An Approach to Human Factors Validation

G. M. Samaras, PhD, DSc, PE, CPE, CQE

Samaras & Associates, Inc., Pueblo, CO USA

george@samaras.eng.pro

ABSTRACT

User and use errors may be the last bastion of equipment and process safety and effectiveness problems. Human factors and ergonomics (HF&E) is fundamentally about designing, building, training, and maintaining products and processes for human use. The contributions that HF&E can make to engineering validation and validation measurements are briefly reviewed. An approach for formulating and validating HF&E requirements is presented. Two practical examples are discussed, one using statistically designed experiments during the development process and the other validating operator use of production equipment. Two important limitations of engineering validation are identified. It is concluded that proper application of HF&E will permit rational and cost-effective validation of medical and pharmaceutical equipment use and the equipment user's training and work environment.

KEYWORDS:

Design of Experiments, Ergonomics, Human Factors, Requirements Formulation, Use Error, Validation

BIOGRAPHICAL NOTE:

George Samaras is a scientist & engineer in private practice. He is a licensed professional engineer (PE), an ASQ-certified quality engineer (CQE) and a board-certified professional ergonomist (CPE). He was trained as an electrical engineer and earned doctorates from the University of Maryland (PhD; 1976; Physiology & Biopsychology) and The George Washington University (DSc; 1992; Engineering Management & I/O Psychology).

Introduction

User and use errors may be the last bastion of equipment and process safety and effectiveness problems. The source of user and use errors can be the result of problems with equipment design and/or can be the result of problems with user training, work structure, and the work environment. While it was previously thought that design problems could be alleviated with training, this is now recognized not to be the case, in general. The discipline of human factors and ergonomics (HF&E) has a body of knowledge to deal with the interaction of humans and their artifacts and with the design of systems in which people participate. The objective of this article is to present an approach to validating human factors (of equipment design) and human actors (training and work environment). Detailed examples for validation of both equipment design (a medical device) and equipment use (a pharmaceutical manufacturing system) will be given.

Cronbach and Meehl [1] have indicated that validation is not essentially different from the general scientific procedures for developing and supporting theories. However, the engineering concept of validation arises from classical systems engineering [2] and is based upon proper, operationally-defined formulation of system-focused requirements. Complete and correct requirements satisfice [3, 204] all stakeholders (obtain a good result that is good enough, although not necessarily the best, for each stakeholder), inform designers, and provide a basis for quantitative validation measurements; this applies both to system development and after system deployment. Requirement Engineering forms the basis of the validation effort – whose purpose is to demonstrate that you cannot reasonably *refute* [4, pg. 48] the assertion that the correct project has been completed. The practice of some in validation attempting to prove or confirm that the system meets the requirements (“affirming the consequent”) is a well-known invalid form in logic. The correct objective of validation is to attempt to refute that the system meets the requirements (“denying the consequent” or *modus tollens*); the objective here is to design empirical tests that might reasonably cause the system to “break” or fail. What constitutes reasonable? It is risk-based and guided by a proper risk analysis. If hazard analysis (HA) indicates mission-critical process failure or loss of life or property or damage to the environment, the standard must be high. Defective validation studies may usually be traced directly to defective requirements formulation (requirements that are incomplete, incorrect, or misleading).

We will review some fundamentals of human factors and related measurements, discuss an approach to formulating user and use requirements that specify quantitative validation and then consider two types of human factors validations: validation of equipment design and validation of equipment use. In both cases, a simplified example (a medical device and a pharmaceutical manufacturing system) will be used to illustrate the approach. We conclude by identifying some of the limitations of empirical validation.

Human Factors & Ergonomics

HF&E is fundamentally about designing, building, training, and maintaining for human use. The discipline is generally divided into micro-ergonomics (design and management of tools for human use) and macro-ergonomics (design and management of human organizations involved in tool use). HF&E considerations [5] participate in a manner similar to hardware, software, and economic considerations in the development of requirements, in compliance with appropriate regulations and standards, and in the engineering of system reliability and system integrity (e.g., periodic re-validations). HF&E considerations are important at all stages of the system lifecycle (development, deployment, and disposal), since validations (and their pre-requisite requirements formulations) occur in each of these stages.

