
Revisiting Usability's Three Key Principles

Gilbert Cockton

School of Computing and
Technology, University of
Sunderland, St. Peter's Way,
Sunderland SR6 0DD, UK
Gilbert.Cockton@sunderland.ac.uk

Abstract

The foundations of much HCI research and practice were elaborated over 20 years ago as three key principles by Gould and Lewis [7]: early focus on users and tasks; empirical measurement; and iterative design. Close reading of this seminal paper and subsequent versions indicates that these principles evolved, and that success in establishing them within software development involved a heady mix of power and destiny. As HCI's fourth decade approaches, we re-examine the origins and status of Gould and Lewis' principles, and argue that is time to move on, not least because the role of the principles in reported case studies is unconvincing. Few, if any, examples of successful application of the first or second principles are offered, and examples of the third tell us little about the nature of successful iteration. More credible, better grounded and more appropriate principles are needed. We need not so much to start again, but to start for the first time, and argue from first principles for apt principles for designing.

Keywords

Software Development Principles, Usability, Evaluation, Design, User Studies.

ACM Classification Keywords

ACM: H.1.2 – User/Machine System

Copyright is held by the author/owner(s).

CHI 2008, April 5 – April 10, 2008, Florence, Italy

ACM 978-1-60558-012-8/08/04.

INTRODUCTION

In 1985, ACM's premier journal CACM published a seminal article, *Designing for usability: Key principles and what designers think* [7], which recommended key principles that still underpin standards by which HCI work is judged. These principles have been revisited twice in each edition of Helander's Handbook of HCI [8, 9], with the claim in 1997 that the principles have "stood up to the passage of time" [9, p. 232]. The original article [7] is most accessible, most cited, and thus better known. Nothing in the Handbook chapters significantly revises the original three principles. A fourth of Integrated Design was added in [8], and addresses the scope of usability work across a range of materials and resources beyond the software user interface. These had already been covered in [7], but with no explicit associated principle. Interestingly [7] does refer on page 301 to *four* principles, suggesting that a fourth principle was edited out. Beyond parading this detective work, the fourth principle will not be discussed here, as it has had less impact on evaluation of HCI practice and research. As John Gould was the only common author across the three key publications [7, 8, 9], combined authors are referred to as *Gould and colleagues* from now on.

Note that the "what designers think" in the title of [7] refers to voluntary attendees responses to surveys at human factors talks and classes at IBM. Reported survey results indicated that attendees' knowledge and understanding of the three principles was limited and often erroneous [7].

This paper presents two interleaved arguments. Firstly, Gould and colleagues misattribute their successes to their principles as they presented them. The principles are post-hoc and applied retrospectively, without reflective self-criticism that could identify their real relative roles in the reported examples of practice. Secondly, the principles no longer embody

contemporary HCI values. Surprisingly, Gould and colleagues failed to realise at the time that their successes were due far more to basic design practices than to initial user research or summative empirical evaluation.

To prepare for these two arguments, the three principles are now individually summarised and critiqued, keeping as close as possible to Gould and colleagues' texts.

Early Focus on Users and Tasks

This first key principle required direct study of the cognitive, emotional, behavioural, anthropometric and attitudinal characteristics of users, as well as the nature of the work expected to be accomplished. The aim is to understand, not to identify, describe, stereotype or ascertain. Direct contact is necessary through interviews, discussions, observations and training by users, ideally institutionalised "from the very outset" via a user panel within the design team, to establish how people go about their work and what their problems are. Gould and colleagues however never defined what "from the very outset" actually meant. What it clearly did not mean from their examples was "before any relevant technology has been invented, incorporated into demonstrators or deployed in a usage setting".

An initial design phase must precede any design iteration. Its purpose is to collect user profiles, and to specify tasks, usage contexts, and measurements of interest. All must be collected for specific development contexts, since user interfaces add "a co-processor of largely unpredictable behaviour" (i.e., a user) and "there is no data sheet for this co-processor ... one does not know what one needs to know about a user until one sees the user in person".

