illustrates is the difference that perspective and standpoint make to opinions on formal systems, appraisals, standards and review processes. The reviewers or appraisers see them as bringing about order and rationality, and as ensuring that "the best" is retained and "the worst" is identified and eliminated. The reviewed or appraised, on the other hand, see formal systems as restrictive and petty, unnecessary interferences with their activities, and as leading to "throwing the baby out with the bathwater."

What other parallels can be drawn? First of course, and top of the agenda for many in the human factors community, is the notion of certification for human factors and ergonomics professionals. This is taking place in proposals for the Centre for Registration of Ergonomists in Europe (CREE) scheme, but has substance with the Board of Certification in Professional Ergonomics in the USA. A recent "Provocations" article in Ergonomics in Design (Senders & Harwood, 1993) took contributions from both sides of the argument. Senders and Harwood themselves point out that one danger is of formality driving out reality, in that any degree or certificate may become more important in itself than the individual's competence it implies.
(Educators are, in fact, often faced with this from their students, when attempts to discuss and explore ideas are met by the students' desire to digest directive information tailored to an

examination setting.) In the same piece, Schumacher and Dorst question certification in terms of need, process and impact. If we summarise and generalise some of their objections, these are:

- certification does not *ensure* quality (or integrity or ethics)
- what position will certification have in law?
- do we want the homogeneity that certification might bring?
- prescriptive criteria may stifle innovative designs.

Perhaps reflecting the difficulty of doing so, the response from Hendrick in the same paper makes little attempt to answer directly some of the questions about whether professional competence certification is needed at all, concentrating instead on defending the processes involved. He does, though, argue that certification can promote growth of a discipline and its image, although admitting that the process cannot guarantee "worthwhile job performance... competency...[or]...conformance to ethical, moral and professional standards" (Senders & Harwood, 1993).

What then can possibly be the arguments for certification, if it cannot even guarantee basic compliancies? To find some value of certification, perhaps we can look at other areas of ergonomics. In the domain of work organisation and job design, and in particular the implementation of changed work structures, it is often argued that the content of any change is of minor importance for successful outcomes compared to the change process. If the best process possible is put in place then we can afford to "change the change", iterating on the actual content as required. Translating this idea to certification in complex aviation systems, we could look upon certification as the means by which an improved development process is enabled, rather than as a limitation and detailed specification of the content of the development.

From the fields of product liability (in the 1970's and 1980's) and health and safety at work (in the 1990's), we can see some of the more systemic benefits possibly accruing from certification. Out of the imposition of regimes of strict liability have come better processes or systems of design amongst producers, and movement to a more beneficial standards regime, that of horizontal standards, amongst the lawmakers (see below). Consequences of the health and safety (ergonomics) legislation implemented in 1992/1993 across member states of the European Community as a result of EC Directives are even more marked. Although it would not be appropriate to be too starry-eyed and naive about beneficial outcomes, it certainly seems as if the need for employer conformance with ergonomics criteria and practices has stimulated the production of tools, techniques, instruments and methodologies for investigation and diagnosis that will be of value across a range of concerns. Not all of these developments are to be widely welcomed of course – I imagine we have all been shocked by some of the so-called "ergonomics aids" now on the market – but the overall effect in general has been one of dynamic growth in the discipline. The ergonomics community itself has had to produce new approaches and techniques, improve the validation of existing ones, and generally ensure greater justification for its guidelines and recommendations. Even colleagues who are dubious about the value or validity of specific requirements in the new regulations have been pleasantly surprised by the consequent pressures for quality in methods.

MANPRINT – Lessons Learned

Can we learn from very close parallels to human factors certification in complex aviation systems? Nuclear or military systems perhaps form comparable domains to an extent.