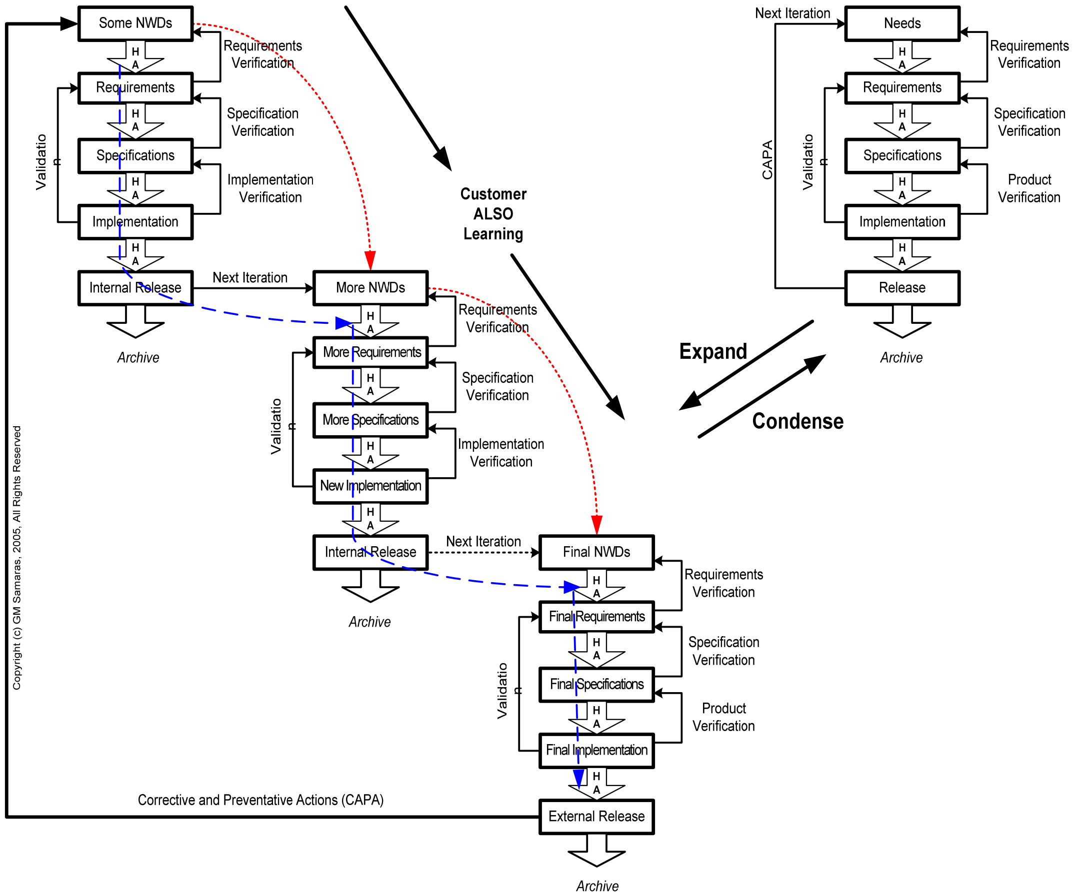

Figure 1: SE Lifecycle Notations - Expanded versus Condensed

While requirements engineering can also occur during deployment and disposal, the product development process perspective offers the most familiar view of requirements engineering. The first step in the iterative learning process (Figure 1, [6]) is identification of the needs of the system users - which presupposes that you have correctly identified the universe of user populations (manufacturers, assemblers, operators, clinicians, patients, maintainers, disposers, etc.)

The assessment of user needs, wants, and desires (NWDs) is a complex activity that often has been implemented by marketing personnel with ad hoc engineering support; in fact, it is a central area of expertise and practice in ergonomics. Some examples of NWD assessment techniques include interviews, questionnaires, and ethno-methodological studies, brainstorming, problem-domain storyboarding, prototyping, literature reviews and ergonomics laboratory research, as well as evolutionary (rapid & iterative) development techniques. Kansei engineering, pioneered by Mitsuo Nagamachi [7], directly addresses the issue of assessment of “desires” in a quantitative fashion. Both from a good business practices perspective and from a FDA regulatory perspective, they all must be implemented in a statistically valid manner [8], so that the results truly represent the populations under study.

Once the NWDs have been determined, the next task is to translate the subset of NWDs, selected to be met, into requirements of a system. This activity also requires the knowledge and skills of ergonomics. Requirements are the foundation of the validation process and a crucial source of the engineering design specifications (Figure 1). When dealing with any system, particularly those in which proprietary software or database content run on generic hardware, there will be issues such as response time, throughput rate, load balancing, disaster recovery, system availability, reliability, and maintainability. It is helpful to treat user interface characteristics in the same manner as these system performance variables, setting usability objectives for the system in measurable terms, typically couched in terms of effectiveness, efficiency, and user satisfaction as identified in ISO 13407:1999 [9].

In the next section, we will discuss, in detail, an approach to formulation of HF&E requirements. However, once the requirements are properly established and verified against the NWDs, the next task is to translate these natural language statements into engineering design specifications. Engineering design specifications are the true basis for the product design and are quantitative product attributes with associated units and tolerances. Ergonomic knowledge can play a crucial role, directly impacting the final design of the product:

1. *Hardware ergonomics perspective*: the ergonomist not only has access to tabulated human cognitive and perceptual data, and as appropriate, anthropometric data, which can dictate physical specifications, but the ergonomist is trained to properly use these data in the realization of engineering designs.
2. *Software ergonomics perspective*: the ergonomist is trained to participate in the design of user interfaces, to conduct task analyses on the proposed logical operation of the product, and to participate in the design of training, operation, and maintenance materials.
3. *Environmental ergonomics perspective*: the ergonomist can assist the design team in assessing how known workspace environmental modalities can impact the use and reliability of the proposed design (e.g., effects of temperature, humidity, lighting, ambient noise, and air quality on user fatigue, perceptual, and cognitive abilities).

4. *Macro-ergonomics perspective*: some ergonomists can assist the organization in harmonizing the design of the product with the way the purchaser organization does business; from inside their own product development organization, these same ergonomists can be called upon to help harmonize their own organization with the product development process, with the manufacturing process, with the product distribution process, and/or with the product field support process.

Prior to the detailed discussion of an approach to engineering human factors-related system requirements, we need to identify the general measurement categories from which operational definitions may be derived in order to permit empirical validation measurements. Figure 2 identifies four general user measurement categories. By overt we mean openly observable, not hidden or concealed; conversely, by covert we do mean hidden or concealed. In order to measure covert phenomena, we need to identify and measure overt resultants (e.g., *force*, a covert physical quantity, is related to the second time derivative of a *displacement*, an overt physical quantity).