The principle was renamed in [9] to *Early – and continual – focus on users*. This was linked in a move from a two to four phase development method, supported by customer experts on the design team and an additional preliminary ‘gearing-up phase’ that advocated preparatory desk research. Some additional methods and a checklist were added. However, the overall philosophy of ‘just finding out’ remained, with only task analysis providing any formal or theoretical structure. Note though that “and tasks” disappeared from the first principle’s name in [9]. Also, the addition of “– and continual –” suggests further reflection on what “early” meant, and whether it was enough. In some ways “sustained” communicates what is needed without the timing issues raised by “early and continual”. It also avoids asking for the impossible, i.e., asking that the human factors or user experience people get in first. They almost never will. Invention always precedes innovation. There will generally be a well developed technology before user experience design can become a concern. Even when envisionment methods such as video, drama or other forms of scenarios are used, the development of new technologies almost always proceeds in parallel, largely uninfluenced by any insights from searching and innovative exploratory design approaches. Although ideal worlds may be possible, in practice HCI methods must be able to cope with the most common situation of technologists getting there first.

Every word in this first principle is open to challenge. *Users* as such may not yet exist, thus *envisaged users* is a more accurate refinement. Lastly, no example of a *focus* is provided in the examples in [7] or [9]. There are vague lists of the sort of things that you could look at, but nothing as tight and reasoned as the approach to focus that was to develop within contextual design [20].

Empirical Measurement

This second key principle required performance of real work to be observed, analysed and recorded, including use of simulations and prototypes. This entails “actual behavioural measurements of learnability and usability”, giving users simple test tasks to allow performance, thoughts and attitudes to be recorded and analysed. Several user testing approaches are possible: pencil and paper tests, video analysis of observed usage, memorisation/recall tests, formal experiments and even demonstrations to visitors and at technical meetings. Note though that the principle is not about just watching, or a more involved recording and analysing, but states that things will be *measured*, a stronger requirement than seeing how people get on with a design.

To be able to collect measurements, the initial design phase should identify measures of interest such as learning time, error rates, or help request counts. Criterion values should be set for them all [7]. These values can be relative to own or competitor products. In [9], testing was extended to reliability. New approaches included on-line communities and long term demonstrators (‘Hallway and Storefront Methodology’). No examples of project measurements however are given for such extensions and new approaches.

Iterative Design

Measurement and user testing (but which?) inevitably finds problems that must be fixed and re-tested “as often as necessary”. Scenarios, prospective user manuals (written in advance of development), prototypes, and simulations all let users *react* to radical designs for novel purposes. Note again, that *reactions* are not necessarily measurements. Simulations or prototypes are “something tangible to use” that stimulate “thought and ideas”. Design from first principles is impossible, so “an empirical approach is essential”: “Designers will have to

make many choices on their own, and be prepared to test their work empirically" [7]. It is interesting how designers' choices do not extend to whether or not to 'test' empirically, or whether to focus early and continually on users (and perhaps tasks). Empirical measurement and a focus on users are advocated principles, without supporting 'first principles' to guide the choice of, e.g., a field study focus or choice of measurements.

Some 'first principles' were however offered for guiding iteration. Design goals such as "user friendly", "easy to operate", "simple", "responsive" and "flexible" are very difficult to reach. A process is required to *ensure* meeting these goals, and that process is an iterative one. And yet among the surveyed 'designers' "almost no-one mentioned establishing testable behavioural specification" [7].

If criterion values cannot be relative to existing products, proposed targets must be iterated alongside the design until appropriate ones are set. Equally or more challenging is finding a remedy when behavioural targets are not met: "this is usually a very tough problem". Indeed, evaluation ("*whether or not* [targets] ... are being met") is not the same as understanding ("*why* [targets] ... are not being met"). The latter is better supported by research instruments such as think aloud. User comments say more than stop watches. This desire for high quality problem explanation indicates that "the need to iterate is not a licence to be sloppy". However, other than using think aloud as an 'explanation' instrument, there was little advice on avoiding explanatory sloppiness. Note though, that the purpose of think alouds appears to be solely to understand why targets are not being met.

A CRITIQUE OF THE THREE KEY PRINCIPLES

There can be no doubt that the values and practices associated with the above three principles have been immensely influential and frequently effective within HCI, but after over 20 years, we should be able to raise our game. Indeed, HCI colleagues in research and practice have raised their game over the intervening decades, but we have not seriously revisited these principles to assess their currency. Some issues with each principle have been raised above. This section critiques the principles as a group.

