CHAPTER 6: TRUSTWORTHINESS, VALIDATION, AND CRITICISM

At the outset of the thesis I stated that an ideal theory for interaction design would describe critical concepts, principles and definitions, and provide an explanation of the relationships, actions, actors and processes. In chapters 4 and 5, I applied the theory to two interaction design projects that demonstrate the descriptive capacity of the theory. I also stated that an interaction design theory would facilitate the generation of new forms of practice, creativity, and discoveries with a prospective orientation grounded in the practice of making. Knowing how we generate trustworthy knowledge in our practice and research in interaction design is at the heart of facilitating these issues. In this chapter I will address knowledge creation by discussing the role of validity in the theory.

As was discussed at length in chapters 1 and 2, interaction design has a particular relationship to human-computer interaction (HCI). I earlier commented that the strength of HCI theory is that it has depth. Underlying HCI is an epistemological grounding in scientific realism such that there is agreement on core concepts. The epistemological viewpoint holds steady and affords a dynamic research and practice space for experimentation and debate. Practice and research methods in HCI are open to revisions and are subject to vigorous and critical debate, which grows the field intellectually. This strength in focus and epistemology creates coherence around principles in research that in
turn allow for flexibility and experimentation with methods of research, practice, and the means to verify claims.

The theory of interaction design proposed here is modelled after such an approach, however there is a substantive difference in epistemological viewpoints: pragmatism over scientific realism. This leads interaction design down a different path that both intersects with and diverges from HCI. I have argued that the pragmatist philosophical orientation is acutely relevant to interaction design whereas scientific realism is not. This difference reframes basic theoretical views of interaction design. This difference also reframes how the discipline is described, namely the concepts, definitions, actors and processes, and how the discipline verifies its claims to new knowledge.

Prior to delving into the issues of validity in the theory it is helpful to take a step back and to position the theory in a broader research context. Scientific realism in HCI is grounded in a positivist paradigm that relies on a quantitative research perspective, whereas pragmatism leads our interaction design theory towards a qualitative research perspective. Despite the differences, the paradigms interact more than the respective advocates would admit, in approaches that are plainly described as “mixed methods” (Creswell, 2007), to more nuanced theoretical positions like postpositivism (Phillips and Burbules, 2000, Denzin and Lincoln, 2005), critical realism (Danermark et al., 2002), and subtle realism (Hammersley, 1992, Seale, 2000). Pragmatism comfortably sidesteps foundational boundaries in openness to experimentation of any kind. As such, our pragmatist theory willingly incorporates a quantitative view within a qualitative orientation. The qualitative orientation is critical to understanding the importance of the descriptive capacity of the pragmatist theory and its approaches to validity.
In order to understand the context of qualitative research in which the theory operates, I quote at length from Denzin and Lincoln:

Qualitative research is a situated activity that locates the observer in the world. It consists of a set of interpretive, material practices that make the world visible. These practices transform the world. They turn the world into a series of representations, including field notes, interviews, conversations, photographs, recordings, and memos to the self. At this level, qualitative research involves an interpretive, naturalistic approach to the world. This means that qualitative researchers study things in their natural settings, attempting to make sense of, or interpret, phenomena in terms of the meanings people bring to them.

Qualitative research involves the studied use and collection of a variety of empirical materials—case study; personal experience; introspection; life story; interview; artifacts; cultural texts and productions; observational, historical, interactional, and visual texts—that describe routine and problematic moments and meanings in individuals’ lives. Accordingly, qualitative researchers deploy a wide range of interconnected interpretive practices, hoping that each practice makes the world visible in a different way. Hence there is frequently a commitment to using more than one interpretive practice in any study (Denzin and Lincoln, 2005, pp.3-4).

The importance of qualitative research to the pragmatist theory is the many characteristics they share. In adapting the shared principles to interaction design they are further shaped and refined. Denzin and Lincoln describe how the researcher or observer is placed in the world. In short, the researcher is implicated in the same world he or she is studying. Qualitative research does not afford the researcher a distant or objective relationship to the research. In Creswell’s view this characterizes the inquiry as interpretive: “The researchers’ interpretations cannot be separated from their own background, history, context, and prior understandings” (Creswell, 2007, p.39). In our pragmatist theory of interaction design, the implication of the researcher goes much further. The designer inquirer actively shapes the phenomenon of the study, the design practice. In the interaction design inquiries in this thesis, the theory relied on self-reflection and auto-ethnography. The theory argues for the embodied designer inquirer and this highlights
issues of data collection methods, as well as the relationship between practice and research. Firstly, in interaction design research, the main data collection is more often than not interpretive, self reflective, and ethnographic. The critical modes of data collection are the self-reflexivity, as in the inquiries in this thesis, or participant observation with a strong emphasis on descriptions and accounts of the designer inquirer. This does not preclude other data collection methods like interviews, document analysis, artifact analysis, experimental studies and so on. Yet the interpretive and self-reflexive nature raises the importance of carefully addressing the verifiability of the claims and findings in interaction design research. The discussions and principles of qualitative research traditions are particularly helpful in this regard.

Given that the designer inquirer influences the phenomenon of study, the question of the practice and research in interaction design has many dimensions. Firstly, as noted by Denzin and Lincoln, in qualitative research there is the aim to “transform the world.” This in itself has many nuances and varies among the different qualitative research strategies. For example, postmodern strategies see the world in the text, and the outcome of the research directly reshapes the textual reality through language. In another example, researchers are seen as enmeshed in a political and ethical existence that inevitably leads to research having both an advocacy role and a goal of advocacy and change, whether through policy or activism. In interaction design and design theories in general, some make the distinction between the descriptive nature of purely empirical research, describing and explaining the world, with what some refer to as normative research, an approach aimed at changing the world in and through research (Robinson, 2001). For example, Schön’s reflective practice can be viewed as a normative theory, which means that the measure of its success or validity is the degree to which it aids professional practice in changing the
world, and not in the validity of its descriptions or findings about professional practice. Hence the practice of design itself is research in the normative sense and the effects of the outcomes are results. The pragmatist theory supports this assumption but argues that an understanding of interaction design requires a systematic and consistent descriptive representation of the phenomenon as well. This level of reflection, in addition to reflection on practice, is how we will best understand interaction design and its role and impact on other disciplines and modes of knowing. A normative view makes the separation of research from practice difficult if not impossible but I suggest that the theory proposed here ably describes both aspects and supports both research and practice without having to completely disentangle the two from each other.

The demand for and ability of the theory to reflectively offer a descriptive representation of interaction design is akin to making the world visible through interpretation in qualitative research. In social sciences, this is a question of the hidden social relations made evident, or in health sciences, the discoveries of social interconnections that determine or affect health or health practices but are typically overshadowed by medical and biological research. Interaction design is a social and human practice that is comprised of many social interactions and relations, but as we have seen it is reliant on designer inquirer judgments and interpretations that have largely been hidden, mystified, or romanticized. This makes it even more imperative to have a theory that aids the explicit and clear representation of interconnections, reflections and decisions in interaction design. In this sense the theory acknowledges its normative framing and also sees equally important the role of descriptive representation and explanation.

A description of the experience of interaction design is a critical step in making the world of interaction design visible, establishing critical cues in normative practice, and
making explicit the phenomenon for future study and to verify new knowledge and findings. Creswell summarizes the aim in establishing a holistic account in qualitative research that largely applies to the theory in interaction design:

Qualitative researchers try to develop a complex picture of the problem or issue under study. This involves reporting multiple perspectives, identifying the many factors involved in a situation, and generally sketching the larger picture that emerges. Researchers are bound not by tight cause-and-effect relationships among factors, but rather by identifying the complex interactions of factors in any situation (Creswell, 2007, p.39).