The well known MANPRINT developed for procurement in the U.S. Army has been adopted by both the British Army and, in modified form, the Royal Navy. Reasons claimed by the British Ministry of Defence for its adoption are as follows (MoD, 1992):

- The perceived success of MANPRINT in the USA, including improved maintainability of equipment and use of analytical techniques to ensure wider and better usability;
- A desire to identify and achieve the best balance between people and equipment. It is recognised that "high quality, multi-capable...better equipped motivated and properly trained" personnel will only result if manpower issues are considered as part of the equipment procurement process.
- Personnel costs now outweigh equipment costs and it is hoped to provide more control over these by being better able to anticipate, budget for or reduce costs through design improvements. It is also anticipated that MANPRINT would give a better specification generally and thus more cost-effective products, reducing overmanning, poor performance and errors.
- Better working conditions and reduced training costs;
- Greater requirements for cognitive skills rather than physical and the shrinking pool of skilled labour available, mean it is desirable to constrain designers to produce operable equipment for existing specified personnel. These reasons are labelled "skills drift" and "demographic trough or slide" respectively (Goom, 1993).
- New health and safety legislation applies to military as well as civilian systems and "covers areas which have traditionally been regarded as usability rather than safety issues", again an argument for the broad approach of MANPRINT.
- MANPRINT covers manpower, personnel, training, human factors engineering, health hazard assessment, and system safety (and habitability and environmental ergonomics for the navy). This ensures a single source of advice across all human factors issues, thus preventing or reducing sub-optimality in systems design.

Such support for MANPRINT and presumably for similar certification systems raises three questions. *First, are these claimed advantages real, secondly are they important, thirdly are they generalisable, especially to civil systems?* A sceptic might answer "no", "partly", and "no" to these questions. Certainly it is easy to be cynical about any claims on the part of the military establishment to be making efforts to reduce costs for instance. Nonetheless, a more reasonable view might be to answer: "the claimed advantages seem reasonable"; "yes, they are potentially very important"; "they might be generalised to other situations in other industries."

In fact, the strongest support for the certification process might derive from the fact that the claimed advantages are as much or more concerned with process as they are with content. If the key gains reported for MANPRINT are summarised and generalised, they look like:

- better human-machine systems designs and improved usability
- improved cost-effectiveness and cost control
- widening of the user base
- compliance with ergonomics/health and safety legislation
- more efficient design process.

Generally translated into systems design then, whilst certification might seem to be a cumbersome route to go down, looking at the potential benefits then the spin-offs might be decisive in judgements as to its value.

Standards for the Certification Process

Certification will require agreement on standards so that it may be useful and feasible. Debate over type and coverage of standards has a long history in the field of product regulation for instance. "The trade-off between voluntary and mandatory standards [concerns] acceptability, applicability and ease of formulation versus possible non-compliance ... even standards enshrined in legislation are of little value unless there is strict enforcement ...." (Wilson, 1984). Much support was given to the notion of performance standards rather than construction or dimensional standards. Problems with safety standards were further identified as: their inadequacy in scope and permissable levels of risk; their not addressing all foreseeable hazards or types of behaviour; specifications that tend to be generalised, partial and inadequate; and a general lack of a standardised format. Although these criticisms are still valid today for product ergonomics standards, there has been a major change in direction away from product or vertical standards and towards horizontal and generally hazard-oriented standards. Advantages of these are said to be faster development, easier updating, greater applicability, more consistency and better clarity and understanding about necessary safety levels (van Weperen, 1992).

Meister (1984) differentiates "attribute" standards, which describe how the product should appear or should function, and "performance" standards, which describe how the design product should perform (pp.215-217, 256-263). He sees the former as being general and applying mainly to the component or equipment level and the latter as particular at the subsystem/system level. He criticises the state of human factors standards in much the same way as consumer product standards have been criticised. As a consequence, Meister (1989) has subsequently stated that "... whether because the standard lacks substantive data support or because human factors is generally viewed ... as a constraint on ... freedom to design, MIL-STD 1472C [for instance] is honoured as much in the breach as in the observance."

In summary, if we are to have certification, then it must be related to some norms or standards or standardised procedures. In other words, a system might be certified if it can be shown to have attributes which meet certain recommended values or if its performance meets acceptable limits on certain recommended criteria. We can add to this a third form of certification, if it is shown that defined analysis or test methods have been applied to the design. In this last case, of course, the methods themselves will have to be certified first. Of relevance to complex aviation technology is a particular case of the last, "there is little doubt that a principal future use of simulators will be for licensing and certification" (Jones, Hennessy, & Deutsch, 1985). Principally used now for pilot training and proficiency approval, there seems no reason why artificial intelligence in the cockpit, and particularly its interaction with crew, cannot also be assessed in simulators. The interesting issue then is certification of the simulation system itself!