One must thus ask at what times the principles actually took the 'test of time'? More reflective disciplines may have been better placed to see what Gould and colleagues really did, which was identifying valuable outcomes and experiences [4] and then providing them. Usability, despite all their claims, constituted a very small proportion of this value. At best, it made the systems worthwhile by reducing usage costs.

The Emergent Approach

The key principles emerged within the long running ADS project (a speech filing application, aka voicemail) [7]. Some of the overall approach was sharpened as early as 1979 within an IBM study group on human factors. However, "the principled type of thinking" had not been developed for ADS' earliest phases, but emerged within an evolving iterative process. Today in HCI a more rational and systematic approach to methodology may be possible, but then then this would have been politically inferior to a grab for critical software development activities, driven by deep convictions about user-centredness. Even so, this invalidates use of ADS as an example in [7], since there could not have been an early focus on key principles here. The examples in [9] are thus a better basis for assessing the impact of principles from the outset of a development project.

Examples apart, the emergent nature of the principles in [7] calls them into question, since they are essentially post hoc generalisations that were influenced by a range of contemporary, organisational and disciplinary values.

Contemporary and Organisational Values

The three principles remain products of their time that had more to do with reining in technological utopianism than following what is currently called design thinking [3]. The tactical outcomes of organisational manoeuvres, the principles were radical battle moves more than reasoned argument. Usability's day would come, but would come quicker with the right organisational redesign. Although Gould and Lewis [7] denied intent to single out "bad folks", and acknowledge that there are alternatives, they pulled no punches against hapless heretics who resisted surrendering "real control of the user interface to the people who had responsibility for the user interface".

There is a palpable crystal ball behind the rhetoric on impediments to the three principles: "competitive necessity will eventually break down these obstacles and traditions" [7]. Liberal economics was part of a usability attack alongside pragmatic liberalism humanism and cut-down pint-pot psychology. With Gould and colleagues so confident of usability's destiny, what use would an attempt at rational derivation of principles serve? Far better to see if destiny can be accelerated through the corporate manoeuvre of establishing a single user interface group to take control of the whole user interface: process, software, manuals, training procedures and deployment workstations [7].

When introducing the three principles above, logical flaws were made already apparent, but these were irrelevant at the time. The principles acted as values for an organisational land grab

for usability within IBM. Similar grabs occurred in Bellcore and Digital, and then Hewlett Packard, BT, Lotus, Sun, Ameritech and Microsoft later in the 1980s, followed in the 1990s by SAP, Philips, Siemens, Oracle and many others. As an ideology, the three principles have secured their place in history. However, their roots were in disciplinary values that no longer dominate HCI. We need new principles that better reflect HCI's current disciplinary balance. For one, reflection, reflexivity and argument need to be taken far more seriously.

There's No Arguing with Not Arguing

Gould and colleagues hold to their three principles in the face of either inconsistent or incompatible self-reported evidence. The principles' emergent nature, coupled with disciplinary preferences, failed to promote their real key discoveries to key principles. A psychologist's preference for facts over argument and reflection is clear, but although facts may speak for themselves, not everyone can always listen.

The aside that "Principles of design are arguable, of course" [7] could be a concession or an oversight. A concession here would follow from liberalism: everyone is entitled to their opinion and you are free to argue. An oversight here would follow from feeling unable to *argue for* principles. With psychology's prejudices towards its abandoned philosophical parent, no argument would be attempted. As the rejection of *introspection* as a research method split psychology off from philosophy in the nineteenth century, *argument* was marginalised in the process. As facts speak for themselves, logical or conceptual assistance is unnecessary. Also, facts keep changing: "ours are not universal truths" [7]. Are there really no universal truths about design or evaluation? If not, what was it then that made these three principles so 'key'?

The surveyed designers just didn't get these 'key' principles. Fancy that! Five categories of reasons explained this. Some would "be resolved one way or the other as the recommendations are more fully tested in practice" [7], but even now such approaches are still rarely fully institutionalised in software development. Perhaps more attempts at argument may have speeded up the principles' spread. Equally however, taking argument seriously may have revealed that the self proclaimed keys couldn't actually unlock all that was claimed for them. Cut down psychology wasn't enough.