A research account of an interaction design inquiry shares this holistic aim and stresses the importance of identifying the interaction of factors more so than determining the causality behind actions. In design practice, the holistic view may not be explicit but no doubt the experienced interaction designer is aware of the contours of the emerging process, keenly mindful of cues leading to the key actions like formative findings. The theory describes these interactions by identifying critical aspects in practice and detailing the holistic view for the research inquiry.

Lastly, Denzin and Lincoln (Denzin and Lincoln, 2005) refer to a commitment to multiple and interconnected interpretive strategies in order to make the world visible in a different way. Multiplicity is found in the different research strategies applied, particularly with verification of knowledge. I will now move on to discuss how the theory adopts multiple plausibility strategies including trustworthiness, validation, and criticism.

6.1 Plausibility of interaction design knowledge

Guba and Lincoln ask the central question about the aims of any social research: “How do we know when we have specific social inquiries that are faithful enough to some human construction that we may feel safe in acting on them, or, more important, that members of the community in which the research is conducted may act on them” (Guba
The claims of validity rest on the establishment of security and trustworthiness in the knowledge generated by research. Guba and Lincoln see the debate about validity cresting over two forms of rigor. One form, rooted in positivism, looks to the rigor in the application of method thus asking researchers to negotiate terms of internal validity, external validity, and reliability en route to a secure claim to new knowledge. The second form of rigor advocates community consent along with a rigor of interpretation that invokes terms like credibility, transferability, dependability, and confirmability (Guba and Lincoln, 2005). The differences lie in the rigor of method versus rigor of interpretation. Needless to say the differences are not steadfast, which has both negative and positive implications. Negatively, Guba and Lincoln argue that warring advocates conflate the separate issues of methods and interpretation thus comparing apples to oranges. On the positive side, research strategies shift from the exclusive positivist perspectives on one side (rigor of method) and postmodern perspectives on the other side (rigor of interpretation) to constructively cohabitate the two forms of rigor in validation. Our own pragmatist theory comfortably adopts a range in qualitative strategies that varies the emphasis on interpretation over method for validity, with an incorporation of quantitative strategies that accedes to a rigor of methods in order to gain a wider interpretive space.

In explaining the role of validity in our pragmatist theory of interaction design, our discussion will traverse across the strategies of credibility, validity, and criticism. All play integral if at times separate roles in securing the claims of knowledge in interaction design.

6.1.1 Trustworthiness in knowing

Interaction design research is a qualitative endeavour. As such it incorporates the many qualitative strategies for validation. Creswell enumerates eight procedures (Creswell,
2007): prolonged engagement and persistent observation in the field; triangulation and multiple sourcing of data; peer review and debriefing for external checks; negative case analysis; clarifying of researcher bias; member checking; thick description; and external audits. Lincoln and Guba (Lincoln and Guba, 1985) use the terms credibility, transferability, dependability, and confirmability to group various procedures together under larger aims, and to offer alternative terms to positivist concepts. They establish that trustworthiness of research and its findings are the central issues in positivist ideals of validity and reliability. In that sense, the terms proposed by Lincoln and Guba are very useful in explaining the trustworthiness aims of the theory.

6.1.1.1 Credibility

The basis of any claim to trustworthy knowledge is credibility. Rather than assume that there is a “truth value” in the research, e.g. the singular truth to be uncovered, credibility establishes that the representation constructed through research is indeed valid and believable. Seale sees credibility as follows:

First, credibility should replace truth value. Through prolonged engagement in the field, persistent observation and triangulation exercises, as well as exposure of the research report to criticism by a disinterested peer reviewer and a search for negative instances that challenge emerging hypotheses and demand their reformulation, credibility is built up (Seale, 2000, p.44).

On the first count of prolongment and persistence, the role of the embodied designer inquirer ensures that this count is met. In the case of a third person study, ethnographic commitments would need to be considered. Critics of first-person studies argue that the accounts are distorted and biased and that a third-person researcher is required to mediate in the case of an informant with an agenda. What of an informant researcher like the designer? As I have discussed, the proactive shaping of the designer is integral to the design experience so the designer’s “agenda” is commensurate or so commingled with the
experience as to not be separable (the pragmatist point of entities-in-interaction forming the experience). However, if we put this issue aside there is a greater weight on the designer inquirer to ensure multiplicity and triangulation of sources and data. In the analysis of the two interaction design inquiries, the sources were diverse: including design documents; process artifacts like videos, sketches, storyboards; peer reviewed publications; personal memories; interviews with stakeholders including design team members, stakeholder views and stakeholder generated artifacts and actions, recollections, evaluation results, notes authored by different team members, and so on. In addition, many of the documents and artifacts were collaboratively produced, manifesting the consolidated but multiple viewpoints of those involved in the making of the design experiences. The descriptions were careful to note the different sources and to depict the connection among the sources that in turn supported actions and findings. The theory strongly supports reflexivity of the researcher, which mitigates the concerns of first-person data collection and analysis (see 6.1.1.3 Dependability and confirmability).

Disinterested peer reviewers are incorporated in the peer review publications in the form of blind reviews. A negative case analysis was not discussed in the earlier accounts but one is clear in the case of socio-ec(h)o. Early in the inquiry we conducted an information ecology charrette, se-J9. The charrette has no connection to any intentions or prior judgments (see Figure 44). The theory claims the common pattern (see 5.2.2.2 The common pattern in interaction design) that judgment actions are initiated by an intention or on the receiving end of a mirror or feedback loop, or at minimum a cascading connection from a previous judgment. None of these attributes applies to the information ecology charrette yet they are part of the inquiry. What has been discussed is that the designer inquirer is implicitly embodied in the relationships (or non-relationships) among the entities.
in the inquiry. In addition, the presence of the designer inquirer at the inquiry layer is expressed as a set of interests, experiences, and aims. The information ecology charette can be explained as an influencing factor of the previous research project, ec(h)o, by the designer inquirer. Information ecology was such a central concept to that project that it is not a surprise that the designer inquirer would carry it forward into a new project. It is also not a surprise that since it was not supported by an explicit intention or rationale it did not have an impact on the inquiry. Noting this, the theory currently does not fully address the longer cycle of inquiry to inquiry. Additionally, this is in part discussed further in the section on criticism (see 6.1.3 Criticism).

Credibility through prolonged engagement, triangulation and diversity of data, disinterested peer review, and negative case analysis establishes the believability in the representation and findings generated by the research.