Currently under consideration is STANAG 3994 AI – the NATO Standardisation Agreement on the Application of Human Engineering to Advanced Aircrew Systems. The intention of this is to "standardise methods for the integration of human engineering procedures with the design and development of advanced aircrew systems." The purpose also is to provide a basis for

agreements between contractors and procuring agencies on human factors scope, context, techniques and criteria. The draft STANAG 3994 AI defines a "general human engineering program model for military systems". The core requirements include those to do with analyses:

- **system analysis** – mission analysis, function analysis, potential operator capability analysis, potential equipment identification, function allocation
- **analysis of operator/maintainer tasks** – timeline analysis, task analysis, critical task analysis, decision analysis, error analysis, loading analysis
- **preliminary system and subsystem design** – information requirements analysis, control requirements analysis, workspace requirements analysis, and environmental analysis.

In addition, a number of other issues are also defined as requiring agreement between agency and contractor. These are: programmes of research involving experiments, testing and dynamic simulations with human subjects; detail on application of relevant human factors standards; development of software and hardware procedures in operation and maintenance; production of mock-ups and models for conformance testing; and development and planning for test and evaluation, including criteria justification and test interpretation details.

Perhaps most critical in terms of any sustainable argument for certification are the requirements to prepare a "human engineering programme plan." This must identify standards of relevance to the system and must identify what human engineering activities will be involved, time-scales and criteria, and should indicate how formal interaction between human factors specialists and other relevant design specialists will be achieved. Finally, provision is made for a tailoring of details in the standard to meet any specific requirements of the system under development.

Taking a domain outsider's viewpoint, several things seem apparent about this STANAG 3994. First, it appears at first sight to be complex and unwieldy, with potential for great overlap in particular amongst the many different analyses. Even co-ordinating these, making sure they are complementary but not too duplicatory (or even contradictory) will be a considerable project management task. Secondly, on the other hand, it is to be welcomed that so much emphasis is placed upon analytical activities and not on prescriptions of design detail. However, within some of these analyses are implied checklist comparison procedures, for instance "... multi-function control modes, and display menu selections shall be analysed ... and the resulting ... structure shall be plotted and analysed for ease and effectiveness of use", or "workspace [requirements] shall be analysed in terms of their access, vision, reach, egress, and emergency requirements, for the range of body sizes, clothing and protective equipment ..." (p5). This will presumably bring into play a plethora of other standards, guidelines and recommendations, the effect of which *may* be to impose a degree of complexity and restriction on design which is not commensurate with finding innovative solutions. Thirdly, again on the positive side, the document does allow for flexibility in its provisions according to circumstances. Finally, and thinking again about systemic gains, it may be that the most important benefit is the integration and collaboration required between ergonomists and engineers.

Any undue complexity, as suggested above, may have excessive cost implications. For instance, even under present regulation systems: "

> It is relatively easy for the profitable airline...but the airline operating in a more competitive area of the aviation system, where economic margins are extremely constrained, may simply be unable to undertake all of the desirable training and

> standardization of equipment without going out of business. The regulatory authority may have considerable difficulties in compelling such airlines to undertake costly procedures as the airlines may accurately point out that by doing so they will be made less cost efficient vis a vis foreign operators (possibly operating in a less regulated environment) with whom they compete directly...The temptation for operator and regulator alike, when faced with an acknowledged but intractable problem, is to undertake some unconscious dissonance resolution by regarding the problem as less serious than they might if it were readily soluble. (Green, 1990, p510)

If this concern applies to aviation systems developers and suppliers as well as to the operating companies, then the impact on the workability of any certification process may be serious.

To certify or not to certify?

So, why is someone who typically dislikes regulation, systematisation and quality assurance generally writing here to support certification in complex systems? The answer lies in what has been stressed above, namely that the systemic outcomes of having a certification programme in place may be advantageous enough to overcome any drawbacks of the regulatory regime and the content of any standards.