Disciplinary Values

The three principles reflected both the psychological backgrounds of key protagonists and also a common sense recognition of limits of applicable psychology. The resulting intellectual vacuum was largely filled through pragmatic common sense, a self-limiting conflation of diagnosis and treatment: the sound common sense diagnosis was that cognitive psychology was no basis for analytical predictive right-first-time design; the less imaginative treatment was to apply common sense in understanding users and usage difficulties, as well as in generating (re-)design options. Such common sense tactics left no open research questions and few deep practitioner dilemmas (challenging yes, but they had answers). The impression is of a development process without need for further improvement, especially given use of words such as "ensure" and "assure" [7], which might be read as the last words on HCI, despite denials elsewhere of universal truths. 12 years on [9, p.240], its status was unchanged: it had "stood the test of time", remaining essential, and with many proud advocates.

Gould and Lewis reassure a key audience that usability methodologies "are sufficiently rigorous and conform to the traditional scientific approach" [7]. Support here is almost

wholly from cognitive psychology. Emotions and attitudes get a mention each, but work is something that people "go about" and "have problems" with. Knowledge has to be "played against the tasks". However, Gould and Lewis [7] anticipated disappointment from psychological peers through their failure to endorse predictive methods from then current cognitive psychology. Cognition is not everything though: "understanding the user becomes all of psychology (cognitive, behavioural, anthropometric, attitudinal etc.)" Presumably developmental, organisational, motivational and social psychology can be assumed, with others, to be part of the 'etc.', but not sociology, philosophy, management, human geography or economics? Despite this recognition of a need for general psychology (rather than an inter-disciplinary team [3]), even 12 years after [7], all psychology examples in [9] concern ease of learning, use, errors and difficulties.

Designing for usability was applied psychology. Other disciplines' values such as "the power of reason", "the power of technology to succeed" are given short shrift, even though careful planning is celebrated for its analytical rigour, and the table driven user interface software for ADS was seen as vital to its success (as was its ultimate successor ITS in [9]). Tensions with other development disciplines are palpable. The glare of psychological empiricism blinded Gould and co-authors to their own dependence on both reason and technology. Again, the post hoc nature of the three principles allowed disciplinary ideals to obliterate the reality of successes with ADS and ITS-based systems. Technological power was critical for effective iteration. Similarly, the initial use of hallway and storefront methodologies in [9] was rationally derived from understandings of local work practices and human circulation within an R&D building.

Trying to Fill a Theory Gap with Common Sense Humanism

Gaps in psychological knowledge were not filled by knowledge from other relevant disciplines, but by an informal humanism that celebrated individual autonomy (except where controlling the user interface process is concerned). There was humility in acknowledgements of scientific limitations: "the most important lesson is the unpredictability of good design" [7]. A pragmatic approach to doing the best you can within a well thought out process was the best way to face such challenges.

Common sense plugged gaps left where cognitive psychology couldn't help. Often it did help, with ADS achieving "zero training" success through self-explanation. Ease of use and learning had been set as demanding design goals (no measurable criteria are given as examples however). Both goals were met. Even so, basic understandings of human memory could have led the design team to add a feature similar to the Pending Message Box "to remind the sender and recipient that an action is needed" [7]. Instead, this was a result of a user suggestion during testing. This is fine as a back stop, but with hindsight perhaps this could have been right-first-time with more searching reflection on action. I still have to tag and label email messages myself to gain such functionality for my Inbox, so this insight did not spread to similar communication applications. To IBM's credit, they did implement it when it was suggested, but even so, some psychological foresight really was possible for this example.