6.1.1.2 Transferability

The next critical step en route to trustworthiness is the degree to which the representation and findings are applicable beyond the particular setting or inquiry. Lincoln and Guba (Lincoln and Guba, 1985) argue that naturalistic inquiries or qualitative research seek transferability by providing the details of the phenomena and setting in formal accounts like research reports. This is achieved by thick description, a rhetorical strategy that richly and “thickly” describes events such that the reader can feel that they experience the events described. This technique arose from rhetorical strategies used in ethnography. Geertz describes how anthropologists persuade readers that they have ‘been there’:

Ethnographers need to convince us…not merely that they themselves have truly ‘been there,’ but…that had we been there we should have seen what they saw, felt what they felt, concluded what they concluded (Geertz, 1988, p.16)
I used this technique in describing the experience of the Canadian Nature Museum from the perspective of a designer inquirer in order to illustrate concreteness in interaction design experiences (see 3.2.1 Concreteness). Other examples of thick description in interaction design are system and interaction experiences that I’ve used in published accounts of the interaction design inquiries in this thesis. For example, here is an account of socio-ec(h)o that accompanied many of the published articles (Wakkary et al., 2005, Wakkary, 2008):

Madison, Corey, Elias and Trevor have just completed the first level of socio-ec(h)o. They discovered that each of them had to be low to the ground, still, practically on all fours. Once they had done that, the space became bathed in warm yellow light and filled with a wellspring sound of resonating cymbals. Minutes earlier, the space was very dim – almost pitch black until their eyes adjusted. A quiet soundscape of “electronic crickets” enveloped them. They discussed and tried out many possibilities to solving the word puzzle: “Opposites: Lo and behold.” At Corey’s urging, the four grouped together on the edge of the space and systematically sent a player at a time to the opposite side in order to gauge any change in the environment. Nothing changed. Madison, without communicating to anyone realized the obvious clue of “Lo” or “low”. She lowered herself to a crouching position. The space immediately glowed red and became brighter. The audio changed into a rising chorus of cymbals – not loud but progressively more pronounced. Corey and Trevor stopped talking and looked around at the changing space. Madison, after a pause began to say “Get down! Get down!” Elias stooped down immediately and the space became even brighter. Corey and Trevor dropped down in unison and the space soon became bathed in a warm yellow light like daylight. The audio reverberated in the space. A loud cheer of recognition came from the group, “Aaaaahhh! We got it!” (Wakkary et al., 2005, p.766)

The theory provides clear guidance to description of the entities in the interaction design inquiry. As a consequence, the details and setting of both the design events and reflections can readily be incorporated as an embedded rhetorical strategy in the research accounts. In another account, we published a reflection on the design process as a way of illustrating the use of participatory workshops for designing the audio display:
The two workshops that we describe came midway through the design process. We had previously hosted several other participatory workshops and conducted concept development meetings where we developed the conceptual foundations of socio-ec(h)o, which included core game mechanics, game progression and structure, and narrative development. We had yet to build a working prototype. Our main concern at this stage was the design of a compelling environment based on user engagement, movements in physical space, immersion, and narrative or game progression. We knew at this point that we needed to investigate specifics in the role that the audio display would have. We had determined that the technical preconditions included location tracking, and an ambient interface that might involve body and object movement, location, and gestures. Given the AmI [ambient intelligence] nature of the project we ruled out a graphical user interface of any kind (Droumeva and Wakkary, 2006, p.37).

The aim of thick description with respect to validity is to describe the event and context in sufficient detail that the reader and researcher can apply relevant knowledge from the account to another setting or event. Seale concisely illustrates the role of thick description and transferability:

Thick descriptions of particular settings are appropriate, giving sufficient detail about context of events so that readers can vicariously experience what it was like to be in the setting. Readers can then conduct their own ‘thought experiment’ in seeking to transfer the lessons learned from this setting encountered through a research text (Seale, 2000, p.41).

The onus is therefore on the interaction design inquirer to detail the accounts of the interaction design experience and research such that it supports a ‘thought experiment’ of applicability in another or similar circumstance. In short, the designer inquirer or interaction design researcher must support the reader in making their own judgment about the relevance of the findings for their own particular design situation or design research problem.

By Lincoln and Guba’s own account, transferability is the natural inquirer’s response to external validity in positivism (Lincoln and Guba, 1985). The idea is that causality proven in a sample study can be generalized to a similar sample or the larger set from which the sample was drawn. In this example, the sample is the typical case, and great
efforts are put into the rigor of method in establishing and testing the sample. Qualitative researchers argue that ecological validity prevents this type of large number sampling, yet transferability shows how single studies with ecological validity intact can be seen as types. There is also what Seale refers to as theoretical generalization that supports single in-depth qualitative cases that can be generalized from the particular, and not from a type (Seale, 2000). I believe this applies to interaction design research. Seale states that theoretical generalization rests on the logic rather than the probability that results from the study. He cites Mitchell (Mitchell, 1983) to say that a case or study may not be representative but the analysis emergent in the case is “unassailable.” He points to analytic induction in which cases are not chosen beforehand as representative examples; rather they serve to illuminate aspects of a theory.

The idea that design of embodied systems is best supported by activities like workshops rather than iterated design representations like scenarios can be regarded as an example of how the theory supports theoretical generalization. In socio-ec(h)o, the overall pattern in the inquiry was for interpretations to build upon each other through participatory workshops (see Figure 44). In the isolated cases where a scenario was used, like se-J3, the trading game scenario does not connect to other actions. Further, a key decision between either puzzles or narratives as the way to structure the game play was resolved by workshops. These workshops occurred late in the inquiry, suggesting that each concept was successfully carried and experimented with through the workshops. The pattern or underlying logic of the actions of workshops and supporting interpretations substantiates the finding on the use of participatory workshops in similar types of inquiries. In comparison with ec(h)o, which was more of a tangible user interface design problem, workshops played supporting roles to scenarios as opposed to being relied upon exclusively.
to explore concepts. Additional descriptions in other inquiries do help support the claim, but the particular example can substantiate a logic quite clearly and fully.

6.1.1.3 Dependability and confirmability

I have brought together the two issues of dependability and confirmability under one discussion since the supporting approaches overlap both concerns. Dependability asks what are the shared or common constructs by which we assess research accounts. Confirmability asks to what degree we can confirm the findings by ensuring a degree of replicability in qualitative terms of any study. Qualitative researchers do not agree upon these terms; some question their import given the underlying constructivist notion of multiple perspectives over absolute truth that underpins qualitative thinking. Nevertheless, they serve a practical end, which is to explicitly show how studies and findings can build on each other through follow-up studies of similar inquiries. Additionally, the concepts when followed up on strengthen the claims to transferability by showing the methodological rigor of the interpretive approaches. This rigor is best characterized as transparency.

The main claim to transparency comes from reflexivity in research investigations. Reflexivity is the presence of a methodologically self-critical account of how the research was conducted. The notion of reflexivity extends to other aspects like credibility where the matter arises through the researcher’s explicit accounting for his or her presence in the research. However, the critical principle underlying reflexivity is the understanding that trustworthiness of research is a discursive matter, e.g. it is always negotiable or open to productive criticism, and not absolute or unassailable as a matter of conclusive proof. The exposure through self-critical accounts offers the most constructive platform for negotiation, criticism, and being shown to be false. This assumes a fallibilism that is
grounded equally in the pragmatism of C.S. Peirce and Dewey as it is in Popper’s fallibilistic presuppositions in quantitative approaches. The embodied and proactive designer inquirer at the centre of the interaction design theory ensures reflexivity in research accounts and data collection, as does the careful categorizing and descriptions of multiple types of data afforded by the theory.

Lincoln and Guba propose a procedure that they call auditing that manifests reflexivity clearly in the research process (Lincoln and Guba, 1985). In essence they argue for an “audit trail” as part of the research process. They propose a systematic accounting of the instances of methodological reflexivity in which data and data analysis produced through the research are catalogued and reviewed during the course of the study by peer auditors. Auditors would look at the “raw data” and examples of analysis, in addition to the triangulation and relationships drawn between the data and evaluations. In addition, auditors would seek out research diary entries, meeting notes for decision-making, analytical strategies and so on in order to provide a complete and thorough methodological formative assessment. According to Lincoln and Guba, researchers and auditors would agree to meet prior to a series of visits at different stages of the research. The proposal is quite thorough and exhaustive. For most researchers it is a challenge to follow the process yet it is less time-consuming than replicating the study (Seale, 2000). Interestingly, the description of auditing and auditors, while less formal, is not unlike the actual collaborative practice of an interaction design inquiry. The auditors are of course peers and members of the team, however in most teams there is this ongoing check of design representations, arguments, and strategies that are not unlike an integrity check of the practice.

