DTU Library

A foundational observation method for studying design situations

Cash, Philip; Hicks, Ben; Culley, Steve; Adlam, Tim

Published in:
Journal of Engineering Design

Link to article, DOI:
10.1080/09544828.2015.1020418

Publication date:
2015

Document Version
Peer reviewed version

A foundational observation method for studying design situations

Philip J. Cash[1]*, Ben J. Hicks[2], Steve J. Culley[3], Tim D. Adlam[3]

[1]Departement of Management Engineering, DTU, Denmark

[2]Department of Mechanical Engineering, University of Bristol, Bristol, UK

[3]Department of Mechanical Engineering, University of Bath, Bath, UK

*Corresponding author

To cite this article: Philip Cash, Ben Hicks, Steve Culley & Tim Adlam (2015): A foundational observation method for studying design situations, Journal of Engineering Design, DOI: 10.1080/09544828.2015.1020418

To link to this article: http://dx.doi.org/10.1080/09544828.2015.1020418

Journal of Engineering Design

Department Of Management Engineering,
Technology Innovation Management
Technical University of Denmark
DK-2800 Lyngby, Denmark
Tel: +45 45254550
pcas@dtu.dk

A foundational observation method for studying design situations

Observational studies of designers play an important role in engineering design research yet there is currently no accepted standard approach for comparing, combining or contrasting studies. Consequentially, reuse, reanalysis, replication, and aggregation of data are limited and the potential impact of individual studies is severely constrained. This paper begins to address this issue by introducing and developing a foundational method for observational design research to improve replicability, reuse, and overall comparability of empirical studies. A three-step foundational method is proposed that covers capture, coding, and analysis. The capture step defines overall and situational context as well as multiple capture streams, generating a broad dataset that can be examined from multiple perspectives. The coding step employs a multi-level approach that seeks to minimise workload whilst describing both detailed and high level information. The analysis step builds on the multi-level approach to provide for a flexible yet standardised examination of the dataset. The overall approach is introduced theoretically and illustrated using a comparison of an industrial study and an experimental study. Finally, it is argued that the proposed method promotes rigour, reliability, and standardisation; and could provide one means for improving comparison and aggregation, ultimately increasing impact in academia and practice.

1. Introduction

This paper develops a foundational method for observational design research with the aim to support the replication, reuse, and comparability of empirical design studies.

Design practice, and design activity specifically, have formed key foci in the design research domain (Cross 2007; Finger and Dixon 1989a; 1989b; Horvath 2004). Core to investigating these areas are observational approaches (Lethbridge, Sim, and Singer 2005), defined here as any approach directly recording the phenomena under study. These approaches aim to provide rigorous and robust characterisation of their subject. In the case of design this includes the practitioner, the process, the artefact, the environment, and wider context.

Observational approaches support theory building (Eisenhardt 1989; Briggs 2006), validation of experimental work (Bolton and Ockenfels 2008), and research impact (Glasgow and Emmons 2007). Fundamental to this role is the ability to bring multiple studies to bear on a single subject, triangulating results, accumulating significant samples, and varied complementary perspectives (Adelman 1991; Seale and Silverman 1997). Thus, it is critical that methods, data, and results can be compared, reused, and built upon. However, in design research there are currently no accepted standard approaches, frustrating such comparisons. For example, contrast Veldman and Alblas (2012) who systematically describe their cases, and Balogun (2006) who uses significantly different case descriptions. The lack of common structure or baseline means that despite similarities these studies are difficult to systematically compare.

There are two possible approaches to developing this basis for comparison. The first uses multiple identical studies, while the second aggregates multiple distinct, but related, studies. In both, standardised capture and analyse procedures can be used to support comparison. The first approach is well understood and thus this paper focuses on providing pragmatic support for aggregating studies. This has been directly inspired by Flay et al. (2005) who describe the positive impact of overlapping standards for methods and metrics in prevention research; and Brennan et al. (2011) and Kitchenham et al. (2002) who respectively propose methodological standards in policy research and software engineering.

The paper first explores the scope of existing approaches (Section 2). The foundational method is then proposed and illustrated via the comparison of an industrial and experimental case (Sections 3 – 7). Finally implications for design research are distilled (Section 8).

2. Background

This section explores current issues and observational approaches in order to develop a foundation for the proposed method.

2.1. Current Issues

An extensive review of design research literature previously undertaken by the authors established six core issues (Cash 2012). These are listed alongside supporting references in Table 1.

Table 1: Core issues affecting observational research

Nº	Issue	Supporting reference
1	Linking to theory	(Blessing and Chakrabarti 2009)
2	Effective contextualisation	(Adelman 1991)
3	Clear characterisation of the whole system	(Kitchenham et al. 2002)
4	Definition and reporting of the method	(Dyba and Dingsoyr 2008)
5	Mitigation of bias	(Kitchenham et al. 2002)
6	The lack of validation, replication, and critical analysis	(Dyba and Dingsoyr 2008)

When considering observational research, the six core issues generally manifest themselves as a number of practical problems. Table 2 provides further description of these problems, which are also discussed by the authors highlighted in Table 1. This explicitly focuses on problems related to method. In particular, characterisation of the system (Issue 3) is decomposed into sampling and research design, while mitigating bias (Issue 5) is split into reflexivity and data analysis.

Table 2: Specific methodological problems

Problem	**Description**
Linking to theory	Effectively fitting the work into the wider field and associated theory
Describing context	Characterizing context to support generalization and links to theory
Sampling design	Avoiding sampling bias to effectively represent the population
Research design	Designing and reporting the research to support replication and validation
Data collection	Avoiding bias and information overload whilst giving a rich dataset
Reflexivity	Managing the research/participant relationship to minimize bias and experimental effects
Data analysis	Minimizing bias while giving results that can be effectively interrogated
Value of findings	Defining the validity, nature and role of the findings in the wider context

2.2. Observational Approaches – Advantages and Limitations

There are many approaches to the characterisation of design practice (Lethbridge, Sim, and Singer 2005). The most common of which are summarised in Table 3, with references provided to exemplify each approach.

Table 3: Observational approaches for characterising practice

Approach	**Description**
Work diary	Participants report events either concurrently or reflectively e.g. Wild et al. (2010)
Work sampling	Participants report events as prompted e.g. Robinson (2010)
Applied ethnography	A combination of observation, interviews, and other studies e.g. Ball and Ormerod (2000)
Autoethnography	Focusing ethnographic techniques on the self e.g. Cunningham (2005)
Shadowing	Researchers follow the participant recording their activity e.g. Bergstrom et al. (2008)
Instrumented systems	Participant activity is automatically record via computer e.g. Lethbridge et al. (2005)
Fly on the wall	Participants record themselves using video or audio e.g. Cooper et al. (2002)

Through consideration of the core issues and the specific methodological problems the advantages and limitations of these approaches can be assessed. This is summarised in Table 4. The highlighted limitations all affect theory building (Issue 1) and system characterisation (Issue 3). Thus these are not listed to avoid repetition. Applied ethnography and autoethnography are included here because they have been specifically developed to be compatible with a realist approach making them suitable for this comparison.

Based on this assessment it is possible to imagine a combination of approaches that could reduce, or even eliminate, many of the limitations while maximising the collective advantages. There are two possible approaches for realising this combinatorial concept. The first is standardised selection, which seeks comparability and traceability by providing a common framework for choosing research, data capture, and analysis methods. This would use a weighting of advantages and limitations, derived from the fundamental research questions being asked and the specific situation under investigation to guide optimal selection. While a standard selection approach is feasible it is the second approach – a foundational method – that is considered here. A foundational method achieves comparability by systematically defining standard study elements to provide a common baseline.

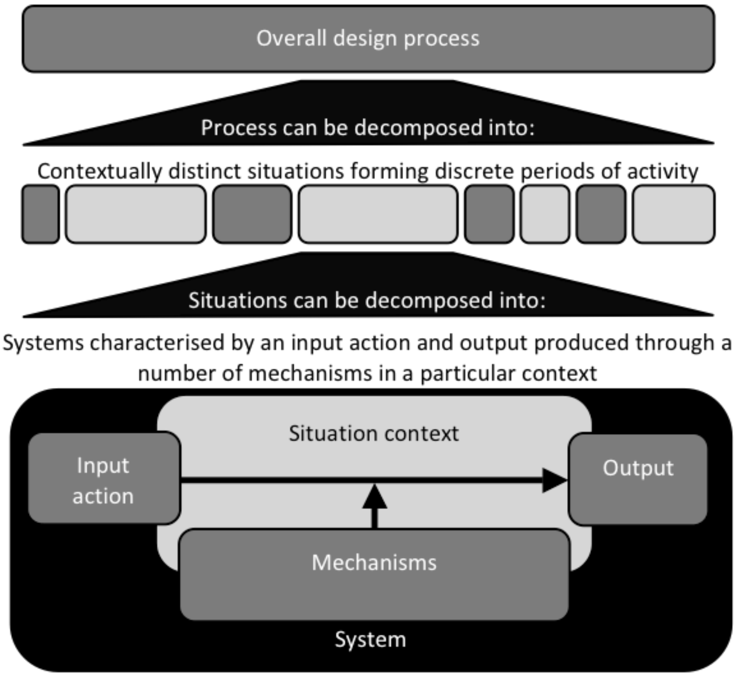
Table 4: Advantages and limitations of current approaches

Approach	**Advantages**	**Limitations**	**Relation to the core issues**
Work diary	Provides insight over a long period without incurring significant demands on the researcher	Difficult to account for bias introduced through self reporting or contextual information	Difficult to account for bias (Issue 5), validate, replicate or generalise (Issue 6)
Work sampling	Generates large amounts of data without incurring significant demands on the researcher	Difficult to account for bias introduced through self reporting or contextual information	Difficult to account for bias (Issue 5), can lack wider characterisation of the system (Issue 3)
Applied ethnography	Provides insight into practice and is not tied to a constructivist paradigm	Difficult to effectively report the full dataset and can be affected by bias	Difficult to account for bias (Issue 5), difficult to validate, replicate or generalise (Issue 6)
Autoethnography	Provides unique insight by making the investigator the focus of the study	Difficult to account for bias, typically of a limited sample size and scope	As above but can also be linked to Issue 3 due to the limited perspective
Shadowing	Can cover a wide range of attributes and requires no additional equipment	Difficult to account for bias and typically of a limited sample size	Issues 5 and 6 play a large role in studies of this type
Instrumented systems	Can provide accurate long term information on specific factors such as patterns of computer use	Difficult to address contextual information or effectively characterise the whole system	Difficult to effectively contextualise system use (Issue 2) and its relation to other work (Issue 1)
Fly on the wall	Unobtrusive and allows participants to acclimatise quickly with little disruption	Difficult to account for bias introduced through self reporting and limited scope	Issues 5 and 6 play a large role in studies of this type

3. Developing the Method

Defining the philosophical and theoretical assumptions underpinning the foundational method is critical to understanding its scope and applicability (Robson 2002). A *critical realist* perspective was selected for three main reasons. First, critical realism and post-positivism (closely related) dominate design research, allowing easier integration with current design research practice. Second, critical realism conceptually decomposes the system under investigation into core elements – *input action*, *output*, *mechanisms,* and *context.* Third, this decomposition can be used to define discreet situations. These two features support flexibility whist retaining commonality and are illustrated with respect to the foundational method in Figure 1. Key terms are defined as follows:

Figure 1: A critical realist perspective with respect to the foundational method



Situation: Contextually discreet periods of activity defined by stage of the design process, focus, general purpose, and setting.