Another failure of common sense is seen in an initial oversight that one's own messages are different to those of others when editing/annotating. The overlooked Self/Other distinction is fundamental in social psychology, and much philosophy [15], theology and psychiatry, so there should have been no surprises: late fundamental changes to backend software were avoidable. A lack of theoretical receptiveness here may have

restricted the designers' ability to see what was in front of them as they watched executives at work. Ethnographic values of "rendering strange" the everyday are better disposed to revealing the profound beneath the mundane [5]. In editing *our own* voice communications, we seek to present ourselves socially and prepare a performance (albeit asynchronously, remotely and potentially impolitely). In annotating *your* voice messages, we organise *our work*. Different motivations apply, but neither volition nor the presentation of self in everyday life is mentioned in [7]. Cognition and attitude-related emotions would do, leaving the other 95% (approximately!) of human agency to common sense, and GREAT [sic] systems to the virtues of the project team and its leader [9, p.240].

If Only They'd Known Those Prototypes Were Probes!

In Molière's *Middle Class Gentleman*, Monsieur Jordain had, for more than forty years, been speaking prose without knowing anything about it. Similarly, Gould and colleagues used *probes* in blissful ignorance. ADS [7] was originally to be a remote dictation system to let executives create memos while travelling, but secretaries could be bypassed if needed to effect a voice mail function: "only after a prototype was in use was it determined that the spoken message communication features... were the really useful ones" [7]. Unless logging is built in, fielded prototypes are not associated with empirical measurement, but are unsupervised and unobserved. With no mention of logging, fielded use was an early example of a technology probe [12] that revealed how users would really use a technology, as opposed to what designers believed they would. Discoveries are, but their near diametric opposition of celebrated to planned scheduled observed behavioural measurement escapes comment. One critical design iteration for ADS thus came from a user suggestion and another from appropriative probe usage, and not to behavioural targets, that

"management tool to assure that system development proceeds properly" [7].

The most recent defence [9] of the principles further undermines the centre stage positioning of empirical 'measurement' and behavioural specifications in the first two of Gould and Lewis's acclaimed principles. There is not one example of measurement leading to good, never mind GREAT systems. Almost all examples address functionality ('empirically determined required improvements ... adding functions' p.245), with the rest appropriating un-pre-specified emotional design achievements. Thus the Illinois system for the unemployed is actually an example of worth-centred [4], rather than user-centred design, by increasing the value to clients while reducing the costs to both them and support staff. The resulting "more dignified way for citizens and staff to interact" (p. 248) is associated with no behavioural specifications or empirical measures of dignity. Dignity may not even initially have been in focus. Success was related to other principles that Gould and colleagues never exposed.

IBM's teams stumbled on probes and stayed there: "we do most of our observations during field studies" [9, p.250], i.e., not in controlled lab studies with measurements against target behavioural specifications. Features such as registering for interviews in the Illinois system were added after *watching* users, as were supportive on-line claims forms. 75% of development effort occurred after field installation, hence *early* focus became *continual* [9]. The principles stood the test of time through endless resits.

THE KEY PRINCIPLES AND CURRENT HCI

Turning to the present, we can see several breaches of current values in HCI research and practice. These explicit design principles are carriers of HCI values, so we should often revisit

Gould and colleagues' principles. Not surprisingly, after 11-23 years much of the principles' presentation and defence is out of step with current HCI. Key subsequent HCI developments are neither anticipated in [7] nor acknowledged in [9]: the turn to the social and theory, the expansion of disciplinary inputs and perspectives, recognition of organisational differences in requirements and capabilities, the expansion of HCI beyond work, and the recent turn to design.

The Rise of Theory

In HCI's second wave, "the turn to the social" [17] replaced the human *factors* at the heart of the key principles with a socially oriented understanding of human *actors*. Theory fared poorly in first wave HCI, but reflective social science approaches brought explicit justified theoretical sensitivities. Second wave HCI was characterised by rising theoretical sophistication, leading to substantial monographs (e.g., [5]) at the cusp of HCI's second and now contested third waves [2]. The third wave turn to human *satisfactors* has hedonic, critical and axiological renderings. Hedonic ones combine affective psychology with pragmatic philosophy and literary theory [e.g., 15]. Critical reflective ones [e.g., 19] bring in moral, ethical and political philosophy. Axiological ones focus on explicit values [6] or more general *worth* [4].