As a start, if human factors certification is to work in any domain we need to consider why it has not been in place there previously and address all potential reasons very seriously. The author is currently working with a very large, reputable transnational company. The design engineers pass their designs through every conceivable review and approval process – HAZOP, P & I Approval, Environmental Impact Assessment, Engineering Audit etc. – but, at the moment, there is not any human factors approval. We asked ourselves why such an absence of any formal human factors standards approval system exists. Possible reasons include:

- Historically, ergonomics has not been seen as important by engineers, managers etc.
- Ergonomics has been assumed to be included in all the other types of approvals and standards.
- Certification of human factors is not seen as cost-effective (in terms of there being few gains); there will still be problems afterwards (since people are seen as fallible), but much time and energy will have been expended meanwhile.
- Certification is genuinely seen, by engineers and/or ergonomists as not required
- Human factors certification may be resisted by ergonomists themselves, perhaps because of the requirements or restrictions it may put on them.
- Certification is seen as impossible to do, or at least impossible to do well.

Before even beginning to introduce ergonomics design review to this company, before even planning what might be included, we need to examine these reasons, see which are apparent in this case, and address the organisation issues involved.

The Case Against

There are two ways to mount a defence of certification: to counter or at least downplay the criticisms, and to promote advantages and benefits. Taking first the criticisms, a number of arguments against certification have been rehearsed implicitly or explicitly in this paper.

Certification may firstly be seen as *unnecessary*; presumably in this view either there is little or no room for improvement in ergonomic design of systems (hardly a sustainable argument) or else that the aviation world is self-correcting. That is, up to a point human factors deficiencies will be rectified anyway during development and commissioning and where they are too large or deep seated then the system itself will not remain in operation. I find this argument unconvincing unless we allow major failures in operation to also be a part of this self correcting process. However, what must be accepted is that, currently, we are talking about remarkably reliable and safe systems of hardware, software, procedures, communications and people.

An extension of the first objection is that certification is unnecessary for simple and/or relatively stable systems and is *impossible* to do adequately for complex systems. There may be some validity to this argument but the advantages of an improved quality of development process discussed below might counteract it to some extent. What must be recognised, though, is that any system of approvals must allow for trade-offs between human factors and between human and technical factors in its operation.

Further criticism might be that a certification regime is *restrictive* and *cumbersome*. If we replace "is" by "can be" then I would agree here. However, if we aim certification at performance and at assessments as against at design specifications, allow tailoring of standards to meet circumstances, and – most importantly – make ergonomics design review and approvals an intrinsic part of development rather than an extraneous add-on, this will all help dilute this complaint. Similarly, we can meet objections that certification might stifle innovation and lead to homogeneity in design.

Certification might also be criticised on the grounds of its being *misdirected*, with standards aimed only at reducing the incidence or consequences of active errors. In this view, standards may be much less help with latent failures or resident pathogens in the system; these are the system problems which may have lain dormant in the system for a long time and which are spawned by the activities of designers, managers and, indeed, regulators themselves (Reason, 1990). One could take a positive view though, that in fact it is these "violations", giving rise to latent errors, which are best attacked through a process of certification, due to the process itself being a good discipline upon all involved in high level planning and decision making.

As for a fourth set of criticisms, that certification is *untestable*, widely *unacceptable* and thus *unworkable* or *unenforceable*, this will largely be a function of the particular regulatory regime. To repeat again an earlier point, if a system of certification can be constructed such that it is seen to improve and streamline the development process and time as well as increase systems integrity, then acceptance will be more widespread.

The Case For

The case for certification can be made positively, as well as by minimising the validity or consequences of criticisms as above. Three areas of benefit may be defined, all systemic in

nature in that they emanate from the fact that human factors certification of complex systems will have effects beyond defining and ensuring compliance with human factors standards.

First, we have the improvements in the design process that might be expected. Knowledge that a system must be certified in terms of human factors may not ensure all correct detail design alternatives are chosen – indeed, there is no such thing as a perfect design given all the trade-offs that must be made. However, it will mean a greater likelihood that all relevant issues are addressed and their consequences assessed much earlier in development. Costly changes after prototyping or even during commissioning trials can be reduced in frequency and extent. A related benefit is that for any suppliers to be able to meet future certification requirements, technical and financial decision makers will have to coordinate much earlier and better with those responsible for the human factors.