	OVERT	**COVERT**
PHYSICAL	**Anthropometry**	**Biomechanics**
BEHAVIORAL	**Verbal, Non-Verbal**	**Affective, Cognitive, Physiological**

Figure 2: User Measurement Categories

Overt physical measurements include such things as length and mass (available as tabulated anthropomorphometric data) related to essentially static human characteristics, whereas covert physical measurements include such things as force and acceleration related to the dynamics of the human body. The technology for making such measurements is well developed [e.g., 10,11]. Properly constructed system requirements that operationally define overt physical measurements might include the dimensions of a cockpit, an infusion pump buttons' dimensions (but not layout), or saw table height; those that operationally define covert physical measurements might include the forces necessary to operate a stick, install a pump cassette, or push a piece through the saw.

Overt behavioral measurements include such things as verbal and non-verbal responses related to external or internal stimuli. The measurement technology is routinely used in experimental psychology. Often these verbal and non-verbal responses are videotaped for later analysis (e.g., see the example presented on pharmaceutical manufacturing). Properly constructed system requirements that operationally define overt behavioral measurements might include the requisite elements of the conversational content with air traffic control (ATC), the layout of push buttons on the pump that would minimize sequence errors, and structures that support moving the piece into the saw at a certain rate.

Covert behavioral measurements include analytical (which rely heavily on prior information), subjective (which rely on self-reporting), performance (using a secondary task), and psychophysiological measures (measuring physiological functions that are believed to covary with cognitive functions). There exists a large body of work in cognitive work analysis (CWA) and cognitive systems engineering (CSE) [e.g., 12,13,14,15]. CSE is not "systems engineering" as its name might imply; it is a requirements engineering approach and, to be consistent with long established nomenclature, it should be termed "cognitive requirements engineering". It is a research strategy whose outputs support formulation of *requirements* for the development of tangible products; these requirements should be system-focused, not user-focused. At present, there still exist gaps between these CWA/CSE outputs and properly formulated system-focused requirements [16,17]. It has recently been pointed out that, "… the test of CSE as a research strategy is its ability to identify basic requirements for how to support cognitive work that must be met, if new technology will be useful to practitioners in context"[18].

An Approach to HF&E Requirements Formulation and Validation

For SE to be successful, "three lines of development – the *user*, *hardware*, and *software* – have to be managed and woven into an integrated product throughout the process [19, pg. 38]." SE validation is based upon properly formulated requirements that operationally define an empirical study of an implementation. Consider a frequently stated "requirement" imposed upon a design team: *The system must be easy to use*! This statement is not a "requirement"; it may be an NWD, but absent operational definitions, it can NOT be a requirement. An SE requirement is a natural language statement that operationally defines the validation measurement(s). Requirements Engineering is that engineering activity of discovering stakeholder NWDs, selecting those NWDs that will be translated into requirements, and formulating the requirements, so that they satisfice stakeholders, inform designers, and provide a complete and correct basis for validation. Let us consider a simplified example of how do we "engineer" the NWD "*The system must be easy to use*" into a system requirement? One approach is a process consisting of three phases: *analysis*, *elaboration*, and *synthesis*; the process repeats, in each iteration, for newly discovered NWDs.

Analysis Phase

Analysis is used in the sense of "*disclosing or working back to what is more fundamental by means of which something can be explained*" [20]. The objective is to take words or expressions with complex connotations and deduce a set of elemental concepts that can be operationally defined. Users may be conceptualized in terms of structures and behaviors (21, pg. 132); we will consider all overt and covert physical and behavioral attributes. In this NWD, we have three complex terms: "system", "easy", and "use".

The term "**system**" in a system development scenario does not include the user(s). From an SE perspective, the "system" is only what the designers can build. It is essential to specify the system boundaries, outside of which exists the environment over which the system designers have no control. The system could be, for example, an infusion pump for administration of intravenous fluids and medications. The boundaries form the operational definition of the "system" [22]. For system development, the system, not the user(s), is the target of validation. However, for validation of user and use (training and operation), the users' knowledge, skills, and abilities would be the object of the validation.

The term “**use**” is synonymous with “operate”. It may be operationally defined as a set of specific behavioral sequences, for a specific set of conditions, which are completed within a specified time. Operation may consist of “covert physical” operation (there is no morphological component, only a biomechanical component) and/or “behavioral” operation. The term “covert behavioral” operation consists of covert observations, computations, and decisions. Their detection by “overt behavioral” operation might include gaze direction, verbal responses, and non-verbal responses (that have biomechanical & physiological characteristics).

The term “**easy**”, from the user’s perspective, may consist of “physically” easy and/or “behaviorally” easy. In both cases, the concept “easy” exists somewhere on the beginning of a continuum from “intuitive” through to “impossibly difficult”. Before we reach the end of this continuum, more and more training and experience, will be required; however, toward the beginning (“intuitive”) of the continuum, little or no training and experience will be required for acceptable “use”. The term “physically” easy consists of a morphological component (e.g., size of pump front panel buttons) and a biomechanical component (e.g., syringe cassette installation force). The term “behaviorally” easy consists of an overt behavioral component (e.g., locating and pushing buttons in a certain sequence) and a covert behavioral component (e.g., deciding the button sequence).