Contenders for HCI's third wave draw on philosophy, much of which deals with problems "requiring conceptual and logical investigation" that cannot "be solved by empirical means" [10], questioning Gould and colleagues' primacy of empirical measurement and data gathering. Third wave approaches move further away from first wave empiricism, continuing the reflection that came with the turn to the social. They bring us up against deeply important philosophical questions of existence, knowledge, truth and value, to which there have

been many approaches (including their rejection [10]). Some may better guide and focus third wave thinking.

Transdisciplinary Tendencies

Gould and Lewis [7] made tentative steps by dropping unproductive cognitive theory, but HCI has yet to make a full move to transdisciplinary approaches. Koskinen and Battarbee boldly reject disciplinary policing of design [13], grounding validity in the *fruitfulness* of design:

Designers do not need to conform to the validity requirements of . . . disciplines: the validity of their work depends on the fruitfulness of their ... design. ... The designers' aim is not to contribute to the conceptual and theoretical development of ... sciences, but to create viable products, concepts and concept portfolios

Gould and colleagues grounded fruitfulness in achievement of behavioural targets. Transdisciplinary approaches admit other groundings: what is fruitful must not be predetermined by rigid principles. Nor, as an aside, must it be hamstrung by intellectual snobbery about using introductory texts such as *Very Short Introductions* [e.g., 14, 15], which I deliberately reference to make ideas accessible. Immediate accessibility for designers has more value than esoterics. Readers who find value in popularised ideas will work towards original material if this proves to be worthwhile, but few will have to reach readings on Bakhtin such as “Этой субъективной игре со временем, этому нарушению элементарных временных соотношений и перспектив соответствует в хронотопе чудесного мира и такая же субъективная игра с пространством, такое же нарушение элементарных пространственных отношений и перспектив” on a long march from the basics of *Technology as Experience* [15]!

Organisational Sensitivity

Gould and Lewis were clearly sensitive to the realities of working within IBM, and alongside a similar team at Digital, they presented HCI principles in a form palatable to software development [20]. The success of IBM and Digital's 1980s Usability Engineering [20] is a good example of adapting methods to specific organisational contexts. However, expectations were for a common destiny for all organisations. Today, seasoned leaders in user experience practice recognise that different organisations have differing needs and capabilities [18]. One set of principles and their associated project management and techniques will not fit all organisations. Interestingly, one of the case studies in [9] was abandoned because the cost of backend extensions was too high. Initial technical feasibility work was based on developer opinion rather than credible technical specifications. This failure to see value, but not what would make it worthwhile, points to a need for a broader view of system worth than that offered by basic usability perspectives [4].

New Frontiers

Work, work, work: that's all Gould and Lewis ever wrote about! For HCI in the 1980s, systems meant work systems. Today, digital media, ubiquity and broadband mobile data communications have taken HCI into almost every imaginable social setting. “Testable behavioural specifications” are simply unacceptable in many of these settings, and impossible in others.

With the EXPO'92 system and other case studies [9] Gould and colleagues' move into non-work settings brought them into contact with laughter, quality of life, empowerment and lingering to learn. Yet their three principles stood the test of time, but no need for new principles for new settings. Phenomenological principles behind technology probes [12]

better explain the achieved design outcomes, doing away with any need for adherence to empirical psychology.

A Growing Design Focus

For some, HCI's earlier turn to the social is being superseded by a *Turn to Design*. Gould and colleagues not only avoided social approaches to work, they also took design for granted. They devalued design approaches such as rationales and guidelines on empirical grounds. The implicit assumption is either that design will just get done anyway, or that there no focus on it is needed, only on users and evaluation. Either way, Gould and colleagues have little positive to say about design. It needs to be delayed by that early focus on users. Once underway, design becomes an object of suspicion. Nor are designers valued: "designers always seem to be in the middle of something – and never at the beginning of something with time to think about global issues" [9, p.241]. The briefest reads of design texts (e.g., [11, 13, 16]) would show this to be more of an exception than a rule.