Overall context: The world in which situations exist, defined by the overall process, company composition, and other generic elements linking all situations in a study.

Situational context: The general features of the activity (interactions, overall focus, and process stage) in a situation. The circumstances in which a situation takes place.

Input action: The impetus for change that results in an activity being undertaken e.g. a decision and move to a meeting room with the intent of brainstorming could be one input action leading to an ideation situation.

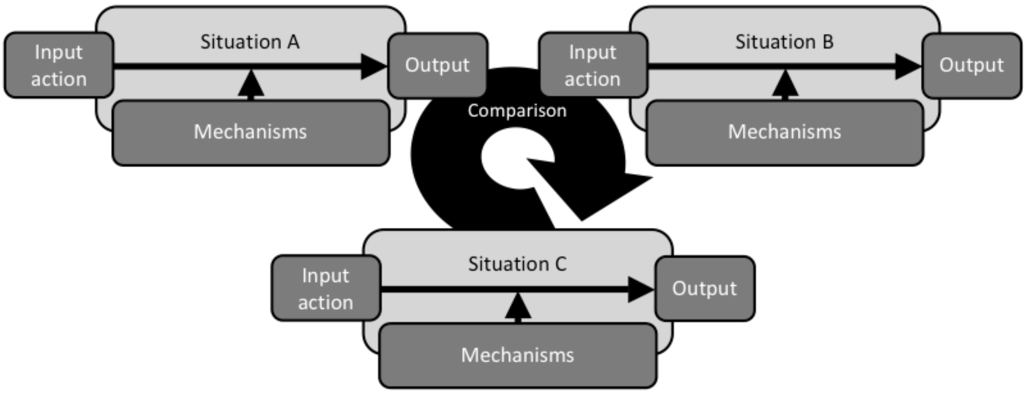
Mechanisms: The processes and variables through which the final output is achieved e.g. the cognitive processes of the designer, or idea iteration in a discussion.

Output: The measurable resultant features of a situation e.g. the number or quality of the ideas produced.

3.1. Standardisation Versus Flexibility

Two theoretical needs underpin the foundational method: identifying the elements necessary for robust comparison, and balancing prescription and flexibility. These are key to supporting effective standardisation without stifling new research approaches. Figure 2 depicts an idealised comparison where the elements in each situation are described in a standard way. This allows for direct comparison and triangulation without significant additional work.

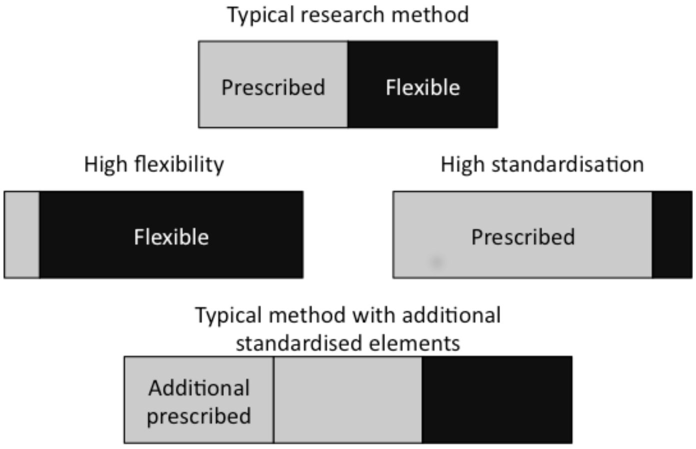
Figure 2: An idealised comparison using a situation-based framework



Comparing systems in this way allows deeper insight into the mechanisms, which underpins effective theory building: “If we understand nothing of the causal mechanisms, then we can only achieve a given outcome by accident at first and by rote thereafter” (Briggs 2006, 581).

Figure 3 illustrates how standardisation and flexibility can be balanced. Standard elements include common methods, metrics, selection systems or training regimes. With respect to the foundational method, standardisation is primarily associated with partial prescription of certain methodological features. Conversely, flexibility describes research specific elements including unique selection regimes, new metrics or new methods. These are brought together with the aim of maintaining scientific rigour. Mays and Pope (1995) state that the basic strategy for achieving rigour is: "*systematic and self conscious research design, data collection, interpretation, and communication*." Thus the ideal case would be standardisation without reduced flexibility or rigour. Here, Figure 3 shows a case where additional prescribed elements give a standard baseline without limiting the scope of the approach, even including grounded 'high flexibility' approaches. Although additional work is required to add the standard elements they do not constrain the construction of new methods, metrics or hypotheses. Hence this forms a standardised foundation upon which a range of methods or datasets can be built and compared.

Figure 3: The idealised role of a foundational method



3.2. Creating a New Method

In order to address the issues and the needs of standardisation/flexibility we take inspiration from key elements of existing approaches, including, Robinson (2010), McAlpine et al. (2011), and Wasiak et al. (2010). Specifically, the combination of capture sources described by McAlpine et al., the multiple perspectives on engineering work used by Wasiak et al., and the multi-level analysis proposed by Robinson. Further, we build on extensive prototyping of the approach.

3.3. Theoretical Framework and Terminology

Several aspects of the framework outlined in Figure 1 require further discussion. In particular, *overall context* has been developed from the work of Prudhomme et al.

(2007) and Visser (2009). Prudhomme et al. define a situation as including both the design process and other non-design activities. Visser defines a situation with respect to the process, designer, and artefact. Both of these conceptions are different from the situation as outlined in Figure 1, hence we have adopted the term 'overall context' to avoid confusion. Further, neither approach is suitable for linking all levels of the framework (Figure 1). As such, they have been combined and expanded in the foundational method.

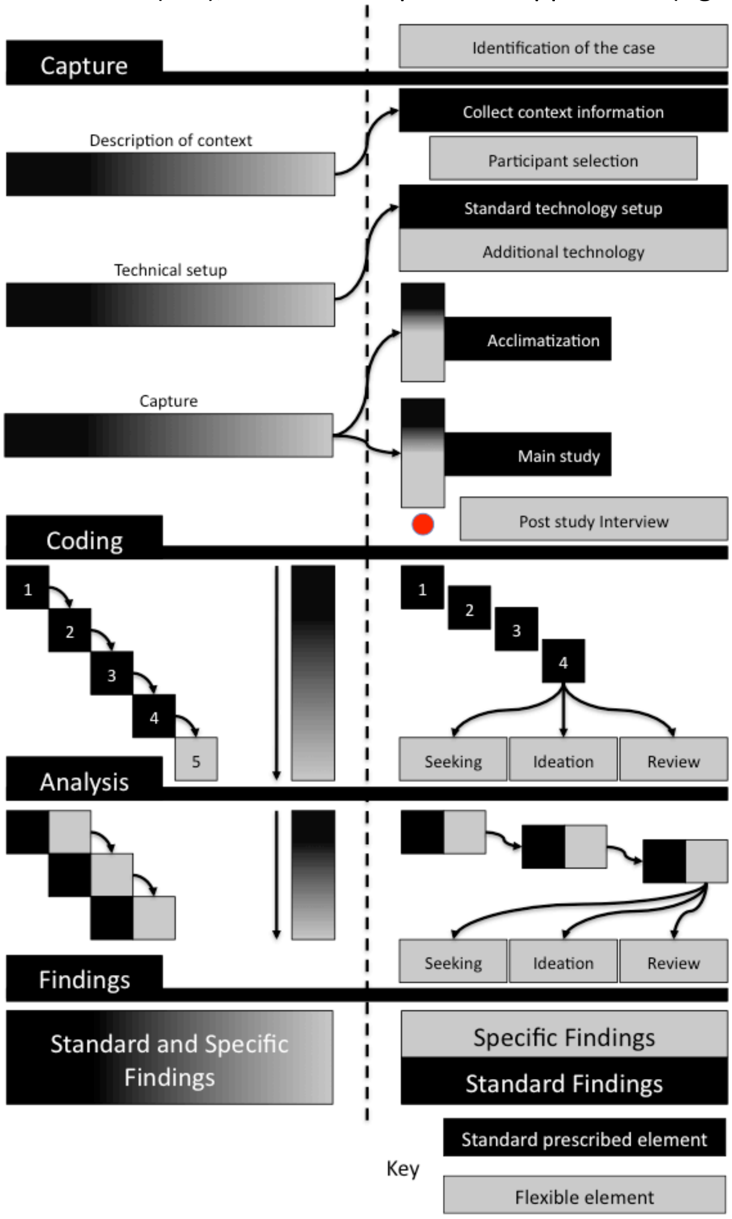
There are three areas to consider in balancing standardisation and flexibility. First, decomposing the observation period into discreet situations defined by common contextual factors allows for systematic description of a situation without constraining the scope of investigation. Further, by defining the granularity of the description it is possible to describe both the overall study and fleeting situations within it using the same spectrum of standardised comparison. This allows for studies at different levels to be compared in a common reference frame. Second, the standard conceptualisation of the situation (input action, output, mechanisms, and situation: Figure 1) gives a common core upon which to build comparability. Finally, *activity* can be used to form the basis for assessing the various situations. This is based on Activity Theory summarised by Bedny and Harris (2005): "*Activity is a goal-directed system, where cognition, behaviour, and motivation are integrated and organised by a mechanism of self-regulation toward achieving a conscious goal*." (p.130)

Operationalizing these concepts, the foundational method uses an integrated three-stage approach: capture (Section 4) – characterising the overall context, and providing the data for situation identification and investigation; coding (Section 5) – characterising the situational context, and providing a basis for detailed comparison; and analysis (Section 6) – exploring the situation with respect to the overall, and situational context. Although combining capture, coding, and analysis in one method is not in itself novel, each stage draws on unique elements that contribute to a more effective overarching method. This allows for research flexibility whilst maintaining standardisation and addressing the identified methodological problems.

3.4. The Foundational Method and Comparison Study

To illustrate each of these stages a comparison study is detailed in Section 7. Figure 4 shows the methods generic steps and links these to the specific work undertaken during the comparison study. It also illustrates how each stage has both standard and flexible elements. It is envisaged that in many cases, standard elements such as the capture strategy will overlap substantially with the specific demands of a particular study.

Figure 4: General method (left), and an example of its application (right)



4. Capture Strategy

There are three major aspects of the capture strategy: description of overall context, technical setup, and data collection.