So far, having analyzed the three original terms “system”, “use”, and “easy”, we have no serious challenge to our current measurement capability.

Elaboration Phase

Elaboration is used in the sense of *providing additional information in intricate and painstaking detail*. Compliance Engineering activities (identifying constraints) participate in the elaboration phase. The objective is to identify clarifying and supplemental information that can be operationally defined and that constrains the requirement(s) to our understanding of the real world. Some simple examples of elaboration are:

1. Identify ***who*** is the user population (e.g., floor nurse, equipment operator) from both a morphological (e.g., age, gender, ethnicity) and experiential (e.g., Registered Nurse, high school graduate) perspective. This will permit use of tabulated human perceptual, cognitive, and anthropomorphometric data and presumption of a specific range of knowledge, skills, and abilities that inform establishing training, operation, and maintenance materials. Cultural and national identifications will permit consideration of fundamental differences in conventions and expectations.
2. Identify the full range of external conditions (e.g., low light levels, mass casualties, equipment failures) and internal conditions (e.g., fatigue at the end of a shift, perceived time constraints) ***when*** the user(s) will be operating. This will permit consideration of whether the users' physical, perceptual or cognitive capabilities may be exceeded in that environment. Make explicit seller/purchaser macro-ergonomic issues (e.g., user work scheduling, product field support, sociotechnical aspects of the development and customer organizations, etc.). Changing technology often necessitates changing organizational policies and procedures and is best accomplished with a macro-ergonomic intervention.

3. Identify ***what*** are the general modes [19, pg. 10] of expected use, unexpected use, misuse, and abuse. The specific modes will identify specific behavioral sequences, for a specific set of conditions, completed within a specified time (known after each intermediate “Implementation”, as shown in Figure 1, is completed and ready for validation).
4. Identify ***where*** the use will occur (e.g., hospital, manufacturing plant, etc.).

While these examples of the elaboration process primarily emphasize the operator (e.g., nurse, operator), it is important that all intended users (e.g., clinical engineer, equipment repair technician) be considered [23].

Synthesis Phase

Synthesis is used in the sense of *recombination of ideas into a complex whole*. The objective is to organize logically the various elements identified or discovered in the analysis and elaboration phases, so that it satisfices stakeholders and informs designers for each iteration.

Satisficing all stakeholders (e.g., customers, users, disposers, developers, producers and managers) requires understanding each group’s NWDs and then prioritizing the resultant requirements, so that appropriate tradeoffs can be made in a systematic fashion. There are two general approaches: design–dependent and design independent.

Design Dependent Approach

Quality Function Deployment (QFD) is a design-dependent approach of formulating requirements. Figure 3 shows QFD in the context of the SE condensed notation identified in Figure 1 [6]. QFD employs a process of listening to the "voice of the customer" [24, pg. 9] to discover, identify and understand NWDs. These NWDs are used to develop the *quality dimension* (synonymous with SE requirements), the "whats". The "whats" are prioritized based upon their importance to each stakeholder. Putative designs (the "hows") are identified and a relationship matrix – relating the "whats" and the "hows" – is constructed. A correlation matrix is also constructed - among the "hows" – permitting identification of conflicts between putative design elements. It is important to note that QFD does not result in design-independent requirements, since putative designs participate in the selection of requirements. However, for existing equipment or acceptance testing without vendor supplied requirements, the QFD approach has great value, since you typically are unable or unwilling to modify the equipment.

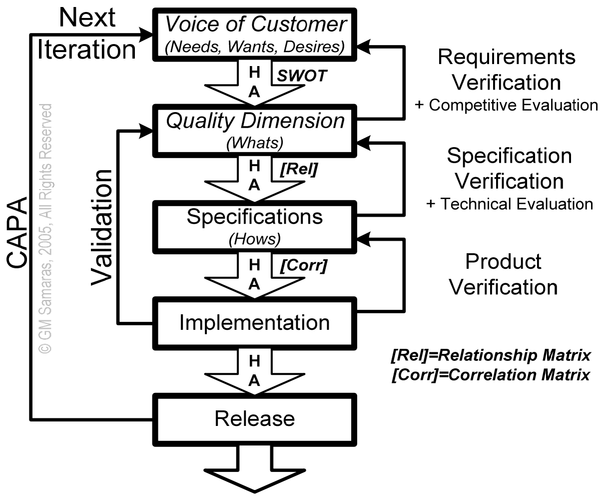

Figure 3: QFD mapped to SE

Design Independent Approach

Informing designers means that the requirements refer to the **system [**19, pg 272], not to the user(s)! There exist at least two equivalent methods of defining design-independent requirements: (a) Use Cases that define requirements in context and (b) Requirements Specification that do not specify the use context. In both cases, the synthesis phase consists of organizing the results of the analysis and elaboration phases into a logical, understandable whole. For Use Cases, the synthesis consists of writing a "set of detailed stories" describing the use of the system; for Requirements Specification, the synthesis consists of enumerating the system requirements in a logical, understandable document. From one perspective, they correspond (respectively) to a top-down (deductive) and bottom-up (inductive) approach to requirements elucidation. While typically only one or the other are used, employing both in parallel greatly contributes to achieving increased consistency and completeness [25, pgs 153, 351]. The same paradigm is used elsewhere (e.g., *reliability engineering* – fault tree analysis vs. failure mode effects analysis, but see [26, 27] on fault detectability; *physics* – thermodynamics vs. statistical mechanics [28, pg 9-17]; *psychology* – cognitive vs. behavioral [29, pg 38]).