Gould and colleagues' experience of software design was from the era of structured methods, with its premature and ungrounded commitments to detailed design specifications. Such engineering values persist in [9, p.245], with mention of "optimising an entire user interface style". It is not clear however how such optimisation is achieved given the lack of any example in [7] or [9] of iterating until satisfaction of an empirically measured prespecified behavioural target. However, design thinking [3], with its roots in historical design disciplines, has always been distinct from the engineering design approaches favoured for software development. Gould and Lewis' broad view of who was and who wasn't a designer reflects an indifference to qualities that may distinguish professionally trained designers from engineers and project managers. The 'designers' in their title were a self-selected

group of system planners, programmers, designers and developers who attended human factors talks and classes, but, when surveyed, misunderstood or failed to mention one or more three key principles [7].

Moving from people regarded as designers by human factors experts to design writers, John Heskett states that designs "result from ... decisions ...Choice implies alternatives, in how ends can be achieved, and for whose advantage" [11]. Related to this, when Norman Potter, a previous generation's well respected design educator asked "What is good design?" he answered that "the 'goodness' or 'rightness' of a design cannot be easily estimated outside of a knowledge of its purpose" [16]. Gould and colleagues are silent on the choice of *concrete means* (how ends are achieved through materials, features and qualities) and *ends* (or worthwhile human purpose). They are also silent on generation or consideration of alternatives. "Testable behavioural specifications" specifies success in terms of means rather than ends: designs may perform as required and yet not be fruitful (i.e., usable but useless). The significant shift of ADS purpose from remote memo dictation to voicemail owed nothing to testable behavioural specifications and everything to user appropriation during prototype usage.

SHOULD WE NOW START AGAIN FROM SCRATCH?

After over 20 years, we should be able to find shortcomings in the three key principles. As continuing foundations of HCI, there are gaps, cracks and loose debris. The gaps include a lack of: ways to track and reflect on design purpose; theoretical receptiveness; underpinning transdisciplinary theory to scope such sensitivities; well thought through approaches for non-work settings [12]; and ways to maintain and compare a diverse range of alternative design means [11, 16] (other than experimental optimisation of interaction

parameters). Cracks include: fault lines between underemphasised design purpose and overemphasised evaluation targets (evaluation purpose unmatched to design purpose [4]); and a failure to equitably consider organisational differences in needs and capability. Debris includes: 'measures to go' from cognitive psychology as a surrogate for evaluations focused on design purpose; and an empirical suspicion of rational normative analysis.

Despite this, evaluations in selected CHI papers are moving towards, rather than away from, the three key principles [1], despite well argued criticism of repressed design creativity [14] in the name of principles that may never have been fit for purpose as originally expressed, and in fact were far less applied than was repeatedly claimed. A more credible account of case studies in [7,8,9] is that users were *consulted*, *observed* and *listened to* in a range of settings.

What can be fairly claimed as key principles from the reported work are much simpler, i.e., *ask*, *watch*, *listen* and *fix*. Measures never cut it in any example. It is interesting to compare what stands up in [7,8,9] with IDEO CEO Tim Brown's five-point model for strategizing by design [3]:

1. Hit the Streets
2. Recruit T-shaped people
3. Build to Think
4. The Prototype tells a story
5. Design is Never Done

Comparing the balance and emphasis in this *design thinking* with [7], there is more on design, but little on evaluation (3?) or measurement (4?, 5?). Still, these 'points' are as emergent as in [7]. As such, they too could marry up poorly against the

concrete examples that inspired them. HCI needs to move beyond such post-hoc and unevenly grounded principle formation to something more systematic. This must be initially focused on design purpose, and only subsequently on process means of methods, techniques and tools. No single academic discipline's values must ever dominate.

Post-hoc principles often do not reapply convincingly to the examples from which they are derived. An alternative approach would be to start with a normative position on the universal purpose of all design. By taking a normative stance, we escape the impossibility of establishing a definition of design that always *will hold*. Instead, we adopt a position that always *should hold*. Designing is doing, and as such, is inescapably ethical and subjective.

Once such a position is declared, we can rationally derive principles for *designing*, and use reflexive self-critique to test whether they are necessary and sufficient, moving principles beyond arguable to *argued*. My attempts at a derivation have concluded that designing should be *committed*, *receptive*, *expressive*, *inclusive*, *credible* and *improvable*.