4.1. Description of Overall Context

Context underpins generalisability and external validity (Kitchenham 1996; McCandliss, Kalchman, and Bryant 2003), and plays a critical role in comparison, reuse, and uptake (Shavelson et al. 2003). Thus this section outlines standard overall contextual factors, although it is expected that additional hypothesis specific factors be recorded as necessary (Section 3).

While overall context is important there are no widely accepted classifications. Terms commonly associated with 'context' are: activity, organizational, cultural, social, and historical (Klein and Myers 1999; Malterud 2001; Wildemuth 1993). When comparing these terms, it is apparent that organizational and cultural are related. 'Cultural' covers aspects associated with the participant including national and developmental background (Janssen, Van de Vliert, and West 2004), and is commonly described via the cultural onion metaphor (Gallivan and Srite 2005). Conversely, 'organizational' is used to describe company culture and thus forms one specific layer of the larger cultural onion. Further, 'activity' has a different definition from that explained previously as part of Activity Theory, hence we subsequently use 'technical environment'. Considering each aspect from a company and participant perspective four main areas emerge: technical environment, social, cultural, and historical. Practically, overall contextual information can be record either pre or post study depending on the specific research focus. The 'context first' approach given in Figure 4 is for illustrative purposes only. The full list of overall contextual features and metrics for each aspect are listed in the appendix.

4.1.1. Technical Environment

The technical environment influences what activities the participant undertakes and their potential modes of action. As such, the participants' environment needs to be characterised in order to establish the technical and structural influences on activity.

For example, a setting with only one meeting room and a densely populated open-plan office might produce an abnormally large number of informal meetings that could be misinterpreted if not properly contextualised. With respect to the participant, the bulk of activity is likely to involve either their personal computer or logbook (McAlpine et al. 2006), dictating that the use of these systems/tools be recorded in a structured manner.

The standard features recorded in the foundational method are the technical layout and resources in the workspace: the physical distribution of the participant(s), other workers, and the overall layout of the working environment; the distribution of working time between the primary workspace and other areas e.g. the home or workshop; and the technical affordances of the space likely to affect activity e.g. the distribution of whiteboards and other equipment.

4.1.2. Social

The factors required for baselining a participant population are measured using socioeconomic status. This has a number of established variables that are used across fields (Adler and Ostrove 2006; Pickett and Pearl 2001). These variables aim to give insight into factors such as social norms (Levitt and List 2007), social status (Jakesch et al. 2011), independence, and interests (Shalley and Gilson 2004). The standard measures used for this are summarised in the appendix.

At the organisational level, there are also a number of comparative factors. Those associated with the social context of the organisation (i.e. affecting job complexity, demands, challenges, and autonomy (Shalley and Gilson 2004)) include: funding level, income source, market pressures, environmental factors, other monetary pressures, and the composition of the organisation's population.

4.1.3. Cultural

Cultural factors have two major aspects considered here, the national cultural background of the participant and the specific organisational culture. Petre (2004) highlights the effect of both these aspects on practitioner behaviour. With respect to national culture, cultural distance measures are well established (Shenkar 2001).

These can be used to generally define the participant population (Kogut and Singh 1988; Dow and Ferencikova 2010), including elements such as collectivism/individualism and group homogeneity (Janssen, Van de Vliert, and West 2004; Shalley and Gilson 2004). These factors are operationalized by Hofstede et al. (2010).

With respect to organisational culture key factors include hierarchy, level of formality, level of socialising, and overall homogeneity (Guzzo and Dickson 1996; Stewart 2006). Factors specifically related to engineering design include: pride in quality of work, competitiveness, type of design work (Wild et al. 2005), organizational aims or areas of support (Janssen, Van de Vliert, and West 2004), and existing projects and practices (Lewis and Moultrie 2005). This is again operationalized using Hofstede et al.'s (2010) cultural measures for organisations.

4.1.4. Historical

Historical factors typically manifest indirectly via the current social or cultural context. As such, there is little to directly assess in this aspect. However, two relevant areas are annual turnover and market maturity. These play a confirmatory role, complementing the factors recorded in the social and cultural areas. In terms of the participant, the key historical factor is their previous experience and knowledge (Shalley and Gilson 2004; Jakesch et al. 2011).

Ultimately numerous variables affect the outcome of a study, however, those highlighted here form a core set of recognised variables necessary for defining a study and its population. Both in a general sense and in relation to engineering design specifically. These standard parts support generalisation without demanding deep research into contextualisation, which comprises a research area in its own right.

4.2. Technical Setup

The standard aspects of equipment selection and setup were based on the work of McAlpine et al. (2011) who assess a range of capture technologies against their level of coverage and data collection/analysis demands. Here, the use of multiple capture pathways allows for the wide variety of situations likely to be encountered in practice. This also partially mitigates recording limitations often imposed in an industry setting by providing a rich record of those periods where data capture is permitted.

The standard technical setup guides equipment distribution with respect to the generic aspects of engineering activity e.g. the workstation and logbook. However, the details of placement are strictly situational, being based on the participants' perception of their working practice (Section 4.1.1). Thus, this provides a standard foundation while research specific additions address the needs of within study validity and insight. To give an effective foundation for reuse and generalisation the standard setup is designed to capture the widest range of possible activities. Table 5 outlines these standard capture elements and how they overlap. This overlap is important for synchronisation, providing redundancy, and allowing triangulation during analysis (H. Robinson, Segal, and Sharp 2007; Seale 1999).

Table 5: Standard capture elements and relevant technical approaches

Perspective	**Approach**	**What it is recording**	**Further information**
Participant	Synchronised camera 1	Front view of participant – high resolution, synchronised with other cameras	www.panopto.com and standard HD web cameras (Panopto 2012)
Workspace	Synchronised camera 2	Wide view of main workspace – audio and video synchronised with other cameras	
Detail of PC work	Synchronised screen capture	Live screen recording – high resolution, synchronised with cameras via e.g. panopto	www.panopto.com
Overall PC usage	Long term data logging	Automatic recording of computer usage – usage, documents and applications	www.manictime.com (ManicTime 2011)
Participant view	Mobile camera	Participants view of situations away from the work station	e.g. Looxcie head mounted camera
Written notes	Recording of logbook	Participants notepad use and audio – writing and audio playback of logbook	www.livescribe.com (LiveScribe 2011)
Participant background	Work diary	Participant records activities not otherwise captured in structured form	Questionnaire e.g. Robinson (2010)

With respect to engineering work each aspect is covered by at least two complementary techniques, given in Table 6. The engineering work activities are taken primarily from Hales (1987), Robinson (2010), and Austin et al. (2001). Key non-technical issues are those of privacy, ethics and confidentiality. Again, the standard capture setup has been designed to mitigate these concerns where possible through redundancy (Table 6). For example, long-term data logging allows a portion of the data to be anonymised on collection rather than after coding. Here, using overlapping, linked recording mechanisms (e.g. multiple synchronised cameras) allows the participant to manage the recording process simply and transparently.

Table 6: Summary of engineering activities and the associated approaches

Engineering activities	**Approaches**	**What is captured**
Collocated meetings and collaboration	Recording of logbook	Meeting notes and audio of conversation
	Mobile camera	Audio and video from the participants perspective
Written communication	Synch. screen capture	E-mail and other messaging activity via computer
	Work diary	Other messaging activity
Distributed communication	Synch. cameras	Audio and visual of phone or computer use
	Synch. screen capture	Computer based video conferencing
Individual design work	Recording of logbook	Personal note making/working
	Long term data logging	Overview of computer usage
	Synch. screen capture	Detail of work carried out on computer
Project management activity	Long term data logging	Overview of computer usage
	Synch. screen capture	Detail of work carried out on computer
Participant detail	Synch. camera 1	Visual of participant demeanour
	Synch. camera 2	Audio and visual participant demeanour
Other	Work diary	Identifies events not otherwise recorded

4.3. Data collection

Data collection is split into three phases; acclimatization, study, and post-study. The standard setup and overall approach aims to minimise researcher/participant interaction throughout this process for two main reasons. First, this reduces the impact of the standard elements on the research specific aspects of the study. Second, minimising interaction reduces experimental effects. Essentially, the act of studying human subjects has a number of effects on their behaviour irrespective of study type (Kazdin 1998). These effects have many specific names and mechanisms of action (Holden 2001; Falk and Heckman 2009) but are generally referred to as Hawthorne type effects using the broad definition: *"… The problem in field experiments that subjects' knowledge that they are in an experiment modifies their behaviour from what it would have been without the knowledge.*" (Adair 1984, 334). These effects can have a significant impact on participant behaviour and must be accounted for either in the study design or through control and normalisation (Diaper 1990; Cook 1962). In observational or descriptive studies this is achieved by minimising researcher/participant interaction (either through reduced contact, blinded research design or preferably both) and acclimatisation. Acclimatisation allows the participant to return to as close to normal behaviour as possible before starting the study. Although this has its own affect on the study (Adair 1984), acclimatisation has been shown to be key in reducing the influence of experimental effects (Podsakoff et al. 2003).

4.3.1. Acclimatisation phase

Acclimatisation serves several purposes in the foundational method:

1. It minimises experimental effects although further study is needed to validate this in the context of engineering design.
2. It allows participants to become accustomed to the research equipment and procedures, such as, the recorded logbook (Table 6). Two weeks was considered the minimum for allowing these to become habit based on McAlpine et al. (2011). In making the research procedure habitual the participant requires less direct monitoring, reducing interaction.
3. It allows the researcher to customize the standard technology setup, integrate any specific elements required, and address any issues raised by the participant. This includes checking the equipment and preliminary data – reducing problems/data loss during the study.
4. It allows the researcher to gather participant feedback on the perceived effectiveness of the capture strategy. Such, reflective feedback is a key tool for improving rigour (H. Robinson, Segal, and Sharp 2007).

It is suggested that participants undertake at least three weeks of acclimatization prior to the main study. This has been shown to be sufficient acclimatisation for the normalisation of Hawthorne and other effects (Leonard and Masatu 2006; Barnes 2010; Podsakoff et al. 2003). However, this can be extended or reduced and validated depending on the specific research design. In all cases the participants should record data and behave as they would during the main study, with the researcher checking the collected data for completeness at regular intervals. When the acclimatization period is very small in e.g. experiments or scenario based studies, control groups or other means should be used to account for the experimental effects. Overall (using acclimatisation or not) participant contact should be minimised and double blind designs used as best practice wherever possible.

4.3.2. Study phase

With acclimatization complete the study phase should start immediately – lasting as long as required for the specific research aim. Before the study starts each participant is given the opportunity to talk through any remaining issues/questions. However, during the study itself participant/researcher interaction should be limited (Section 4.3.1). This minimisation is explicitly designed into the standard setup and overall method, with data collection automated where possible. At this point it is sufficient to recommend that researchers consider this when developing the research specific elements as further constraint could potentially limit the scope of possible research.