A second benefit is predicted, based upon experience in other domains where ergonomics standards have been introduced or toughened. In the act of formulating, specifying and testing the process and procedures necessary to allow systems to be certified, the human factors community will have to respond to pressures for better methods, techniques and criteria, and will have to validate, justify and report them better.

Finally, this improved "professionalism" in human factors, the perceived benefits to the design process, and – if experience in industrial health and safety and ergonomics is any guide – increased interest of engineers in the resulting need for increased problem solving, will all act to produce a more human-centred design approach. Thus, through both the fact and process of certification, as much or more than by its content, the design of complex aviation systems will be improved.

Conclusions

We should not talk of certification only as a choice of two options – to certify or not to certify. If we draw an analogy in politics, like saying that electorates have a choice between the authoritarian right (prescription, control, punitive consequences of non-compliance) and the libertarian right (the individual has an absolute right to do as he/she pleases, and the 'market' will ensure instability is kept in bounds), such debates see anarchy as the only outcome if a choice between the two options is not made. There *are* other paths in government however, whereby individuals have rights or freedoms but also responsibilities towards society, and where society attempts to redress imbalances in power. Thus, a regime of certification *can* be implemented such that it provides a framework for complex systems design, a benchmark to aim for, and a bulwark against very poor design, whilst still allowing for innovation and not imposing too costly or cumbersome a design regime.

What must be addressed before instituting a formal certification system are the needs of such a programme. Benefits will be realised and disadvantages or problems minimised only when decisions are made on:

- Distinctions to be made and balance to be found between attribute, performance, personnel and process certification
- Desirable degree of prescription or latitude for design
- Identification, definition, agreement and validation of test methods, measures, criteria etc.

- Provision for flexibility and updating of requirements
- Examination of trade-offs between value and cost/time of the certification process
- Systems to certify the certifiers/certification systems
- Communication of outcomes of the certification process in more useful terms than just a pass/fail, yes/no judgement
- Consideration of implications of non-conformance, and thus enforcement

It must be stressed once again that the process of certification can be of value even if we are unsure of or unhappy about the content when it is first instituted. More than this, if we get the process right, then content problems – in appropriate requirements or missing tools or data for instance – will be rectified as part of the process being put into operation. We must remember, though, that the individuals who produce certification processes or who test and approve systems are themselves fallible, as also will be any intelligent systems built to help with certification. Perhaps this is the key issue for acceptance of certification – Quis custodiet ipsos custodes?

References

Goom, M.K. (1993), An industrial view of MANPRINT. Proceedings of the Annual Conference of the Ergonomics Society (ed E J Lovesey). Heriot Watt University. London: Taylor & Francis.

Green, R. (1990), Human error on the flight deck. *Phil. Trans. R. Soc. Land.B., 327,* 503-512.

Jones, E.R., Hennessy, R.T. & Deutsch, S. (eds), (1985). *Human Factors Aspects of Simulation.* Washington D.C.: National Academy Press.

Meister, D. (1985). *Behavioural Analysis and Measurement Methods.* New York: J. Wiley & Sons.

Meister, D., (1989). *Conceptual Aspects of Human Factors.* Baltimore: John Hopkins University Press.

Ministry of Defence. (1992). *The MANPRINT Handbook (2nd ed.).* London: HMSO.

Reason, J. (1990). The contribution of latent human failures to the breakdown of complex systems. *Phil. Trans. R. Soc. Land.B., 327,* 475-484.

Senders, J.W. & Harwood, K. (1993). Provocations: To certify or not to certify. *Ergonomics in Design, 1,* April, 8-11.

STANAG 3994 AI, *The Application of Human Engineering to Advanced Aircrew Systems.* Draft NATO Standardization Agreement.

van Weperen, W. (1992). A hazard-oriented approach to product safety criteria. *Proceedings of ECOSA Workshop on Product Safety Research in Europe,* Amsterdam, July, 10-19.

Wilson, J.R. (1984). Standards for product safety design: A framework for their production. *Applied Ergonomics, 15,* 203-210.