Six worth-centred principles (i.e., commitment, receptiveness, expressiveness, inclusivity, credibility, improvability) can underpin a framework of methods and approaches to worth-centred design and evaluation. This moves us beyond post hoc principles to direct support for transdisciplinary principles with their roots in design, rather than in pint-pot psychology. There is an alternative to after the fact generalisation. If we can put aside objective empiricist prejudices, then we can embrace the inherent subjectivity of design and argue systematically for what the principles of designing *should be*, and abandon all pretence of recognising what *they are*.

REFERENCES

1. Barkhuus L., and Rode J. A. *From Mice to Men – 24 years of Evaluation in CHI*, alt.chi 2007 paper, www.viktoria.se/altchi/submissions/submission_barkhuus_0.pdf, last accessed 2/2/08
2. Bødker, S. (2006) "When second wave HCI meets third wave challenges" in *Proc. 4th Nordic Conference on HCI*, eds. A. Mørch, et al, ACM, 1-8
3. Brown, T., (2005) "Strategy by Design" in *Fast Company*, [www.ideo.com/pdf/FastCo-StrategyByDesign\(TimBrown\).pdf](http://www.ideo.com/pdf/FastCo-StrategyByDesign(TimBrown).pdf). last accessed 2/2/08
4. Cockton, G. (2007a) "Make Evaluation Poverty History" alt.chi 2007 paper available from www.viktoria.se/altchi/submissions/submission_gilbert_0.pdf, last accessed 2/2/08.
5. Dourish, P. (2001) *Where the Action Is: the Foundations of Embodied Interaction*. MIT Press.
6. Friedman, B. and Kahn, P., "Chapter 61: Human Values, Ethics and Design", in *The Human-Computer Interaction Handbook*, eds. J. Jacko and A. Sears, 1171–1201, LEA.
7. Gould, J., and Lewis, C. (1985) "Designing for usability: Key principles and what designers think," *Communications of the ACM*, 28(3), 300-311.
8. Gould, J (1988) "How to Design Usable Systems in M. Helander (Ed.) *Handbook of Human-Computer Interaction*, 1st Edition, North-Holland, 757–789.
9. Gould, J., Boies, S.J. and Ukelson, J. (1997) "How To Design Usable Systems" in M. Helander, T.K. Landauer, and P.V. Prabhu (Eds). *Handbook of Human-Computer Interaction*, 2nd Edition, 231-254.
10. Grayling, A.C. (2001) *Wittgenstein: a Very Short Introduction*, Oxford Paperbacks.
11. Heskett, J. (2002) *Design: A Very Short Introduction*, Oxford Paperbacks.
12. Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B. B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N., and Eiderbäck, B. (2003) "Technology probes: inspiring design for and with families," *Proc. CHI '03*, ACM, 17-24
13. Koskinen, I. and Battarbee, K. (2003) "Introduction to user experience and empathic design" in *Empathic Design – User Experience in Product Design*, eds. I. Koskinen, K. Battarbee and T. Mattelmäki, Edita Publishing, Finland, pp. 37–50.
14. Lieberman, H. (2003) *The Tyranny of Evaluation*, CHI Fringe, web.media.mit.edu/~lieber/Misc/Tyranny-Evaluation.html, last accessed 2/2/08
15. McCarthy, J. and Wright, P. 2004 *Technology as Experience*. The MIT Press.
16. Potter, N. (1989) *What Is a Designer: Things, Places, Messages*, 3rd Edition, Hyphen Press.
17. Rogers, Y., Bannon, L., and Button, G. (1994) "Rethinking theoretical frameworks for HCI:" in *SIGCHI Bulletin*, 26(1), 28-30.
18. Rosenbaum, S. (2007) "The Future of Usability Evaluation: Increasing Impact on Value," in *Maturing Usability: Quality in Software, Interaction and Value*, eds. E. Law, E. Hvannberg and G. Cockton, Springer.
19. Sengers, P., Boehner, K., David, S., and Kaye, J. 'J'. (2005) "Reflective design," in *Proc. Conference on Critical Computing*, eds. O.W. Bertelsen et al., ACM, 49-58.
20. Whiteside, J., Bennett, J., and Holtzblatt, K. (1988) "Usability engineering: Our experience and evolution," in *Handbook of HCI*, 1st Edition, ed. M. Helander., North-Holland, 791-817.