4.3.3. Post study Phase

Post study reflection – both immediately after the study, and with respect to the final analysis – is an important part in validating the completeness and accuracy of the data and findings (H. Robinson, Segal, and Sharp 2007). When applied after data analysis post-study interviews can be used to assess the validity of the results and conclusions, as well as to reflect on the success of the method itself. The foundational method employs semi-structured interviews to explore the following factors and conclude the study:

- It allows the researcher to check if the participants' perceived their working practices to have been in any way unusual during the study.
- It allows the researcher to check that participants were still hypothesis blind where appropriate.
- It allows participants to explain/expand on any incidents reported in the work diary and relate any issues or unrecorded events encountered during the study.
- It allows participants to provide one type of validation with respect to the conclusions drawn from the analysed data.

5. Coding Strategy

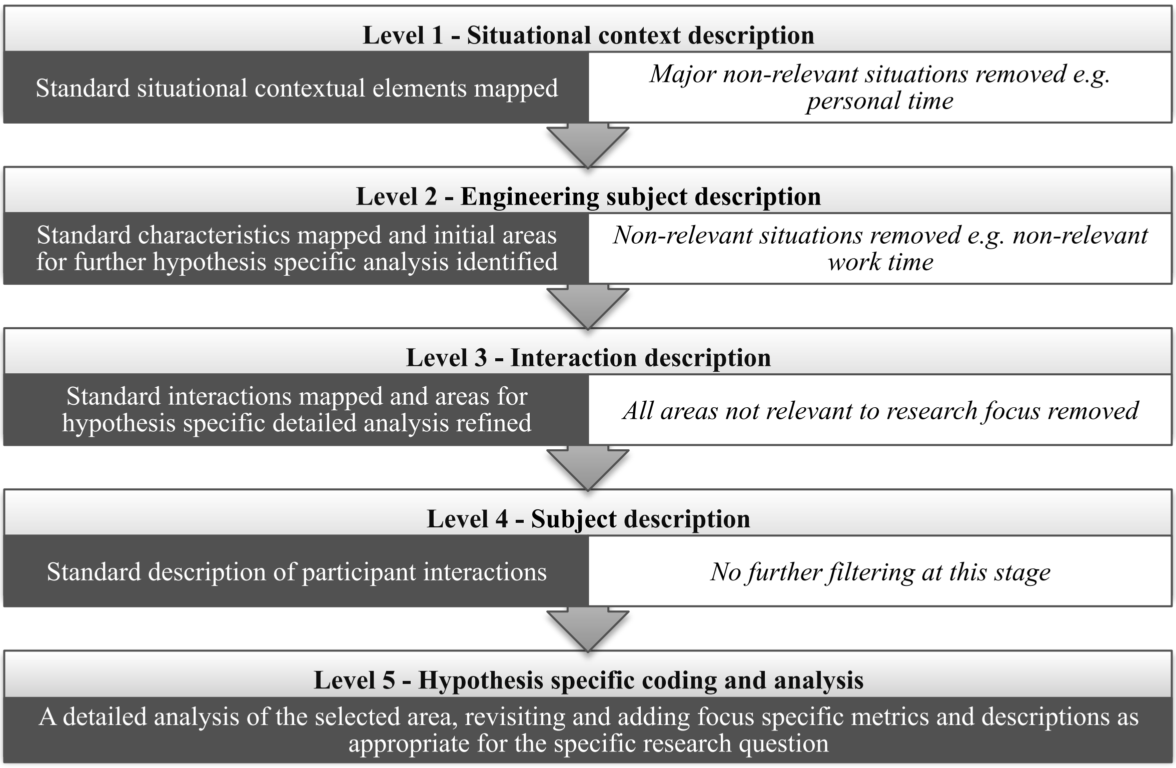
The foundational methods multiple capture streams generate a large amount of data. It is emphasised that not all of this information need be immediately utilised, instead it forms the foundation for varied, multi-perspective reuse and reanalysis. As such, a streamlined approach is necessarily adopted in the proposed coding strategy, minimising workload whilst supporting comparison. This is realised using a multi-level strategy that facilitates the rapid narrowing of scope. Systematically linking specific situations with the wider information contextualising such periods.

5.1 Multi-level coding and analysis strategy

The multi-level coding strategy consists of five levels of increasing detail. The sequential levels act as filters, isolating periods that the researcher does not wish to explore further. This ensures that both the wider context and the detail required for the specific research aim are treated without overloading the researcher. Thus it is possible to describe the entire data corpus at Level 1, and then subsequently narrow the scope by removing less relevant elements – as dictated by the researchers focus.

Figure 5 outlines the five levels, describing the focus and the filtering strategy at each level (filtered elements are italicized). Each level guides the selection of data to be coded at the next level, thus reflection on what should be removed at each stage is essential to the strategy's effectiveness. This is conceptually linked to Activity Theory, which describes a system where discreet periods of activity are defined using sequential levels of increasing detail, down to unconscious operations (Bedny and Harris 2005). The five levels proposed by the coding strategy complement this model. Here the situation in which an activity is taking place is described at various levels of detail. This supports comparison at any of the specified levels without prescribing or restricting the investigation of the hypothesis specific activity. In the context of engineering design, five levels have been defined and are considered to provide an appropriate balance between resolution and workload. More levels were considered excessively prescriptive whilst offering little further generalisability.

Figure 5: Multi-level coding and analysis strategy



5.2 Coding

The multi-level coding has been designed to maximise its practicability whilst also giving maximum benefit to the researcher and the wider field. Only Level 1 is applied to the whole data set, with subsequent levels being applied to increasingly limited time periods. Further, the multi-level contextualisation of the final period (defined by the specific hypothesis) explicitly supports and promotes the triangulation of different studies, data, and approaches. Finally, the generality of the codes make them ideal for characterising a broad range of design situations, whilst also being applicable to a variety of research specific foci.

This has been achieved by the levels being designed to fulfil the key requirements for understanding and contextualising activity (defined by Activity Theory). Bedny and Harris (2005) describe two key characteristics for describing activity: object (a tool or material object that the subject or subjects interact with), and subject (two or more subjects are characterised in terms of information exchange, personal interaction, and mutual understanding). Combining this with the context discussion (Section 4.1) four areas emerge for defining an activity:

Situational context – the immediate work environment, the type of interaction being undertaken, and the participants' focus in terms of the generic engineering design process (Hales 1987). This also reflects a distinction between object and goal as discrete aspects of activity (Bedny and Harris 2005).

Engineering subject – the specific engineering design characteristics of the exchange between subjects: problem solving and information exchange. These have been established as applicable in the engineering domain by Wasiak et al. (2010) and Blandford and Attfield (2010). Here they have been adapted for generalisability by reflecting on the underpinnings of Activity Theory.

Interactions – the object(s) forming the primary focus of the activity, both individual and group. This has been generalised based on Cash et al. (2010).

Subject – the characteristics of exchanges between subjects: type of information exchange, personal interactions, and mutual understating (Bedny and Harris 2005). These have been based on the works of Horvath (2004) and Wasiak et al. (2010), and have again been generalised with regard to Activity Theory.

Table 7: The four levels of standard codes

Level 1 Situational context			
Group	**Nº**	**Code**	**Code options**
Interaction type 1	1	Individual/ group	0 - individual, 1 - group
Interaction type 2	2	Synchronous/ asynchronous	0 - synchronous, 1 - asynchronous
Interaction type 3	3	Co-located/ distributed	0 - co-located, 1 - distributed
Environment	4	Location	0 - normal, 1 - other

Focus 1	5	Design process stage	1 - brief creation, 2 - feasibility, 3 - design development, 4 - manufacture, 5 - testing, 6 - reporting, 7 - other
Focus 2	6	Focus: people / product / process	0 - other, 1 - people, 2 - product, 3 - process
Level 2 Engineering subject			
Group	**Nº**	**Code**	**Code options**
Problem solving	7	Goal setting	0 - not goal setting, 1 - goal setting
	8	Constraining	0 - not constraining, 1 - constraining
	9	Exploring	0 - not exploring, 1 - exploring
	10	Solving	0 - not solving, 1 - solving
	11	Evaluating	0 - not evaluating, 1 - evaluating
	12	Decision making	0 - not decision making, 1 - decision making
	13	Reflection	0 - not reflecting, 1 -reflecting
	14	Debating	0 - not debating, 1 - debating
Information exchange	15	Recognising need	0 - not recognising need, 1- recognising need
	16	Interpretation	0 - not interpreting, 1 - interpreting
	17	Validation	0 - not validating, 1 - validating
	18	Seek/ request	0 - neither, 1 - seeking, 2 - requesting
	19	Using information	0 - other, 1 - informing, 2 - clarifying, 3 - confirming
Management exchange	20	Managing	0 - not managing, 1 - managing
Level 3 Interactions			
Group	**Nº**	**Code**	**Code options**
Audiovisual	21	Audio only	0 - not interacting with X, 1 - interacting with X
	22	Visual only	
	23	Audiovisual	
Documentation	24	Formal	0 - not interacting with X, 1 - interacting with X formal/informal split defined by Hicks et al. (2002)
	25	Informal	
Physical	26	Environment	0 - not interacting with X, 1 - interacting with X
	27	Tools	
	28	Design representations	
Level 4 Subject			
Group	**Nº**	**Code**	**Code options**
Type of exchange	29	Opinion/ orientate/ suggest	giving or receiving: 0 – other, 1 – opinion, 2 – orientation, 3 – suggestion
Understanding	30	Agree/disagree	showing: 0 – other, 1 – agreement, 2 – disagreement
Personal 1	31	Antagonism/ solidarity	giving or receiving: 0 – other, 1 – antagonism, 2 – solidarity
Personal 2	32	Tension/ tension release	showing: 0 – other, 1 – tension, 2 – tension release

In order to characterise each area, codes are defined over four sequential levels, summarised in

Table 7 and defined in the appendix. Each level is split into groups for clarity. Within each group codes are mutually exclusive. Level 5 is flexible and is thus not included in

Table 7.

6. Analysis Strategy

The intent of the analysis strategy is not to fully analyse all the data captured and coded in Sections 4 and 5. Instead, analysis is again tackled in sequential stages. This avoids overloading the researcher whilst maintaining traceability and analytical rigour. In order to achieve this result there are a number of standard steps required to ensure rigour and completeness: alignment, layered analysis, and reflection.

First, the data sources need to be aligned on a single persistent timeline as emphasised by Torlind et al. (1999; 2009). For maximum benefit both standard and research specific sources should be aligned to a common timeline. This allows the researcher to increase the potential of complementary data sources in three ways:

- It allows gaps in one source to be filled by another e.g. using mobile camera footage to follow the participant when they leave their desk – developing a more complete record.
- It allows multiple coded sources to be compared for a single event e.g. the coding for the participant's logbook could be compared to the coding for the camera in order to refine the final result – developing a more rigorous record.

- It forms a better foundation for generalisability, replication, and reuse by relating the standard and flexible elements of the study to a single core unit – in this case the common timeline.