Some Practical Validation Examples

HF&E requirements are formulated ideally during the development process. However, absent a complete and correct set of user and use requirements, these requirements must be formulated prior to the deployment validation studies used to establish installation, operation, and process validation and may be used for the periodic re-validations that ensure continuing system integrity.

The validation engineering activity is not unlike the quality engineering activities of system, parameter, and tolerance design [30, pg 536] optimizing a process (during development) and monitoring the process (during deployment and continuing operations). In both cases, statistically designed experiments can facilitate understanding the relationship among controlled (design) factors, uncontrolled (noise) factors, and the desired output(s) (Figure 4). We have two types of independent variables. The factors controlled by the designers (derived from the specifications) include both the system components and non-operator system inputs. The uncontrolled factors include both **system external** environmental factors (e.g., temperature, humidity, vibration, illumination, etc.) that may influence system operation and **human internal** environmental factors (e.g., fear, boredom, anger, fatigue, etc.) that may influence operator behavior. The dependent variable(s) is (are) the system output(s) that must be validated against the requirements. It is to these outputs that we apply the statistically designed experiments.

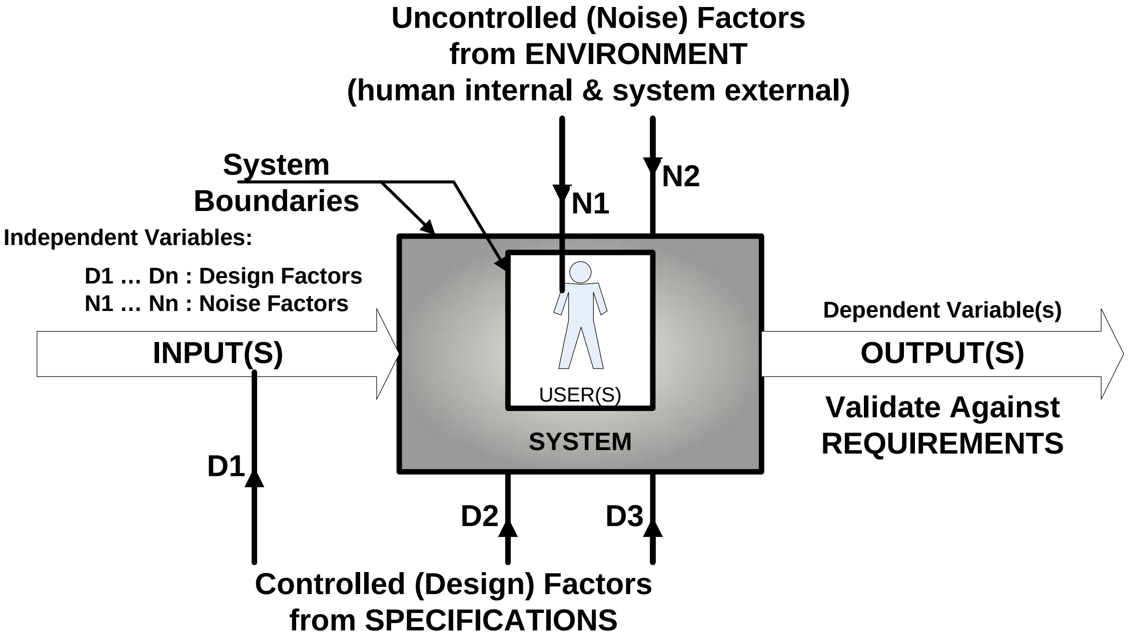

Figure 4: A Simplified Process Diagram

The Taguchi simplification is based on the assumption that there is no aliasing (confounding) between design factors and noise factors (please refer to Figure 4); if this holds true, the experimental procedures may be significantly simplified [31, pg. 101], dramatically reducing the requisite number of measurements. Further reductions in experimental complexity may be accomplished using Shainin's diagnostic tools [31, pg 67-74].

Regardless of whether the design factors and the noise factors are orthogonal, statistically designed experiments are preferred to the one variable at a time approach used historically by many engineers in industry [32, pg 93]. The experimental techniques are well known to the quality engineering and experimental psychology communities. Not only are they more economical, but also constructed properly, they more readily identify optimality conditions and sources of variability. In the development phase, their use in iterative validations will help reveal relationships among and between controlled and uncontrolled variables, as well as helping identify sources of output variability. In the deployment and operations phase, their use in re-validations has been more frequent, presumably due to the modern training of quality engineers.

Let us consider two validation examples: an intermediate validation during the development of a medical device (an infusion pump) and a re-validation during the use of pharmaceutical manufacturing equipment (producing tablets). The examples have been intentionally altered to mask their origins.

A Simplified Medical Device Example

Consider the development of an infusion pump. A general outline of how we might approach a portion of an intermediate validation (some iteration during the development process; please refer to Figure 1) of our requirement derived from the simplified example NWD "*The system shall be easy to use*" would be as follows.