Returning to research aims, there is a range of practical strategies to auditing a research study. This is similar to the range of strategies with inter-rater reliability that
include hiring external coders, to calculating statistical reliance between codings, to working within the research team to arrive at a consensus with the codings. The critical point in interaction design is to formalize evidentiary data collection that is diverse and ensure the data is sufficiently detailed to offer an explicit and public accounting of the research process. Similarly, Amanda Jane Coffey and Paul Atkinson write:

“Transactions and the ideas that emerge from them should be documented. The construction of analytic or methodological memoranda and working papers, and the consequent explication of working hypotheses, are of vital importance. It is important that the processes of exploration and abduction be documented and retrievable. Their documentation is part of the transformation of data from personal experience and intuition to public and accountable knowledge” (Coffey and Atkinson, 1996, p.191).

The proposed interaction design theory provides the taxonomical descriptors that cover the range and diversity of data, and also cover the transactions and ideas that result from the design inquiry. This in effect provides the tools for an intellectual and research audit trail of the interaction design inquiry. The accounting extends to the “triangulation” or the relationships among transactions and entities. While the theory does not call for external auditors, the explanation of relationships among the inquiry data is expressed as a matter of integrity with respect to practice and research. The higher-level taxonomy conveniently reduces these elements to actions and layers that ultimately create descriptive flexibility. The theory provides the interaction design researcher and practitioner with procedural support. Chapters 4 and 5 amply demonstrated the descriptive or auditing functions of the theory.
Equally important is that the theory provides the shared construct necessary for confirmability. In Table 21, the framework is represented to show the structure based on a hierarchy of layers, actions, and outcomes. I have detailed the interactions and the descriptive purview of each element in earlier chapters.

In addition to the taxonomy, I discussed structural descriptors that analyse the design inquiry and help the researcher measure the structural quality (see 5.2.2 The experience layer view). These offer added support for reflexivity. Table 22 shows the common pattern, which is a basic structure of any inquiry (see 5.2.2.2 The common pattern in interaction design). Table 22 also shows descriptions of relationships among entities that include formative findings, mirrors and feedback loops, and cascading (see 5.2.2.1 The relationships in an interaction design inquiry). At a higher level, the inquiry can be seen to be composed of the anatomical structures of inter-integrity, intra-integrity, and extra-integrity (for more details see 5.2.2.3 The dynamic structure of an inquiry).

<table>
<thead>
<tr>
<th>Layers</th>
<th>Actions</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>Theoretical Reflections</td>
<td>Accounts</td>
</tr>
<tr>
<td>Inquiry</td>
<td>Designer Inquirer</td>
<td>Design Intentions</td>
</tr>
<tr>
<td></td>
<td>Judgment</td>
<td>Representations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Activities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Models, Artifacts, Systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Evaluations</td>
</tr>
<tr>
<td>Actions</td>
<td>Interpretations</td>
<td>Accounts</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stakeholder Views</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Findings</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Criticism</td>
</tr>
</tbody>
</table>

**Table 21 Interaction design theory framework**
1. An intention is required to initiate the design inquiry. Ideally it is supported by a design rationale.
2. The intention initiates a judgment that is the first interaction design action.
3. The action is interpreted and the reflection is mirrored or fed back into the inquiry in the form of another judgment.

**Relationships of the inquiry**

Formative and summative findings
Mirrors and feedback loops
Cascading

**Anatomical structures**

Inter-integrity (Mirrors and feedback loops, formative findings, and initiations): relationships across types of actions like interpretations and judgments; relationships across layers like intentions crossing from the inquiry layer to the actions layer; and formative findings that prescribe judgment actions. In all cases, these relationships cross kind in either action or layer.

Intra-integrity (Cascades): relationships that connect similar actions together like a series of judgment actions leading up to a final prototype, or a series of design charrettes to orientate the design team.

Extra-integrity (Findings and criticism): a series of findings that cascade into forms of criticism external to the designer inquirer. Criticism takes on the form of a “meta feedback loop” that influences subsequent design inquiries. This pattern is more speculative and posited theoretically. Given the state of interaction design it is not evident in the projects described.

**Table 22 Structural descriptors of interaction design inquiries that can support reflexivity and auditing**

The framework itself is open to further negotiation and becomes a discursive object in itself. However, the clear articulation of a construct for description and analysis of interaction design is necessary to ground and manifest the researcher’s reflexivity. The framework functions as the agreed upon (or further negotiated) construct by which the knowledge generated can be said to be confirmed and dependable.

**6.1.2 Quantitative validation**

In pragmatism there is no single methodology that can be utilized to cover the scope and complexity of experience. Inquiries share the general pattern of beginning with a genuinely problematic situation that can come to a resolution via processes of manipulation and experimentation. I have detailed this in the earlier accounts and described at the framework level a common pattern for interaction design (see 5.2.2.2 The common pattern in interaction design). This pattern and the theoretical framework encourage wide latitude
for specific methodologies that are judged only by their matching with particular types of problematic situations. There is no absolute or foundational language to which all other languages must be reduced. A constructivist approach is adopted in which quantitative methods are included if not welcomed.

The judgment actions are expansive and inclusive. While the inquiries analyzed in the thesis show no evidence of this, an experimental study is as much at home at the beginning of an inquiry as a charette or workshop. Such a study could be followed by an interpretation and its findings could be based on statistical analysis. Design fields have been open to and have used scientific realism findings in psychology, sociology or computing science, yet less so is scientific realism used within its own practices nor is it typically seen as part of the various practices. In our own practice we have incorporated quantitative analysis where it was sensible. For example we examined the coefficient alpha, which is often used as a measure of the reliability of a questionnaire instrument, for our own questionnaire as a test of internal validity (Jiang, 2008).

For the most part, our use of quantitative analysis is part of a “mixed methods” approach (Creswell, 2003). In this section, I will discuss two examples. The first is the use of data from an evaluation of the ec(h)o system, the second discusses the statistical analysis of video coding of a socio-ec(h)o user study.

6.1.2.1 User perception testing in ec(h)o

In this example, our evaluation mixed quantitative statistical analysis with qualitative analysis of a semi-structured interview of each participant in order to confirm and support the findings. This analysis was previously published in (Wakkary and Hatala, 2007). We evaluated the final prototype of ec(h)o in a user study at the Canadian Nature Museum. The formal user evaluation effort involved sessions with six participants that
included two men and four women, from 25 to 53 years old. Table 23 shows the characteristics of each user session: total length of the interaction, number of interaction cycles, number of selected and listened to audio objects, and number of location changes.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Length</th>
<th>#Cycles</th>
<th>#Selections</th>
<th>#Locations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1</td>
<td>10:36</td>
<td>27</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>Participant 2</td>
<td>6:19</td>
<td>11</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Participant 3</td>
<td>8:56</td>
<td>22</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Participant 4</td>
<td>9:53</td>
<td>21</td>
<td>16</td>
<td>5</td>
</tr>
<tr>
<td>Participant 5</td>
<td>9:18</td>
<td>22</td>
<td>17</td>
<td>5</td>
</tr>
<tr>
<td>Participant 6</td>
<td>5:01</td>
<td>16</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Expert 1</td>
<td>15:03</td>
<td>32</td>
<td>23</td>
<td>9</td>
</tr>
<tr>
<td>Expert 2</td>
<td>17:58</td>
<td>36</td>
<td>29</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 23 Test session characteristics per participant and two expert reviewers

We evaluated user experience through observation, a questionnaire, and a semi-structured interview. The questionnaire included sixty-three questions that assessed the overall reaction to the system, the user interface, learning how to use the system, perceptions of the system’s performance, the experience of the content, and degree of navigation and control. The majority of the questions in the questionnaire were on a Likert scale (5 points), while some were open-ended written responses.