Synchronisation and alignment requires a core timeline for consistency. For example, using the standard record of the computer screen (Table 6) as a master timeline in VCode (Hagedorn, Hailpern, and Karahalios 2008) (or similar annotation tools) all other sources, both standard and flexible, can be combined. Although the primary source is not prescribed, and need not be one of the standard sources, it is recommended that the selected source is the most individually complete and comprehensive – minimising additional combinatory work. Further to the methodological advantages of combining the sources onto a single master timeline this also streamlines the analysis, export, and comparison tasks. With the various data sources aligned it is possible to start the analysis. The foundational method utilises three levels of detail and complexity.

The first and least complex level is the high-level quantification of the standard codes. This can include the total time each code accounted for, the number of instances, and overall trends. This high-level analysis follows the same approach and structure as outlined in

Figure 5 i.e. analyse codes level by level, sequentially omitting areas not of interest as required. This allows for a standard baseline to be created, against which other studies using the foundational method can be compared.

Second, with the high-level analysis complete consider groupings of related standard codes. This level can be used to draw out deeper comparisons and to define more complex activities or situations. For example, using a combination of standard codes to describe a key situation allows for the subsequent identification of similar situations in other datasets utilising the foundational method. As such, this provides the basis for multi-perspective examination and triangulation. This again allows pattern, frequency, total time or other aspects to be analysed for each group of codes. Groups are identified based on the following standard steps; illustrated using the comparison study (Section 7) as an exemplar:

1. Develop descriptive definitions of areas of interest – in this case tasks within the engineering design process as defined by Hales (1987).
2. Allow groups of codes to emerge from the data for the defined areas of interest (this can include multiple groupings) – In this case, conceptual design is defined using six combinations of codes. For example, two groups are: *'group', 'design dev', 'focus – product', 'exploring'*, referring to a group brainstorming activity; and *'individual, 'design dev', 'focus – product', 'exploring'* referring to an individual ideation activity.
3. Reflect on the allocation of the groups of codes to ensure that the selected definitions are appropriate and further definitions do not need to be considered for the given research focus. This is an important step as there can be large numbers of combinations for a single definition (depending on the code level to which the groupings are defined).

Third, the standard codes can form the basis for the detailed analysis if they are considered sufficient for the research specific focus of the study (Level 5).
Once the analysis is complete it is necessary to reflect on the validity, reliability, and limitations of the data. However, as the focus of the foundational method is on supporting replication and comparison rather than explicitly addressing internal validity, the means by which the researcher establishes these parameters (validity, reliability, limitations etc.) is flexible. With respect to the foundational method it is sufficient to establish that the information that has been coded is representative of the data. As such, appropriate inter-coder reliability checks should be undertaken.

7. A Comparison Study

In order to explore the utility of the foundational method in the context of design research the comparison study explores two perspectives. These perspectives are necessary in order to more fully assess the ability of the method to support comparison between studies that are related but distinct – one of the primary motivations for this work. Both perspectives are supported by the comparison of an industrial study and an experimental study (the same studies are used for both perspectives).

The first perspective considers the method's potential for comparing and triangulating studies of different formats. Here the format is considered to include the setting (laboratory, intermediary or practice), the population (e.g. student or practitioner), and the type of approach used (fully contrived experiment, quasi-experiment, practice based case etc.). The second perspective considers the method's applicability to varied research foci. Here the research focus is considered to mean the main subject of investigation as well as the overall aim of the research (e.g. descriptive or prescriptive). Format and research focus were considered the key differentiators between studies and thus selected as the basis for the comparisons outlined in this section. Further, the comparison between the reported studies constitutes a real research need. Specifically, the research aim was to compare design activity in a range of design situations across research settings. However, this is not the focus of the work here and has been fully reported elsewhere in Cash et al. (2013). A summary of the two studies has been included below in order to contextualise the comparison.

Summary of Industrial Study

The population (seven designers in 18 staff) was introduced to the research through a series of introductory meetings and a sample size of three was selected. Three was considered appropriate because the company used a matrix structure where all engineers work across multiple projects. Participants were then asked to volunteer due to the intensive nature of the observation. Five volunteered and three were selected randomly, representing each seniority level i.e. junior, midlevel, and senior.

Each participant was then observed for four weeks total (three weeks of acclimatisation and one week of study). Only one participant was observed at a time and there was no overlap in terms of projects or collaboration – each participant worked independently of the others. In terms of technical setup the full suite of capture approaches was used as described in Section 4.2. This resulted in a total of three weeks of study data amounting to circa 300 hours of video (due to the multiple sources) and circa 100 GB of data after compression. Participants were observed, coded, and analysed in a random order, with all data sources synchronised.

Summary of Experimental Study

The population for this study was selected from a group of 40 final year masters level engineering students. From this population, twelve students were selected and randomly assigned to one of four teams. All participants were given creativity and sociometric tests in order to form a baseline.

Each team was then given the following brief: "*You are to design a universal camera mount for use on an aerial vehicle. The aerial vehicle is to be used by an amateur photographer, primarily to take still photos.*" In completing this brief the session was split into four phases: (1) 50 minutes individual information seeking, (2) 50 minutes group ideation, (3) 90 minutes individual design development, and (4) 50 minutes design review.

7.1 Perspective 1: Study Format

The objective of this comparison is to verify the applicability of the foundational method across a range of formats as well as to establish its utility in the comparison and aggregation of data across related studies. Both studies were carried out based on the foundational method with adaptations for the specific setting and research aims.

For the industrial study the situations described fully at Level 4 are contextualised by the preceding levels and can be defined in terms of combinations of codes. The sub-aim of the industrial study was to identify and characterise key design situations, such as, ideation and design review meetings, contributing to the overall comparative aim described above. As such, the coding schema allowed for the rapid narrowing of scope while retaining the overview of the whole study period.

This progression can be illustrated using the industrial study example. First, the majority of the study (240 hours – see Table 8) was coded coarsely at Levels 1 and 2 (average time per code = 680 seconds compared to Level 4 where the average time per code = 17 seconds). Results were tabulated and a Visual Basic script used to automatically identify and list all the unique combinations of codes. This resulted in 147 combinations, which could be described as unique situations. Next, the combinations were grouped by defining key features for further analysis (described in detail in Section 7.2) and confirmed by checking the identified combinations against the original footage. This example led to ten high-level situation types, with the design related situations summarised in Table 9. Finally, an automated check was used to reveal any periods that could be assigned to more than one overall group (14 in total), which could then be rapidly identified and clarified.

Conversely, the experimental study was defined in detail by its associated research aims (see Section 7.2) and could thus be immediately characterised by the standard coding elements. To elaborate, the higher levels of the coding strategy were predefined or highly limited by the research aim and were therefore used descriptively. Further, as the higher levels could be coded rapidly specific codes could be examined with little additional effort.

Table 8 describes how the foundational method was applied to each of the studies. This highlights how it can be adapted, streamlined, and applied to different contexts whilst retaining comparability. For example, overall participant context was captured for both studies – allowing for comparison of populations. In the industrial study this information was captured prior to the study using questionnaires as the aim was descriptive and thus hypothesis blindness was not an issue. In the experimental study overall participant context was recorded post study to avoid biasing the participants and was complemented by specific tests required for the research aim.

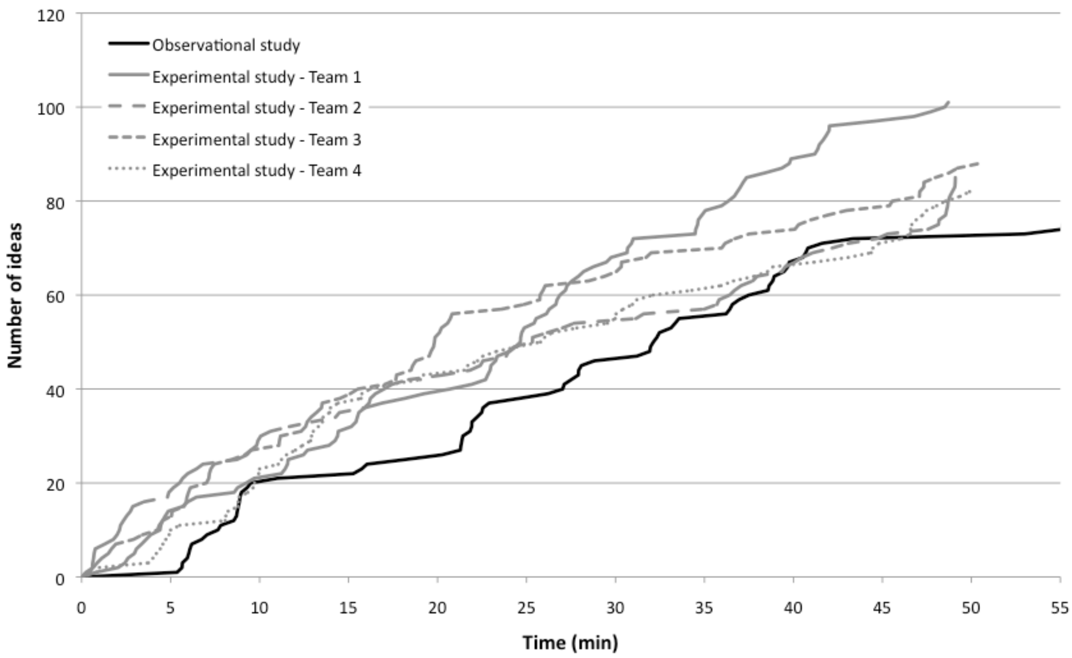
A key feature of the foundational method, highlighted by Table 8, is that the experimental study can be immediately and directly related to similarly contextualised situations from the industrial study. For example, periods of ideation from the industrial study could be explicitly identified and compared to ideation during the experimental study. This is born out when the data from the experimental and industrial studies is compared. An example comparison is shown in Figure 6, which shows the similarity in ideation between the industrial study and the experimental study.