For purposes of this example, let us assume that the "system" is a single channel infusion pump for administration of intravenous fluids or medicines. An extended hierarchical task analysis [33] draws our attention to one particular set of pump set-up tasks. Naïve subjects are used (experience with infusion pumps, but no previous experience with this particular system); a maximum acceptable time (Tmax) for correct task completion is chosen as the output threshold. A 12-run Plackett-Burman [34] screening experiment with two replicates (used to identify main factors, while ignoring confounding) identifies five main factors. The design-controlled factors are display contrast, number of button pushes and the presence vs. absence of a flip card "crib sheet" with step-by-step instructions; the uncontrolled factors are syringe dimension and beginning vs. end of 12-hour shift. Assuming that the controlled and uncontrolled factors are independent, an orthogonal array [35] design is acceptable. Assuming that the effects are linear, a 2-level study of the five factors may be used (32 runs); the dependent output is the time (T) necessary to complete the task correctly. Pareto analysis ("separating the vital few from the trivial many", [30, pg 17]) finds that the factors and interaction effects that have most influence on the output are syringe size and the display contrast.

The result of the observed failures (T >Tmax), in this intermediate validation, is that two amendments to the existing system requirements are added for the next iteration:

- *the system liquid crystal display contrast shall have a user adjustment; the adjustment shall be large, prominent, and adjacent to the display on the front panel OR a high contrast, backlit display will be employed;*
- *the system shall calibrate the syringe output volume, for a plunger step, prior to each infusion OR generic (un-calibrated) syringes will not fit the pump.*

Note that these are human factors-related system-focused requirements, not user-focused requirements formulated as:

- *The operator must be able to discriminate the displayed information at a variety of ambient light levels;*
- *The operator must use only calibrated syringes for infusion of intravenous fluids and medications.*

It is crucial that an ergonomist involved with the product development effort provide the design team system-focused, not user-focused, requirements. System-focused requirements inform designers; user-focused requirements only inform other ergonomists. It is generally understood that infusion pumps with HF&E design flaws have resulted in injuries [36], recalls (e.g., ECRI, ISMP), federal fines (e.g., FDA), and significant corporate financial losses. It is considerable less expensive to include the requisite HF&E requirements formulation and validation.

A Simplified Pharmaceutical Manufacturing Example

Now consider a very different type of validation problem; the manufacture of tablets. The process involves grinding raw materials, bulk mixing of powders, tableting powder, and packaging. Healthy young men and women with minimal education were taught the manufacturing process "on the job" by senior, experienced equipment operators. Operators functioned on standard 12 hour shifts. This was a brand new facility; a major expansion to meet increased demand. In order to avoid problems, the current manufacturing equipment suite was duplicated; identical equipment from the same manufacturer was placed in the new facility. The validation team, working with two senior operators, successfully completed the installation qualification (IQ), the operational qualification (OQ) and the process validation (PV); the suite of equipment was released to manufacturing in record time. In order to minimize operational problems, it was decided that about half the operators from the old facility would be moved to the new facility and both facilities would hire and train new operators to meet the full complement.

The manufacturing process was started and, very shortly thereafter, the quality group began to have serious headaches. Sampling indicated that there were two families of product quality – one well-within the specifications and one just outside the specifications – and both populations with relatively little variability. The validation team successfully repeated the PV, but once released, the problem arose again with unpredictable periodicity. The senior operators spent time monitoring the work of the inexperienced operators and even spent some time monitoring the work of the experienced operators – to no avail.

External help was secured and the IQ, OQ, and both PVs were reviewed in detail, but with no indication of a problem! Discussions were held with the senior operators who reported that they did not see how it could possibly be an "operator error". Nevertheless, the results suggested some type of subtle operator-related error. A study was designed to correlate operator and equipment behavior with product quality. Briefing sessions were held with both manufacturing shifts to describe the experiment and the reasons for conducting the study. Video equipment was set up at each workstation (overt behavioral measurements, as previously described), with the field of view only including the manufacturing equipment – so that the operators did not feel that they were being spied upon. The study was conducted until four incidents of unacceptable product were produced, the study was stopped, and the video data analyzed to identify any possible correlations between product quality and operator behavior. The only correlation between product quality and operator behavior was that the sequence of control panel actions differed. But, the senior operators insisted that this was not an issue and occurred routinely in the older facility – this particular equipment was insensitive to the operational sequence.

After considerable discussion (primarily related to manufacturing delays, consultant costs, and personnel time), it was decided to take the new facility off line and validate the operational sequences – validation of one aspect of human use. There were three controls involved, so the permutation (sequence matters) permits six (3!=6) unique operational sequences, which require a six run screening experiment (product quality was measured as usual). The result was that all, but two, sequences produced acceptable product. The equipment vendor was contacted and, ultimately, it was discovered that a minor software "improvement" had been made, but had not been validated internally or identified with a new revision number. The combination of the software change, and all operators "knowing" that the control panel sequence did not matter, resulted in the two different product quality populations! It is still unknown why both senior operators always followed the same identical procedure, thus permitting the two PVs to succeed.

As a result of this difficult experience, a series of procedural changes were instituted in the standard operating procedures for the facility engineering and the validation groups. The most important change for the validation group was the inclusion of training/use validations that included documenting required user procedures and studying non-standard "operational sequences" with actual operator observations; in essence, apply Figure 1 in the deployment phase. In this simplified example, no detailed operator training requirements had been established, no formal operating procedures had been validated, and an unanticipated use error – precipitated by a software flaw – resulted in new facility launch delays, inability to meet production quotas, and undetermined corporate financial losses.

Limitations of Validation

If the validation engineering is properly conducted during the development phase, the post-deployment (re-)validations are mere checks of system integrity; if not, then the post-development validation activities become far more complex. Reliability is a necessary, but not sufficient, condition for validity [37]. If the system is unreliable, validation is difficult, if not impossible.