The results showed that participants found the system enjoyable and stimulating. The general sense of satisfaction was split between those participants who liked the playful approach and those who did not. While our sample was small, we noted a clear age difference in that the “younger” participants rated satisfaction higher based on their liking of the playful approach (this was confirmed in the semi-structured interviews).
Among the factors that stood out as most positive for the participants was that the cube and audio delivery were seen as playful. The open-ended written comments and semi-structured interviews made this point clear as well. The tangible user interface was also well received especially in terms of ergonomics and ease of use. These two activities were documented in Chapter 4 (see 4.2.1 ec(h)o: judgment). In addition, learning to use the
interface and navigation were both rated highly and participants felt the system had a low learning curve and that it was easy to get started:

Umm, I found it was really easy. Sometimes I got so engaged in listening to what they were saying that I forgot in which orientation I was holding the cube. And I found that I would have to occasionally look down. But the way it was designed with the round part to go in your palm… it was really easy to quickly reorient myself to how I was holding that cube (Participant 5).

The questionnaire did point out some challenges and areas for further research. We expected some of the challenges, for example the headphones were uncomfortable, yet to such a degree that it led us to rethink the trade-off between personalized spatial audio and use of headphones that resulted in the audio display design choices in socio-ec(h)o. Other results point to a threshold in the balance between levels of abstraction and local information. Some visitors had difficulties at time connecting what they were listening to with what was in front of them (in part this was an inherent challenge in the exhibition since the display cases had dozens to over a hundred artifacts). In addition, we see both a threshold point in play and focused attention on the exhibit, which is evident because the question relating to the content asking if it was “distractive-synergistic” scored 2.83. This raises the issue of balance in play and the possibility that play shifts attention away from the environment rather than acts as a means of further exploring the environment.

In an open-ended question in the questionnaire and through the interviews, we explored the issues of liminal play and engagement further. The results here are quite clear that play was a critical experiential factor in using the system. It was often remarked how the experience was similar to a game:

The whole system to me felt a lot like a game. I mean I got lost in it, I found myself spending a lot of time in a particular area then I normally would. And just the challenge of waiting to hear what was next, what the little choice of three was going to be. Yeah… So I found it over all engaging, it was fun, and it was very game-like (Participant 4).
The playfulness did in most instances suggest a quality of engagement that led to learning, even through diverse types of museum visits; from the visitor who browses through quickly but is still looking to be engaged to the repeat visitor who experiences the audio information differently each time:

I learned a lot and well you know I’m a scientist here, and I think anybody going through, even people who are in a real rush, are going to pick up some interesting facts going through. And… I mean, that was good, the text was great and was short enough that somebody in a rush is still going to catch the whole thing (Participant 1).

As mentioned earlier, there is a threshold between play in support of the exhibit on display and play with the system that can be an end in itself and even a distraction. For example, one user’s enthusiasm for the game-like quality led her to at times pay more attention to the interaction with the system than the exhibition. In addition, people respond to play differently and can be argued to belong to different types of players (Bartle, 1996), a point we investigated later in socio-ec(h)o. One participant would have preferred a more serious and “non-playful” approach.

The prefaces were playful, but the text was not at all, you know, that contrast between them…. but I find it was too playful and I think maybe, either you, or maybe you could give people the choice between you know choosing a playful or a non-playful version” (Participant 2).

In addition, participants’ observations on the liminality of the experience manifested in comments suggesting that play was more natural for children than themselves, however as expressed below, they soon overcame this issue:

At first it felt a little bit strange, especially holding this cube that looked like a children’s toy, and I felt a little bit awkward about doing that, but I got over that pretty quickly (Participant 5).
It was quite chatty, which was kind of fun. I kind of felt like 'Oh, I bet like a twelve year old would really like this (Participant 3).

In this example, we used a mixed method approach in which quantitative findings were supported and refined with qualitative analysis. The combined approaches led to the findings described in Chapter 4 (see 4.3.2.3 ec(h)o: findings).

6.1.2.2 Evaluation in socio-ec(h)o

In evaluating socio-ec(h)o we had vast amounts of data in the form of video and audio transcripts, system logs, and questionnaire data. This evaluation and analysis was previously published in (Wakkary, 2008). Interaction aspects of the research were exploratory and so we used quantitative analysis to identify points of foci that we could later examine in greater detail through richer qualitative analysis. We analyzed questionnaire data but the bulk of the analysis was conducted with video coding data. In the overall final study, we conducted an experiment involving 56 participants divided into teams of four. The experiment included a two-hour session of playing in the socio-ec(h)o environment. The teams were divided into two groups that each followed a different protocol (see 5.1.1.4 socio-ec(h)o: evaluation). The analysis here reports on only one of the protocols that included groups of thirty-six (36) participants and the play.

Each session was recorded with three cameras and audio, providing ample data to code group actions and behaviours. Our coding scheme was based on two main factors: cohesion and goal focus. I earlier described the scheme (see 5.1.2.1 socio-ec(h)o: accounts) but it is worth detailing further in the context of understanding the data analysis. The combination of these factors in a two-dimensional matrix shows the degree of descriptive capacity, see Figure 46. Two researchers independently coded the videos and for inter-rater reliability they negotiated the differences to reach a consensus.
Figure 46 Matrix showing the descriptive capacity of the two factors

Cohesion can be described as the extent to which players appear to be acting as a team (all members coordinating together); whether that is working on a game solution, playing, thinking, or talking to each other. Cohesiveness is a measure of team dynamics and does not necessarily reflect their focus on the game but only whether they are acting in unison as a team. We analyzed different degrees of cohesion:

- **Low**: players are not together as a group or they are temporarily fragmented. They are not communicating or are individually exploring;
- **Medium**: players are in the process of becoming a group or are regrouping. Players are negotiating roles and establishing leadership or consensus;
- **High**: players constitute an established team. They make several agreements and are coordinated in their movements or are communicating with each other about strategy and solving the puzzles.

Goal focus can be described as the extent to which players appear to be or are attempting to “play the game” the way they understand it. Game activity is not dependent on whether players are working as a team or not. In addition, game activity does not necessarily only mean that players are actively playing, i.e. in our case moving. If players
are still because they believe the game requires them to be still, then they are “playing the game.” We analyzed different degrees of goal focus:

- **Low**: players are not involved in playing the game. They are resting, or are distracted, or engaged in activities not related to the game;
- **Medium**: players are in the process playing the game. They are experimenting with different actions, and communicating with each other about or reflecting on the effects of their actions;
- **High**: players are actively and consciously playing the game and attempting to solve the puzzle at hand. This is reflected in concerted efforts and good communication related to their performance in the game. Many ideas are shared on actions for solving the puzzle.

For rationales and discussion of related literature in the development of the schemes see (Wakkary, 2008).

As I discussed in Chapter 5 (see 5.1.1.4 socio-ec(h)o: evaluation), our analysis looked at the different levels of cohesiveness and goal focus over duration to determine a density value in percentages:

\[
density = \frac{\text{factor(min)}}{\text{duration(min)}}
\]

We looked for combinations of density values of the different degrees (high, medium, low) of the two factors (cohesion, goal focus) and compared these to team performance or duration of the game level. Additionally, we correlated the different degrees of cohesion and goal focus factors with team performance (duration) using the Pearson correlation coefficients. The Pearson correlation coefficients measure the degree and the direction of the linear relation between two variables. That is, how much are changes in one of the variables related to changes in the other variables. Correlation can be used to estimate the extent to which teams’ performance, cohesion and goal focus factors were related. Lastly,
we compared the video coding results with the intensity data from the logs (see Figure 47). Based on these comparisons we isolated key events for further study through transcripts and videos.