Table 8: The two studies in relation to the foundational method

Foundational method	Industrial study	Experimental study
Overall context		
Personal	Carried out prior to the study using questionnaires as no fixed hypothesis	Carried out post study to maintain hypothesis blindness using questionnaires and other tests for hypothesis specific information
Wider population	Carried out prior to the study using interviews with company management	Carried out independently based on available data from the host university

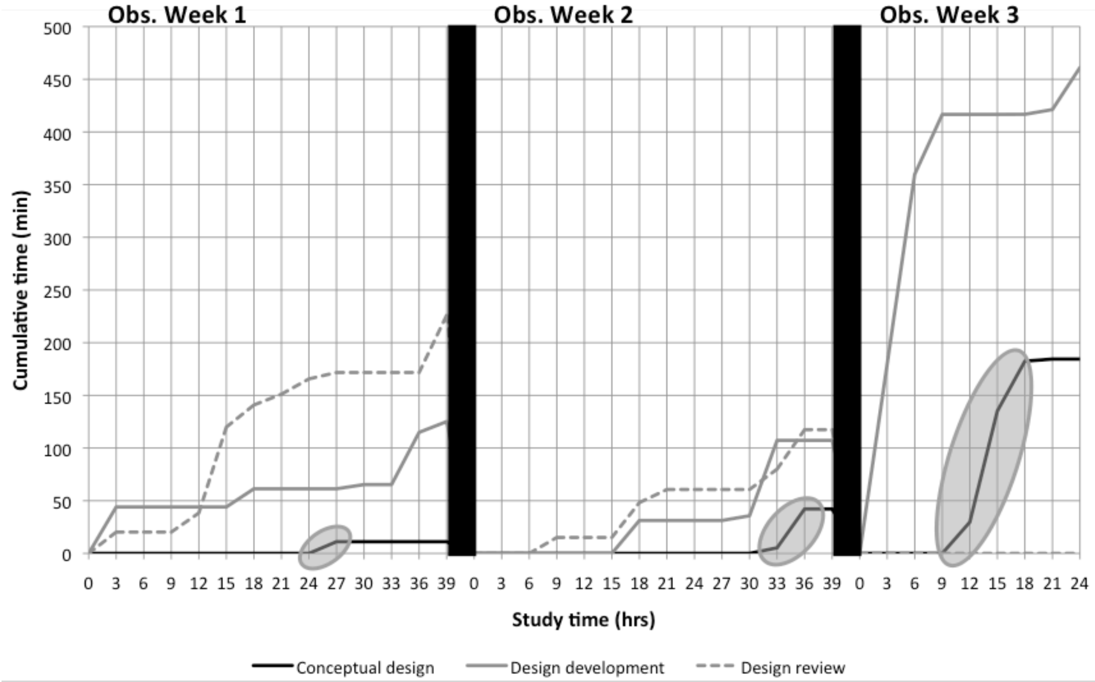
Technical setup	As prescribed	As prescribed but forgoing mobile cameras due to the restricted setting
Capture		
Acclimatisation	Three weeks for each participant to minimise effects	None due to the study design, instead control groups could be used
Study	One week per participant with full freedom (98 hours total for the three)	Four hours with each experimental team – predefined group and individual work
Post study	Interview assessing the data, and reported work of the participant	None
Coding		
Level 1	100% of time coded at this level	Specified by the study design thus not coded
Level 2	80% of the time coded at this level	Guided by the study design, only a selection of codes were encountered (4 of 14 Level 2 codes used)
Level 3	Focus reduced to group work with a focus on the product: 34% coded	Guided by the study design, only a selection of codes were encountered (2 of 8 Level 3 codes used)
Level 4	Specific situations: only 4.2% coded: one ideation, one information seeking and one review situation (250 min)	Coded fully for each of the studies
Specific	None originally – then specific codes from the experiment applied situations in the industrial study	Additional codes added for ideation, information seeking and design review based on the research questions
Analysis		
Synchronisation	As prescribed, using the participant camera as the central timeline	As prescribed, using the participant camera as the central timeline
High level	Individual codes used to describe overall design activity and process	Level 1 used to compare experimental context to industrial study
Groupings	Groups of codes used to describe specific situations for comparison – ideation, seeking and review	Guided codes (Level 2 and 3) used to link to the specific situations observed in practice
Detailed analysis	Specific codes analysed and then applied to the identified analogues situations from the industrial study – ideation, seeking and review	
Reliability	Cohen's Kappa applied to check inter-coder reliability	Cohen's Kappa applied to check inter-coder reliability

Figure 6: Ideation in the industrial study and in the experimental study



By enabling this comparison the method allows for an improved assessment of the likely impact of findings from the experimental study in practice. Here, the features of the experimental study (e.g. information seeking) can be explicitly mapped to similar periods in practice. The potential impact from changes in these periods can then be assessed holistically. Figure 7 highlights three examples of such mappings, which are denoted by the circular overlays. These are explicitly and directly linked to the situations described in the experimental study. Specifically, Figure 7 shows information seeking, design development, and design review activity in the industrial study (based on Levels 1 – 4). This shows how these isolated regions can now be assessed with respect to their role in the wider process.

Figure 7: Identifying relations between the studies



7.2 Perspective 2: Research Focus

In this context four different research foci were considered: the overall design process, information seeking, ideation, and design review. The objective of this comparison is to verify the flexibility of the foundational method in supporting a range of different research foci.

7.2.1 Design process

The industrial study was used to assess the ability of the foundational method to support a purely descriptive focus. Here, the aim was description of the design process based on the standard analytical steps (Section 6). First, the individual codes allowed for a raw assessment of the types of work undertaken using the total time spent on each activity e.g. product, process or people, and the design phase.

Second, combining the codes allowed for a more nuanced description of the design process and participant activity. With respect to the example of information seeking it allowed for the whole range of information behaviours characterised by Robinson (2010) to be described in terms of combinations of codes. This resulted in approximately 45% of the participants' time being associated with information seeking activities of various types. This closely, links to other estimates of information seeking in the extant literature (Robinson (2010) – 56%, King et al.

(1994) – 40-60%, Puttre (1991) – 32%, and Cave and Noble (1986) – 30%). This suggests that the combination of standard codes was in fact sufficient to fully represent this specific research focus. An example of a combination of standard codes used to describe one type of information seeking activity is (the number of the relevant code is given in brackets (Table 7)): Individual (1), distributed (3), feasibility stage (5), product focus (6), solving (10), and requesting information (18). In this case the standard coding could allow Robinson or others to reanalyse the data with respect to their own work without significant recoding effort. Instead reanalysis is achieved either by defining combinations of codes or by identifying areas of interest and then recoding them specifically.

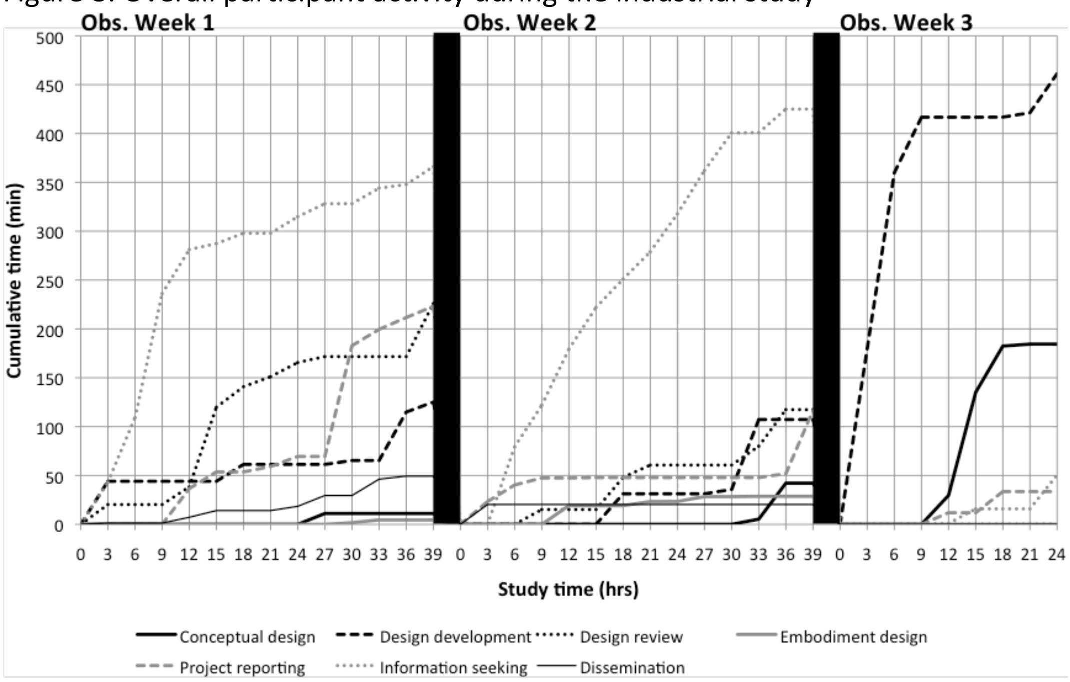
This process of identifying extant research foci from the literature and then using these to define code groupings was used to assess the flexibility of the foundational method for each stage of the design process (Hales 1987).

Table 9 summarises the stages and the literature used in assessing the foundational method's application to each. This allowed for each stage to be mapped across the study period and to be evaluated both individually and collectively (Figure 8). Although additional work was required to code the extra information at each level of the schema, it subsequently gave significant scope for reanalysis and reappraisal. More specifically, the foundational method increased the time required for the initial coding by approximately a third, expanding on the original codes rather than demanding a whole new analysis. However, the multifaceted coding reduced reanalysis time – allowing the data to be fed into a range of different research foci including ideation, information seeking, and design review. Building on the multi-level coding allowed these to be assessed automatically with additional codes added only where necessary – reducing the overall workload significantly.

Table 9: Hales' (1987) stages of the design process related to the foundational method

Stage	Description
Conceptual design	Ideation and concept development tasks inc. brainstorming, idea selection and concept exploration (Howard 2008; Cash et al. 2011)
Design development	Development of a specific final concept inc. design refinement and problem solving (Carrizosa and Sheppard 2000; Kim and Maher 2008)
Design review	Reviewing existing work or future planning inc. review meetings and reflection on current designs (Huet, McMahon, et al. 2007; D'Astous et al. 2004)
Embodiment design	Technical layouts and CAD configurations inc. CAD, prototyping and configuration (Scaravetti and Sebastian 2009; Chenouard, Sebastian, and Granvilliers 2007)
Testing	Not considered as not present in the industrial study
Project reporting	Formal collation and dissemination of structured reports inc. lessons learned, reports and formal presentations (Wild et al. 2005; Haas, Weber, and Panwar 2000)
Information seeking	Searching, requesting, synthesizing and evaluating information inc. examination of records and applying data (M. A. Robinson 2010; King, Casto, and Jones 1994)
Dissemination	Informal communication of decisions, plans or progress inc. email, conversations and shared workspace (McAlpine 2010; McAlpine, Hicks, and Culley 2009)

Figure 8: Overall participant activity during the industrial study



With respect to the different research foci three main areas were considered for comparison across the studies.

7.2.2 Information seeking

This example examined the role of information seeking activity and sources on design performance. Here, the situational context was described as 5 (design development stage), 7 (product focused), and 10 (solving) with either 18 (seeking/requesting) or 16 (interpretation) denoting information exchange. The Level 5 codes, based on the work of Robinson (2010), served as the basis of comparison when examining the results.

7.2.3 Ideation

This example examined the need for creative stimuli by assessing the change in the rate of idea generation over time. Here, the situational context was defined as codes 1 (group), 4 (in a meeting room), 5 (feasibility stage), 6 (product focused), and either 9 (exploring the problem) or 10 (solving the problem). Further, the work of Howard et al. (2010) was also characterised using the standard method – facilitating a comparison to this existing dataset from practice. As such, the only Level 5 code was for idea generation. An example of the results is given in Figure 6.

7.2.4 Design Review

This example focused on the use of artefacts during a design review meeting. Here, the codes 1, 4, 5, and 6 were used to define the situational context while Level 2 and 3 codes were used as the basis for the analysis. The results were then compared to the work of Huet et al. (2007).