There exist two very important limitations to engineering validation related to missing or defective requirements. First, if a requirement is absent (a latent failure [38, pg 173, 208] – a hole in Figure 5 [39]), the system will incorrectly pass the validation. In the graphical example (Figure 5), routine use of standardized specifications obfuscate the existence of the missing requirement and block the hazardous state; when the specification is subsequently changed (such as during a manufacturing quality engineering optimization), the unanticipated hazardous state “unexpectedly” appears. A means of minimizing this is the iterative risk analysis; however, ruthless enforcement of Requirements Engineering is fundamentally the best approach. Second, “drift” [40, pg 35-37] during manufacturing or post-deployment maintenance will expose unanticipated hazards that may not be susceptible to traditional validation studies. An example of this might be a variation in maintenance either that was not envisioned (a missing requirement) or that was outside of the control of some of the stakeholders (a defective requirement e.g., [40, pg 31-33]. Such unexpected events (e.g., Hazards #1 and #3 in Figure 5) are typically not accessible during validation of the system implementation; they may only be accessible during the periodic re-validations (system integrity checks) and this will be highly dependent on the design of the re-validation protocol. Once again, the iterative risk analysis and ruthless enforcement of Requirements Engineering are indicated, especially in the case of re-validation protocols.

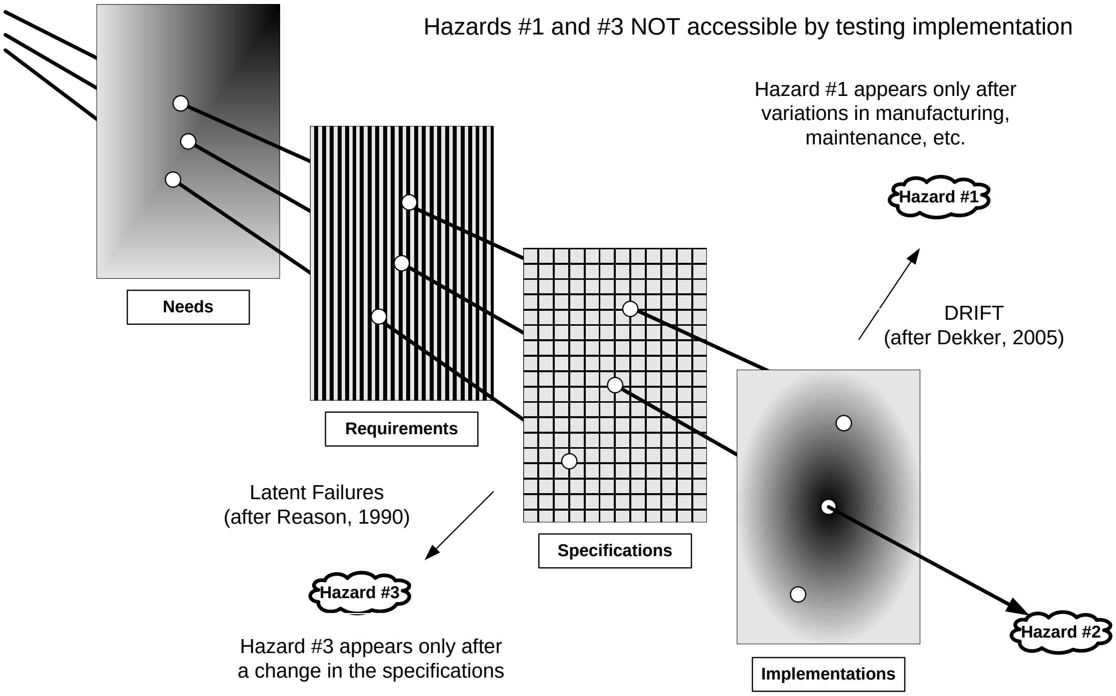

Figure 5: Hazards due to Latent Failures and Drift

Conclusion

We have reviewed some fundamentals of human factors and related measurements, discussed an approach to formulating user and use requirements that specify quantitative validation and then considered two types of human factors validations: validation of equipment design and validation of equipment use. In both cases, a simplified example (a medical device and a pharmaceutical manufacturing system) was used to illustrate the approach. We concluded by identifying some of the limitations of empirical validation. Proper application of HF&E will permit rational and cost-effective validation of equipment use and the user's training, the work structure, and the work environment. Proper consideration of HF&E issues is crucial for proper implementation of Process Analytical Technology approaches that attempt process understanding, process validation, process improvement and process optimization.

References

---

1 Cronbach LJ & Meehl PE. Construct validity in psychological tests. Psych. Bull, 52:281-302. 1955

2 Hall AD. Systems Engineering from an Engineering Viewpoint. IEEE T-SSC. 1:4-8. 1965

3 Simon HA. Models of Man: Social and Rational. New York: Wiley. 1957

4 Popper KR. Conjectures and Refutations: The Growth of Scientific Knowledge. London: Routledge. (1963/2005).

5 Samaras GM & Horst RL. A Systems Engineering Perspective on the Human-Centered Design of Health Information Systems. Journal of Biomedical Informatics 38:61-74. 2005

6 Samaras GM. Engineering Complex Systems: Validating the Human Factors. Proc. 7th Annual Symposium on Human Interaction with Complex Systems. Greenbelt, MD. November 17-18, 2005, *in press*.