![Comparison of video coding results with system logs]

**Figure 47** A comparison of the video coding results (color bars on the top of the figure) with the system logs that measured intensity (graph).

Our results show correlations between high degrees of the two factors: the role of transitions, and players’ perceptions. Table 24 and Table 25 show correlations between cohesion and goal focus. Note that in each table, column numbers refer to the same values as rows, for example in Table 24, row 7 and column 2 show a significant correlation of .871 between the medium degree of cohesion and completion time.

One might expect that a team that showed high density values of both cohesion and goal focus factors would lead to a fast performance in the game. Indeed, we found that Team H held density values of 93% for goal focus and 97% for cohesion in level 4, and completed the level in less than a minute. However, Team D had significantly more modest density values for level 4, 63% for goal focus and 67% for cohesion, yet the team was able to complete the level in just under a minute, (see the comparison in Figure 48). To further the point, a team like Team C, which had a density value for goal focus of 66% and cohesion of 89%, required over 39 minutes to complete the levels.
Statistically, we found no significant correlations between high degrees of cohesion or goal focus factors and team performance in game level 3 (see Table 24). We had virtually the same results for game level 4 (see Table 25). However, Table 24 shows a significant correlation between medium degree of cohesion and performance (.871). Table 25 shows a strong correlation between medium degrees of cohesion and goal focus and performance (.892; .927). This led us to examine the role of transitions, where factors change in degrees such as a team shifts from a high degree of cohesion to medium degree of cohesion.

![Image of bar graphs](image.png)

**Figure 48** Bar graphs showing that high density values of high cohesiveness and high goal focus do not correlate to fast completion as in the example of teams H and D.
<table>
<thead>
<tr>
<th>Level 3</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>High degree of cohesion</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Medium degree of cohesion</td>
<td>.577</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Low degree of cohesion</td>
<td>-.355</td>
<td>.828 **</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>High degree of goal focus</td>
<td>-.346</td>
<td>-.609</td>
<td>.284</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Medium degree of goal focus</td>
<td>.303</td>
<td>.799 **</td>
<td>-.511</td>
<td>.743*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Low degree of game focus</td>
<td>.158</td>
<td>-.010</td>
<td>.172</td>
<td>-.651</td>
<td>-.023</td>
<td>-</td>
</tr>
<tr>
<td>Whether or not completion time is &lt; 5min</td>
<td>.494</td>
<td>.871*</td>
<td>-.743</td>
<td>-.439</td>
<td>.651</td>
<td>-.012</td>
</tr>
</tbody>
</table>

* Correlation is significant at 0.05 (2-tailed), ** Correlation is significant at 0.01 (2-tailed).

**Table 24 Correlations in level 3**

<table>
<thead>
<tr>
<th>Level 4</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>High degree of cohesion</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Medium degree of cohesion</td>
<td>.837**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Low degree of cohesion</td>
<td>-.785*</td>
<td>.358</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>High degree of goal focus</td>
<td>.854**</td>
<td>.834**</td>
<td>-.566</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Medium degree of goal focus</td>
<td>-.785*</td>
<td>.926**</td>
<td>.370</td>
<td>.882**</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Low degree of goal focus</td>
<td>-.590</td>
<td>.450</td>
<td>.586</td>
<td>.816**</td>
<td>.486</td>
<td>-</td>
</tr>
<tr>
<td>Whether or not completion time is &lt; 5min</td>
<td>-.037</td>
<td>.892**</td>
<td>-.334</td>
<td>-.647</td>
<td>.927**</td>
<td>.297</td>
</tr>
</tbody>
</table>

* Correlation is significant at 0.05 (2-tailed), ** Correlation is significant at 0.01 (2-tailed)

**Table 25 Correlations in level 4**

We found that transitions from different levels of coherence and goal focus held statistical significance when compared against performance throughout level 3 of the game, except for transitions to high cohesion, see row 9 in Table 26, and significance in transitions from all degrees of both factors in level 4 of the game except for transitions to low cohesion, see row 9 in Table 27.
### Table 26 Correlations between transitions and duration in level 3

<table>
<thead>
<tr>
<th>Level 3</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total transitions game focus</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transitions to low game focus</td>
<td>.985**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transitions to medium game focus</td>
<td>.854**</td>
<td>.771*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transitions to high game focus</td>
<td>.947**</td>
<td>.963**</td>
<td>.649</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total transitions cohesion</td>
<td>.860**</td>
<td>.823**</td>
<td>.848**</td>
<td>.744*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transitions to low cohesion</td>
<td>.818**</td>
<td>.839**</td>
<td>.670*</td>
<td>.770*</td>
<td>.923**</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transitions to medium cohesion</td>
<td>.883**</td>
<td>.838**</td>
<td>.889**</td>
<td>.754*</td>
<td>.994**</td>
<td>.893**</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transitions to high cohesion</td>
<td>.820**</td>
<td>.775*</td>
<td>.825**</td>
<td>.703*</td>
<td>.995**</td>
<td>.906**</td>
<td>.984**</td>
<td>-</td>
</tr>
<tr>
<td>Completion time</td>
<td>.939**</td>
<td>.960**</td>
<td>.732*</td>
<td>.916**</td>
<td>.688*</td>
<td>.687*</td>
<td>.725*</td>
<td>.625</td>
</tr>
</tbody>
</table>

* Correlation is significant at 0.05 (2-tailed)
** Correlation is significant at 0.01 (2-tailed)

### Table 27 Correlations between transitions and duration in level 4

<table>
<thead>
<tr>
<th>Level 4</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total transitions game focus</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transitions to low game focus</td>
<td>.989**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transitions to medium game focus</td>
<td>.958**</td>
<td>.911**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transitions to high game focus</td>
<td>.971**</td>
<td>.987**</td>
<td>.863**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total transitions cohesion</td>
<td>.897**</td>
<td>.842**</td>
<td>.980**</td>
<td>.769*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transitions to low cohesion</td>
<td>.778*</td>
<td>.690*</td>
<td>.922**</td>
<td>.613</td>
<td>.964**</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transitions to medium cohesion</td>
<td>.897**</td>
<td>.841**</td>
<td>.978**</td>
<td>.770*</td>
<td>.999**</td>
<td>.966**</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Transitions to high cohesion</td>
<td>.942**</td>
<td>.913**</td>
<td>.976**</td>
<td>.845**</td>
<td>.977**</td>
<td>.888**</td>
<td>.972**</td>
<td>-</td>
</tr>
<tr>
<td>Completion time</td>
<td>.950**</td>
<td>.971**</td>
<td>.831**</td>
<td>.988**</td>
<td>.722*</td>
<td>.553</td>
<td>.717*</td>
<td>.817**</td>
</tr>
</tbody>
</table>

* Correlation is significant at 0.05 (2-tailed)
** Correlation is significant at 0.01 (2-tailed)
We examined the relationship between players’ perception of the helpfulness of the system and their performance. No correlation was found, thus fast players did not necessarily believe the system to be more helpful than slow players. However, there was a significant relationship between players on teams who completed the most levels and the players’ perception of the support of the system. This suggests that the more “skilled” players (those who could complete the higher levels) perceived the system to be more helpful. The overall rating of the system was quite good, for example on the question of how helpful the system was, the median score was 4.0 (SD 1.02) on a scale of 1-5 (5 high).