8. Discussion

This section outlines the theoretical contributions associated with each aspect of the method before offering an overall critique of its limitations and its potential impact on the design research field.

8.1 The Foundational Method

The foundational method proposed in this paper aimed to *improve the replication, reuse, and comparability of empirical design studies*. This was achieved by addressing the specific problems identified in Table 2: linking to theory, describing context, sampling design, research design, data collection, reflexivity, analysis, and value of findings. The foundational method combines the benefits of both standard and flexible elements using multi-level capture, coding, and analysis. This allows the flexible examination of research specific detail whilst also providing rich contextualisation of the situation under study, a standardised dataset, and a means of comparison and triangulation.

The capture step first formalises the reporting of overall context in four areas – activity/technical, social, cultural, and historical. Second, the standard multi-perspective capture approach defines numerous complementary sources. Finally, an acclimatisation period is incorporated into the typical observational approach to reduce experimental effects. These support the generation of a broad and robust dataset, which can be analysed at multiple levels of detail from a wide range of specific research foci.

The multilevel coding strategy allows for a streamlined contextualisation of the wider study and minimal restriction of flexibility by progressive filtering at each of the four coding levels. This enables a rapid interrogation (and comparison) of the dataset at multiple levels of detail whilst maintaining context and methodological robustness, and minimising additional workload.

The corresponding multi-level analysis provides a standard foundation for replication, reuse, and comparison by aligning and baselining the dataset. Further, the analysis strategy allows the researcher to interrogate the data at increasing levels of detail with relatively little additional effort. This enables an analysis of the coded data, which supports both high-level contextualisation and rapid analysis of large bodies of data while also supporting flexibility and overall rigour.

8.2 Improving design research methods

The three steps combined in the foundational method support the standardisation of key comparative data for a wide range of studies. Section 7 considers a comparison between an industrial and an experimental study. This capability is critical to improving reuse and laying the foundation for meaningful comparison and triangulation of data/findings – all key areas for the improvement of design research methods (Blessing and Chakrabarti 2009). Further, the foundational method offers the pragmatic benefit of allowing the researcher to more effectively structure and navigate through the large amounts of data generated in observational studies and significantly expands on the recommendations of Blessing et al. (1998). For example, the narrowing of focus from 32 codes and 240 hours of study, through to 147 unique code combinations, and finally ten situations can be examined and reanalysed at any stage without any further coding. Finally, the multilevel approach allows the foundational method to be flexible in terms of research focus without sacrificing the benefits of standardisation or rigour, as highlighted by the comparison study and discussed throughout.

It is proposed that by taking the first step towards a standardised approach for design research the foundational method contributes directly to the development of methodological and evidential standards in the field. It is unlikely that serious progress will be made in these areas without a bottom up drive for improved methods and standards, motivated by researchers and underpinned by tangible benefits – as has been the case in other fields e.g. education (Gorard and Cook 2007) or policy research (Brennan et al. 2011).

8.3 Reducing research bias

A second area where the method contributes is in the mitigation of researcher bias. Despite there being significant scope for researcher flexibility, the multi-level coding and analysis facilitate the recording of data selection at each level, and the tracing of analytical logic as researchers focus in on the main area of interest. For example, the 147 code combinations of the presented comparative study (Section 7) can each be explicitly described and examined in order to assess coder reliability and facilitate reinterpretation of the data. As such, systematic bias can be more easily traced and accounted for when reusing the data. Further, it provides a fuller record of coding and analysis steps than the typical approach where only the key focus area is considered – lacking the rich contextualisation and foundation provided by the foundational method. Typically this is not reported, making it difficult to assess how situations have been selected or are linked to the wider process. It should be noted that as this method builds on coding based approaches it is not possible to eliminate researcher bias and, as such, promotes a philosophy of transparency.

8.4 Evolution of the foundational method

In terms of longevity, this paper forms one of the first tangible propositions in the discussion of standard methods in design research. This makes for an inherently evolving dynamic environment. Thus, despite the foundational method being designed based on fundamental research issues (Table 2) it is to be expected that as the scientific debate evolves, so will the foundational method. Just as the Randomised Controlled Trial (RCT) (a major standard for clinical research design (Devereaux and Yusuf 2003)) did not emerge in its final form upon conception, the foundational method contributes to an on going debate key to the methodological future of design research. Thus, despite the RCT evolving over time its conception is fundamentally robust, because it was built on a foundation of explicitly addressing core research issues. As such, it is envisaged that even as new methods and good practices evolve in design research, the baselining and aggregation of studies will never be a negative. Further, although research needs and standards may change, fundamental aspects of design will not i.e. the multi-level coding. Thus, the very nature of the foundational method provides a baseline for future comparison, forming a foundation of data and good practice that can be built upon and matured as the field grows, without restricting the development of new and improved methods.

Finally, there are numerous alternative solutions to the issues described in Table 2, hence the complexity of the debate surrounding good research practice. Two key alternatives include improved education of researchers and standardised selection systems for methods or variables. Although the first is clearly a viable alternative it does not supersede the proposed work, instead it offers a complementary means of improvement. Indeed, a key element in education is identifying best practice and training towards effective and meaningful standards, as well as, in the use of recognised methods e.g. the RCT in medical research. As such, the foundational method both guides and complements teaching efforts. Second, both a standardised selection system and the foundational method aim to address the same issues and are therefore complementary whilst offering alternative solutions depending on the evolving research context. However, at time of writing, no such selection system exists in the design research domain, and thus offers a major area for further work. While out of the scope of this paper, meaningful field wide methodological standards and overall improvement are only likely to come from focused discussion and the proposal, comparison, and synthesis of a range of approaches. For example, consider the debate between alternative approaches articulated by Gorard and Cook (2007). Here, despite each individual approach being successful it took many decades and iterations before they were recognised as being complementary and subsequently combined.

In summary it is posited that the foundational method addresses many of the problems identified in Table 2, however, there is still need for further work. This is summarised in
Table 10, which highlights how the problems (Table 2) have been addressed and where the need for further work has been identified.

Table 10: Issues and their mitigation by the foundational method

Problem	Description of mitigation
1. Linking to theory	Contextualisation and multi-level analysis allow situations to be linked to existing work and wider theory by offering a standard basis for comparison
2. Describing context	The overall contextual information and multilevel coding built on situational context significantly improves description of the hypothesis specific elements
Further work	This requires further development in order to identify what specific information is most valuable when recording context in the design domain
3. Sampling design	This is addressed by the standard contextualisation of population and allows for more effective comparison and triangulation of similarly described studies
Further work	There is a need to develop and validate the links between sample design and the elicited contextual information
4. Clarity of research design	Description of the coding schema and the ability to define the level or area of analysis from combinations of codes supports standardisation and clarity
Further work	There is need for significant work in the development of links between levels and the development of relationships between individual and groups of codes
5. Mitigation of bias in data collection	The acclimatisation period and multimodal capture allow for reduced experimental effects and triangulation of multiple sources, reducing bias
6. Reflexivity	The semi-automated nature of the capture strategy eliminates the need for researcher/participant interaction during the study period
Further work	Work is needed to understand the impact of experimental effects over time in the engineering design domain and to subsequently optimise acclimatisation

7. Data analysis	Multilevel coding and analysis coupled with multimodal capture allow characterisation of the system at multiple levels of detail reducing bias
8. Value of findings	The ability to give detailed analysis for selected situations while retaining high-level contextual information supports replication, reuse, triangulation and critique – key areas for improving theory and research uptake

8.5 Constraints and limitations

There are several limitations of the foundational method. The primary weakness is in validating the range of possible participant/case sample sizes to which the foundational method can be applied. However, the multilevel approach allows the researcher to define the sample size required (from statistically significant to single case) and then apply the appropriate level of coding and analysis without losing the advantages of standardisation, contextualisation, and additional detailing.

A second issue requiring further investigation is the period of acclimatization. Although this has been the focus of some investigation in other fields there is little information on the amount of time needed and specific effects encountered in the engineering design domain. An improvement would be to carry out a series of studies to explicitly determine the extent of the disruption caused by experimental setup and the length of time required for participants to return to normal practice. In the context of the comparison study the acclimatisation period was considered sufficient as evidenced by participant's checking private emails and other personal activities. However, for each study the acclimatisation period should be designed accordingly. It should also be emphasised that this ought to be supported with double blind design as best practice.

Third, although the foundational method does introduce some additional methodological and standardisation demands the strategy allows the method to be rapidly adapted to most observation research contexts. This has been demonstrated via the comparative study (Section 7) where information seeking, ideation, and design review were all examined. However, the true scope of the flexibility of the foundational method in terms of sample, compatible research topic, and approach are yet to be fully validated, and form a key opportunity for further study.

Fourth, due to the nature of grounded, qualitative, and mixed-method studies systematic statistical comparison of results is philosophically difficult – particularly grounded qualitative studies. However, a central thesis of this approach is that with better description of the context in which results are generated and a systematic means of baselining populations, more considered aggregation of data could be attempted. This allows for cross study comparisons as opposed to the current situation where this is rarely, if ever, attempted. For example, we highlight the lack of meta-analytic studies in design research. Although comparison (particularly statistical aggregation) is by no means trivial (even using methods such as that proposed) without the elements in the foundational method significant aggregation or comparison attempts are not even possible, severely hampering development in design research. In doing this it is important that the researcher understand the philosophical approaches that define various methods and thus how they are compared. Otherwise incompatible perspectives may be combined leading to erroneous conclusions. In particular the foundational method supports the aggregation of data and does not give direct insight into causal mechanisms. In the causal domain different approaches are needed to establish and explain causal relationships. Thus, although the foundational method may be used to ground causal studies via the generation of baseline data, further development to include casual methods is beyond the scope of this work.

Finally, a more specific limitation of the work reported here is the scope of the comparison study. In the context of the foundational method true validation would require two elements: a systematic comparison across all possible variables and study contexts demonstrating each aspect of the method; a rigorous comparison of the foundational method against all relevant alternative approaches to improving reuse, replication, and comparison. Both of these are significantly beyond the scope of a single study and are likely to only be established reflectively after multiple years of uptake, critique, implementation, and comparison. As such, the comparison study presented in this paper does not claim to validate the method, instead it illustrates the utility of the foundational method and provides an example of how the comparison process can be used to give new insight.

9. Conclusions

This paper outlines the creation of a foundational method for supporting the aggregation of observational studies in the engineering design domain. The method introduces a multi-level approach to capture, coding, and analysis building on pervious works. This comprises multi-stream data capture, combined with a 5 level coding and analysis scheme. Together these promote the wide contextualisation of data and the progressive filtering and focusing of analysis efforts to support standardisation and research specific flexibility.