7 Nagamachi M. Kansei Engineering. Tokyo: Kaibundo. 1989

8 21 CFR 820.250 Statistical Techniques. 1997.

9 Human-centered Design Processes for Interactive Systems. International Standards Organization, ISO Standard 13407:1999.

10 Kroemer KHE. Engineering Anthropometry. in Handbook of Human Factors and Ergonomics, 2nd Edition, (G. Salvendy, Ed). New York: Wiley-Interscience. 1997.

11 Chaffin DB & Andersson GBJ. Occupational Biomechanics, 2nd Edition. New York: Wiley-Interscience. 1991.

12 Rasmussen J, Pejtersen AM, Goodstein LP. Cognitive Systems Engineering. New York: Wiley. 1994.

13 Vicente KJ. Cognitive work analysis: toward safe, productive, and healthy computer-based work. Mahwah, NJ: Lawrence Erlbaum Associates.1999.

14 Hollnagel E (Ed.). Handbook of cognitive task design. Mahwah, NJ: Lawrence Erlbaum Associates. 2003.

15 Hollnagel E & Woods DD. Joint cognitive systems: foundations of cognitive systems engineering. Boca Raton, FL: Talyor & Francis/CRC Press. 2005

16 Eggleston RG, Roth E, Whitaker R, & Scott R. Conveying Work-centered design specifications to the Software Designer: A retrospective analysis. Proc. Human Factors and Ergonomics Society 49th Annual Meeting, Orlando, FL. pgs. 332-336. 2005

17 Lintern G. (2005) Integration of cognitive requirements into system design. Proc. Human Factors and Ergonomics Society 49th Annual Meeting, Orlando, FL. pgs. 239-243.

18 Woods DD. (2005) Generic support requirements of cognitive work: Laws that govern cognitive work action. Proc. Human Factors and Ergonomics Society 49th Annual Meeting, Orlando, FL. pgs.317-321.

19 Chapanis A. Human factors in systems engineering. New York: Wiley. 1996.

20 Beaney M. "Analysis", *The Stanford Encyclopedia of Philosophy (Summer 2003 Edition)*, Edward N. Zalta (ed.), Available at: http://plato.stanford.edu/archives/sum2003/entries/analysis/. Accessed October 19, 2005.

---

21 Dowell J & Long J. Conception of the engineering design problem. Ergonomics, 41(2): 126-139. 1998.
22 Kossiakoff A., & Sweet Wm.N. Systems engineering principles and practice. New Jersey: Wiley-Interscience. 2003.
23 Samaras GM. Validation engineering in ergonomics: Theoretical perspectives. Proc. Human Factors and Ergonomics Society 47th Annual Meeting, Denver, CO USA. pgs.1454-1457. 2003.
24 Madu C.N. House of quality in a minute. Fairfield, CT: Chi Publishers. 1999.
25 Lewis C.I. & Langford, C.H. Symbolic Logic. New York: Dover.1932
26 Del Gobbo D, Cukic B, Napolitano R, & Easterbrook, S. Fault Detectability for Requirements Validation of Fault Tolerant Systems. Proc. 4th IEEE International Symposium on High Assurance Systems Engineering. 17-19 Nov. 1999. p 231-238
27 Schmidt, MW. The Use and Misuse of FMEA in Risk Analysis. MD&DI. 26(3):56-61. 2004
28 Pathria RK. Statistical Mechanics. Oxford: Pergamon Press. 1972.
29 Grossberg S. How does a brain build a cognitive code. in: Studies of Mind and Brain. Dordrecht, Holland: D. Reidel Publishing Co. 1982.
30 Pyzdek T. Quality Engineering Handbook. Tucson: QA Publishing. 2003.
31 Bhote KR. World Class Quality. New York: AMA Membership Publications Division. 1988.
32 Antony J. Design of experiments for engineers and scientists. Amsterdam: Butterworth-Heinemann. 2003.
[33] Chung PH, Zhang J, Johnson TR & Patel VL. An extended hierarchical task analysis for error prediction in medical devices. Proceedings of AMIA 2003, pgs. 165-169. 2003.
[34] NIST/SEMATECH. Plackett-Burman Designs, NIST/SEMATECH e-Handbook of Statistical Methods, Available at: www.itl.nist.gov/div898/handbook/pri/section3/pri335.htm Accessed on March 2, 2006.
[35] NIST/SEMATECH.What are Taguchi Designs? NIST/SEMATECH e-Handbook of Statistical Methods, Available at: www.itl.nist.gov/div898/handbook/pri/section5/pri56.htm Accessed on March 2, 2006.
36 Vicente KJ, Kada-Bekhaled K, Hillel G, Cassano A, Orser BA. Programming errors contribute to death from patient-controlled analgesia: case report and estimate of probability. Can J Anesth 50: 328-32. 2003.
37 Pedhazur EJ & Schmelkin LP. Measurement, design, and analysis: An integrated approach. Hillsdale, NJ: Lawrence Erlbaum Associates. 1991.
38 Reason J. Human Error. Cambridge: Cambridge University Press. 1990.
39 Samaras, GM. Systems Engineering for the Human Factors Engineer: A Workshop. 16th World Congress on Ergonomics. Maastricht, NL. July 10-14, 2006. *in press*.
40 Dekker SWA. Ten questions about human error: A new view of human factors and system safety. New Jersey: Lawrence Erlbaum Associates. 2005.