All of the findings discussed here are described in Chapter 5 (5.1.2.3 socio-ec(h)o: findings). This example shows how we validated our findings quantitatively and how quantitative analysis was used to make sense of rich data in order to establish foci for subsequent qualitative analysis.

6.1.3 Criticism

Criticism is an activity central to philosophy for Dewey, who understood it as appraisal or evaluation (Dewey, 1929b). In pragmatism, the value of criticism arises in many respects, from understanding the value of an inquiry in the context of everyday experiences, to the mediation of intended operations over those of chance or whim, and finally to address the role of temporality in experience and its validation over time.

The basis of Dewey’s understanding of the relationship between philosophy and democracy is that knowledge and practice are harmonized in the security and freedom that the values embodied in experience are widely shared (Dewey, 1929b). This wider issue aside, criticism enables the sharing of the values articulated by an inquiry to move beyond the limitations and constraints of the initial inquiry. In pragmatist terms, they become intrinsic goods when they are valued in everyday experiences. In other words, criticism
extends beyond the particular inquiry into a wider set of social interactions in a way that qualitative and quantitative validation strategies do not.

Life experience will happen regardless of our intentions and aims. In some sense this is true of design: things will be made or design actions will occur whether or not someone calls himself or herself a designer and consciously assigns intentions and rationales to their actions. This is what Dewey would consider as life experience with a minimum of regulation. Dewey argues that criticism minimizes these acts of “fortune or providence” by affording intelligent actions that are experienced and come with meaning. It is not that chance does not have a role in design or life but criticism enables the accrual of experience that mediates and guides intentions and actions (Dewey, 1929b). The idea of discernment and evolving design intelligence over time is what makes interpretation central to interaction design. I have shown in the description of inquiries how interpretation is manifest in reflection and action simultaneously. Criticism is the explicit extension of the interpretive and dialogic actions of the inquiry beyond the particular inquiry.

Returning to the idea that criticism is inherently democratic, Dewey saw in democracy the increased participation in a diverse set of activities that could be understood and valued in everyday experience: “Criticism is discriminating judgment, careful appraisal, and judgment is appropriately termed criticism whenever the subject-matter of discrimination concerns goods and values” (Dewey, 1958, p.298). The democratic underpinnings to criticism argue for the participation of many and the access to freely discriminate the “goods and values” of an inquiry in an open exchange. It is this open dialogue in an everyday context focused on the value of an inquiry that makes criticism so central. In pragmatist terms, knowledge becomes negotiated within the affairs of ordinary experience. Criticism mediates the intentional and intelligent operations of experience.
Temporality is infused throughout pragmatism. It has many dimensions that affect interaction design processes, from the interaction of actions, intentions, and rationales to the accrual of experience and development of the designer inquirer. Temporality extends actions into experience, for example past actions and interpretations become embodied in the designer inquirer and are then carried from project to project. It is this longer cycle of inquiries in which criticism plays a strong role. As I have discussed, criticism moves the notion of validation beyond the inquiry into lived experiences. It maintains the dialogic aspect of interpretation and in more fluid and explicit ways it sees interpretations to be more negotiable than even the qualitative approaches to validation that I discussed earlier (see 6.1.1 Trustworthiness in knowing).

In addition to discursivity, lived experiences are contextual and temporal, exposed to the vicissitudes of change and the unforeseen. And so the “value and goods” of the inquiry must also be seen as temporal. As such they are extant and therefore assessed or validated over and through time. The real challenge and promise of criticism is to secure and extend the trustworthiness, validation, and ultimately the value of the goods of the inquiry by discerning which are appropriate and which are not at any given time.

And so I posit the vital importance of criticism in interaction design but I have or we have little to show with respect to examples of criticism in interaction design. In the wider context of design, this absence of criticism, criticism practice, or critical theories in interaction design is glaring. Architectural criticism dates back to Vitruvius in the time of Augustus (63 BC- AD 14) and was present in, if not instrumental to, architectural practice during the Renaissance in key figures like Leon Battista Alberti (1404-1472). Industrial design, graphic design and architecture developed a clear practice of criticism through the late 19th and 20th centuries (Margolin, 1989). Richard Buchanan has suggested a liberal
arts foundation for design built on the critical theories and practices of industrial and graphic design (Buchanan, 1995). Adrian Snodgrass and Richard Coyne (Snodgrass and Coyne, 2006) argue to reclaim the term of interpretation from criticism in architecture: “Whereas we agree that architecture is a discursive practice, and is abetted by talk and writing, we will demonstrate that to design is to interpret” (Snodgrass and Coyne, 2006, p.4). My proposed theory centralizes the actions of interpretation into the practice of interaction design, yet there is no talk or writing “abetting” the practice of interaction design.

The lack of a critical practice has not gone entirely unnoticed. One early exception is Steven Johnson’s Interface Culture (Johnson, 1997) in which Johnson discusses graphical user interfaces in the context of film and literary criticism. More recently, McCarthy and Wright (McCarthy and Wright, 2004) employ rhetorical approaches similar to criticism to illustrate experiences of technology. We are far from a practice of criticism in interaction design but at least this has been identified as a problem by some interaction design and HCI researchers (Löwgren and Stolterman, 2004, Löwgren, 2002, Blythe et al., 2008). Löwgren writes: “An issue of particular interest is the possible role of critics in interaction design. One can imagine a field of interaction design criticism in analogy with more mature design fields such as architecture or graphic design” (Löwgren, 2002). He sees a problem in that the scientific realism orientation of HCI prevents a criticism approach and cites Preece et al’s Interaction Design: Beyond HCI:

This appears problematic from a HCI perspective: ‘Finding measurable characteristics for the user experience criteria is even harder, though. How do you measure satisfaction, fun, motivation or aesthetics? What is entertaining to one person may be boring to another; these kinds of criteria are subjective and so cannot be measured objectively’ [p. 182] (Löwgren, 2002).
I have identified how HCI has overshadowed and curtailed the development of interaction design theory and practice and this includes minimizing the role of criticism (see 1.1 The visionary and status quo versions of interaction design). Despite the clash of an interpretive approach with scientific realism there is an activity in HCI and broader scientific practice that engenders all the critical attributes of quality criticism. This activity is blind peer review.

Blind peer review has all the qualities one would look for in criticism, an informed peer critically analysing the value of a research contribution or design. Blind peer review is the anonymous review of a journal or conference paper submission by several peers. The reviews range in formality but often under the dictates of a journal’s editorial board or a conference program committee, there are minimum criteria to follow and this typically involves a mix of quantitative but primarily qualitative comments and analysis. In the case of journals, manuscripts, and some conferences, a shepherding process is used in which the reviewers maintain anonymity but continue to review changes and ongoing iterations of papers until publication. I have included a review from an early attempt of mine to publish on the ec(h)o project in 2003:

In general the reviewers found this work to be interesting, to build well on prior work (especially that of Woodruff), and to introduce the novel user interface concept of audio icons that invite further interaction at a specific museum location. However, the reviewers also agree that the work is too preliminary (design only), and not well-enough motivated to warrant publication. As it stands the work is a modest addition to existing work; full explication of the design rationale, and a thorough account of connections between the participatory workshops and resulting design features might add enough of a research contribution to make the work publishable, but the authors are also strongly encouraged to implement and test their ideas first.

Several of the reviewers provide important pointers to work not recognized or cited (see in particular the listings by Reviewer #2). The project does a nice job of building on specific related work in museum systems, but do not go beyond to consider the role of other gesture-related work with audio-based systems.
All of the reviewers were intrigued by the role of the conversation model in motivating the audio selection mechanism, so it seems likely the authors have hit upon a fruitful research direction. But this proposed motivation raised as many questions as it answered, so clearly more work needs to be done examining and better articulating the motivation provided by a conversational model. Reviewers #2 offers a detailed critique of the rationale that can be understood (or inferred), while Reviewer #4 describes difficulty in even seeing the connections between many of the proposed interaction criteria and a conversation model. Reviewer #4 also provides a nice summary of what is missing in this paper, namely what are claimed to be the competencies of the users the design is targeting, and what benefits are actually realized for such users in actual museum activities.

In general the writing of the paper was fine, though there were a few comments for improvement, e.g. more careful introduction and use of special vocabulary. The consensus is that the work has promise but must be more developed to qualify as more than a modest incremental contribution.

At the program committee meeting, there was agreement that the arguments summarized above were grounds for rejecting the paper, despite the general interest in this as a research area (ACM CHI 2004 reviewer comments – personal correspondence).

Evidently the paper was rejected but this is not the point of the example; suffice to say that a subsequent version of the paper was later accepted and was directly influenced by the comments of the review. The review presented here is a meta-review written by the review chair that consolidates and summarizes the comments by all the reviewers, in this case four other reviewers. The reviews have all the functional components of criticism:

• Overall assessment of the value of the work in general terms;
• Assessment of the relevancy of this effort with other research or practice efforts such as building on and extending existing practices or overlooking existing works;
• Identification of unique aspects or values in the work or identification of redundancies or derivative aspects or values;
• Assessment of the contribution value of the work in terms of strengths and weakness;
• Identification of promise latent in the work or not.
Lastly, the reviews offer multiple views that show differences in their assessments. The constructivist principle of multiplicity and interpretations reigns in criticism but this does not lead to relativism. The differences are negotiated through the public exposure of the reviews, at least among the reviewers and chairs. In addition, the quality of the reviews differs. At stake is a craft of constructing a review that supports one’s impact in the implicit negotiation of the validity of the work. I have shown how in Dewey’s view, the mediation itself (that is criticism) is subject to establishing its own value. Snodgrass and Coyne write: “Interpretation, then, is ‘the working out of possibilities projected in understanding’, that is, it is the working out of how something figured in the context in which it stands” (Snodgrass and Coyne, 2006, p.38). This view of criticism is very much in keeping with pragmatism. Rather than a retrospective view, criticism prospectively discovers possibilities in the work (positively or negatively) and therefore shares the potential risk as to whether or not these possibilities will either come to fruition or are equally perceived by others. For those who have reviewed other works I’m sure you have experienced a case where you champion a particular paper or work and therefore share the risk that the perceived value is or is not in fact valid or will or will not manifest. Conversely, you may have negatively reviewed a work but had thoughts that you were simply unable to see the value that will inevitably be unequivocal to others.

The functional qualities of blind peer reviews are directly relevant and supportive of critical practice, yet the principles and practice of blind peer reviews are antithetical or at minimum severely limit the practice of criticism in interaction design. In pragmatist terms, the strength of blind peer review is in its mediation of intended operations. Its weaknesses lie in its inability to mediate in the context of everyday experiences. This particular weakness limits the criticism to the inquiry and not beyond the inquiry. The summary of
the challenge is the lack of democratic practice within a blind review process. The access to and participation of the role of critic is limited and controlled, as is the distribution and audience of the critique. It is again a conflict of epistemological assumptions and safeguards. Blind reviews serve the needs of scientific realism and certainly have the same role in interaction design research as they do in HCI. However, if we look to the functional strengths of blind peer review, we can see how we can leverage the current knowledge and routines of current interaction design researchers. We can also see how the functionality of blind peer review can be modified to bootstrap a practice of criticism. This would involve:

- **Publicly declared authorship in reviews:** Reviews are classically authored and so they are not anonymous. This maximizes the shared risk in negotiating the values of interaction design practice and research. Additionally, the declared author establishes credibility in the criticism;

- **Reviewers are not blind to the designer inquirer:** The designer inquirer is an integral part of the interaction design inquiry and is part of the object of review;

- **Public dissemination of reviews:** Reviews are public and are also publications in their own right. Criticism is itself a discursive object and so is open to negotiation and interpretation;

- **Critics are self-defined:** The role of critic is not predetermined; it is largely self-defined and emerges from a range of backgrounds of which interaction design is only one;

- **Broader definition of the object under review and the contributions:** Blind peer review is focused on research and the publication of archival text. Criticism in interaction design by definition has a broader focus on practice, designers, research, and artifacts;

- **Developing the practice of writing beyond service writing:** Criticism is a practice mostly manifest in writing. The act and artifacts of writing are central to the craft of criticism.

In many respects, criticism is the one validation strategy in which interaction designers are potentially most at home. In interaction design practice, inherited from the
traditions of design, there is ongoing negotiation and forms of criticism. For example the
design studio critique is a longstanding form of criticism that establishes interpretation,
reflection, discursivity, and validation as part of a central ongoing routine in practice. Other
existing forms of criticism include the use of heuristics, design exemplars, and on occasion
critical theory from other design traditions.

Emerging discussions in HCI have argued for a reflective and critical stance in HCI
research. For example, Paul Dourish argues that ethnographic approaches offer a path to
theoretical and critical inquiry within HCI (Dourish, 2006). In various writings, Phoebe
Sengers and colleagues argue for an interpretive and reflective practice in HCI grounded in
intellectual traditions of design and critical theory (Sengers et al., 2005, Sengers and
Gaver, 2006). These strategies lay the groundwork for the validating strategies of criticism.

6.1.4 Summary

This chapter has discussed how the pragmatist interaction design theory
comfortably adopts a range of validation strategies within a qualitative orientation.
Emphasis is placed on the role of interpretation. The theory relies on a multiplicity of
validation strategies that ultimately looks to open widely the interpretive space. In
explaining the role of validity in our pragmatist theory of interaction design, I discussed
and demonstrated the roles of strategies with the aims of securing knowledge across criteria
of trustworthiness, validity, and criticism. This supports the fifth proposition that
interaction design research is guided by qualitative and interpretive strategies to validation.

---

3 For a similar argument see Jeffrey Bardzell's blog Interaction Culture: Musings on interaction
design ad culture at http://interactionculture.wordpress.com/
The fifth proposition is that interaction design is guided by qualitative and interpretive orientations in validating new knowledge. The pluralism of pragmatism leads to the proposed theory promoting multiple strategies of validation that set out to be rigorous in interpretation and supportive of the practice of criticism, and inclusive of quantitative strategies where relevant.

The examples of trustworthiness show how credibility is achieved through prolonged engagement of the designer inquirer, collection and integration of diverse data, disinterested peer review in publications and interested peer review through collaboration, and negative case analysis, in order to establish the believability in the representation and findings generated by the research. Transferability is achieved through thick description and theoretical generalization. The descriptive auditing and other reflexivity mechanisms in the theory provide procedural support and a shared theoretical construct that together ensure dependability and confirmability of data and findings.

With respect to quantitative validation, the examples show how quantitative analysis and validation has been effective in different contexts: the internal validity check of a data collection instrument, the “mixed methods” approach that combines quantitative and qualitative analysis to support findings, and a quantitative analysis of video coding data. The underlying constructivism in qualitative research together with the multiplicity underlying pragmatism opens the theory to a broad use of validation strategies that includes quantitative validation.

Lastly I posited the importance of criticism in interaction design. This strategy of validation best assesses the value of interaction design inquiries beyond the inquiries
themselves into everyday experiences and over time. I described how criticism should be at home in a design discipline like interaction design as opposed to the contrary, and how blind peer review could be modified and leveraged into the beginnings of an interaction design criticism practice.