The foundational method offers several key advantages for improving replication, reuse, and comparison. First, the capture approach formalises the reporting of context and the use of multiple complementary sources in order to produce a broad and robust dataset – allowing for *both* standardised contextualisation and research specific flexibility. Second, the multi-level coding and analysis strategies combine to promote theory building, standardised contextualisation, comparison, triangulation, and reporting. In particular, the coding and analysis strategies allow successive degrees of detail to be examined whilst maintaining a cohesive structure. Third, there is a significant pragmatic benefit in the reduction of the coding and reporting workload whilst maintaining the contextual grounding and flexibility of the research specific elements. Finally, combining standardisation and flexibility allows effective comparison and triangulation of studies in a standard and transparent manner – key to developing a wider base of research data in the community.

As highlighted in
Table 10 further work is necessary to: identify the significance of various contextual factors and formalise their reporting in design research; develop and validate the links between sample design and the required contextual information; and examine the significance and extent of experimental effects in the engineering design domain. Further, and most critical to this work is the on-going requirement to validate the foundational method in practice. However, as true validation can only come through multiple applications in numerous contexts and by many different researchers, this is beyond the scope of any single study. It is hoped that by providing the basis for such comparisons the foundational method will be validated

through examination, critique, and adoption by the engineering design research community itself.

Acknowledgements

The work reported in this paper has been undertaken as part of the EPSRC Innovative Manufacturing Research Centre at the University of Bath (grant reference GR/R67507/0) and has been supported by a number of industrial companies. The authors gratefully acknowledge this and express their thanks for the advice and support of all concerned. We would also like to express our thanks to the reviewers for their insightful input.

References

Adair, J G. 1984. “The Hawthorne Effect: A Reconsideration of the Methodological Artifact.” *Journal of Applied Psychology* 69 (2): 334–345.

Adelman, L. 1991. “Experiments, Quasi-Experiments, and Case Studies: A Review of Empirical Methods for Evaluating Decision Support Systems.” *IEEE Transactions on Systems, Man, and Cybernetics* 21 (2): 293–301.

Adler, N E, and J M Ostrove. 2006. “Socioeconomic Status and Health: What We Know and What We Don’t.” *Annals of the New York Academy of Sciences* 896 (1): 3–15.

Austin, S, J Steele, S Macmillan, P Kirby, and R Spence. 2001. “Mapping the Conceptual Design Activity of Interdisciplinary Teams.” *Design Studies* 22 (3): 211–232.

Ball, L J, and T C Ormerod. 2000. “Putting Ethnography to Work: The Case for a Cognitive Ethnography of Design.” *International Journal of Human-Computer Studies* 53 (1): 147–168.

Balogun, J. 2006. “Managing Change: Steering a Course between Intended Strategies and Unanticipated Outcomes.” *Long Range Planning* 39 (1): 29–49.

Barnes, B R. 2010. “The Hawthorne Effect in Community Trials in Developing Countries.” *International Journal of Social Research Methodology* 13 (4): 357–370.

Bedny, G Z, and S R Harris. 2005. “The Systemic-Structural Theory of Activity: Applications to the Study of Human Work.” *Mind, Culture, and Activity* 12 (2): 128–147.

Bergstrom, M, A Ericson, M Larsson, H Nergard, T Larsson, and B Renstrom. 2008. “Needs as a Basis for Design Rationale.” Design 2008.

Blandford, A, and S Attfield. 2010. “Interacting with Information.” *Synthesis Lectures on Human-Centered Informatics* 3 (1): 1–99.

Blessing, L T M, and A Chakrabarti. 2009. *DRM, a Design Research Methodology*. New York: Springer.
Blessing, L T M, A Chakrabarti, and K M Wallace. 1998. "An Overview of Descriptive Studies in Relation to a General Design Research Methodology." In *Designers - The Key to Successful Product Development*, edited by E Frankenberger, 42–56. Berlin: Springer-Verlag.
Bolton, G E, and A Ockenfels. 2008. *Does Laboratory Trading Mirror Behaviour in Real World Markets?: Fair Bargaining and Competitive Bidding on EBay*. CESifo.
Brennan, L, S Castro, R C Brownson, J Claus, and C T Orleans. 2011. "Accelerating Evidence Reviews and Broadening Evidence Standards to Identify Effective, Promising, and Emerging Policy and Environmental Strategies for Prevention of Childhood Obesity." *Annual Review of Public Health* 32: 199–223.
Briggs, R O. 2006. "On Theory-Driven Design and Deployment of Collaboration Systems." *International Journal of Human-Computer Studies* 64 (7): 573–582.
Carrizosa, K, and S Sheppard. 2000. "The Importance of Learning Styles in Group Design Work." In *Frontiers in Education Conference, 2000. FIE 2000. 30th Annual*, 1:T2B/12–T2B/17 vol.1.
Cash, P. 2012. "Characterising the Relationship between Practice and Laboratory-Based Studies of Designers for Critical Design Situations." Bath: University of Bath, Dept. of Mechanical Engineering.
Cash, P, B J Hicks, and S J Culley. 2010. "An Information Requirement Strategy for Capturing and Analysing Design Activity and Behaviour." In *Design 2010*. Design 2010 International Design Conference. Dubrovnik, Croatia.
———. 2013. "A Comparison of Designer Activity Using Core Design Situations in the Laboratory and Practice." *Design Studies* 34 (5): 575–611.
Cash, P, B J Hicks, S J Culley, and F Salustri. 2011. "Designer Behaviour and Activity: An Industrial Observation Method." In *ICED 11 International Conference on Engineering Design*. ICED 11 International Conference on Engineering Design. Copenhagen, Denmark.
Cave, P R, and C E I Noble. 1986. "Engineering Design Data Management." In *1st International Conference on Engineering Management, Theory and Applications*. Swansea, UK.
Chenouard, R, P Sebastian, and L Granvilliers. 2007. "Solving an Air Conditioning System Problem in an Embodiment Design Context Using Constraint Satisfaction Techniques." *Principles and Practice of Constraint Programming CP 2007*: 18–32.
Cook, D L. 1962. "The Hawthorne Effect in Educational Research." *The Phi Delta Kappan* 44 (3): 116–122.
Cooper, R G, S J Edgett, and E J Kleinschmidt. 2002. "Optimizing the Stage-Gate Process: What Best-Practice Companies Do." *Research-Technology Management* 45 (5): 21–27.
Cross, N. 2007. "Forty Years of Design Research." *Design Studies* 28 (1): 1–4.
Cunningham, S J. 2005. "Autoethnography: A Tool for Practice and Education." In *6th ACM SIGCHI International Conference on Computer-Human Interaction Making CHI Natural - CHINZ 05*, 1. New Zealand.

D'Astous, P, F Detienne, W Visser, and P N Robillard. 2004. "Changing Our View on Design Evaluation Meetings Methodology: A Study of Software Technical Review Meetings." *Design Studies* 25 (6): 625–655.
Devereaux, P J, and S Yusuf. 2003. "The Evolution of the Randomized Controlled Trial and Its Role in Evidence-based Decision Making." *Journal of Internal Medicine* 254 (2): 105–113.
Diaper, G. 1990. "The Hawthorne Effect: A Fresh Examination." *Educational Studies; Dorchester-on-Thames* 16 (3): 261–267.
Dow, D, and S Ferencikova. 2010. "More than Just National Cultural Distance: Testing New Distance Scales on FDI in Slovakia." *International Business Review* 19 (1): 46–58.
Dyba, T, and T Dingsoyr. 2008. "Empirical Studies of Agile Software Development: A Systematic Review." *Information and Software Technology* 50 (9-10): 833–859.
Eisenhardt, K M. 1989. "Building Theories from Case Study Research." *Academy of Management Review* 14 (4): 532–550.
Falk, A, and J Heckman. 2009. "Lab Experiments Are a Major Source of Knowledge in the Social Sciences." *Science* 326 (5952): 535–538.
Finger, S, and J R Dixon. 1989a. "A Review of Research in Mechanical Engineering Design. Part I: Descriptive, Prescriptive, and Computer-Based Models of Design Processes." *Research in Engineering Design* 1 (1): 51–67.
———. 1989b. "A Review of Research in Mechanical Engineering Design. Part II: Representations, Analysis, and Design for the Life Cycle." *Research in Engineering Design* 1 (2): 121–137.
Flay, B R, A Biglan, R F Boruch, F G Castro, D Gottfredson, S Kellam, E K Moscicki, S Schinke, J C Valentine, and P Ji. 2005. "Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination." *Prevention Science* 6 (3): 151–175.
Gallivan, M, and M Srite. 2005. "Information Technology and Culture: Identifying Fragmentary and Holistic Perspectives of Culture." *Information and Organization* 15 (4): 295–338.
Glasgow, R E, and K M Emmons. 2007. "How Can We Increase Translation of Research into Practice? Types of Evidence Needed." *Annual Review of Public Health* 28 (1): 413–433.
Gorard, S, and T D Cook. 2007. "Where Does Good Evidence Come From?" *International Journal of Research & Method in Education* 30 (3): 307–323.
Guzzo, R A, and M W Dickson. 1996. "Teams in Organizations: Recent Research on Performance and Effectiveness." *Annual Review of Psychology* 47 (1): 307–338.
Haas, R E, F Weber, and K S Panwar. 2000. "Engineering Knowledge Management - Current Status and Future Challenges." ICE 2000 International Conference on Concurrent Enterprising.
Hagedorn, J, J Hailpern, and K G Karahalios. 2008. "VCode and VData: Illustrating a New Framework for Supporting the Video Annotation Workflow." In *Conference on Advanced Visual Interfaces*, 317–321. New York, USA: ACM.
Hales, C. 1987. "Analysis of the Engineering Design Process in an Industrial Context". Cambridge UK: University of Cambridge, Department of Engineering.
Hicks, B J, S J Culley, R D Allen, and G Mullineux. 2002. "A Framework for the Requirements of Capturing, Storing and Reusing Information and Knowledge in

Engineering Design." *International Journal of Information Management* 22 (4): 263–280.
Hofstede, G, G J Hofstede, and M Minkov. 2010. *Cultures and Organizations: Software of the Mind, Revised and Expanded 3rd Ed*. Vol. 3rd. McGraw-Hill, New York, NY.
Holden, J D. 2001. "Hawthorne Effects and Research into Professional Practice." *Journal of Evaluation in Clinical Practice* 7 (1): 65–70.
Horvath, I. 2004. "A Treatise on Order in Engineering Design Research." *Research in Engineering Design* 15 (3): 155–181.
Howard, T J. 2008. "Information Management for Creative Stimuli in Engineering Design." Edited by E Dekoninck. Bath: University of Bath, Departement of Mechanical Engineering.
Howard, T J, S J Culley, and E Dekoninck. 2010. "Reuse of Ideas and Concepts for Creative Stimuli in Engineering Design." *Journal of Engineering Design* 22 (8): 565–581.
Huet, G, S J Culley, C McMahon, and C Fortin. 2007. "Making Sense of Engineering Design Review Activities." *Artificial Intelligence for Engineering Design, Analysis and Manufacturing : AI EDAM* 21 (3): 243–266.
Huet, G, C McMahon, F Sellini, S J Culley, and C Fortin. 2007. "Knowladge Loss in Design Reviews." In *Advances in Integrated Design and Manufacturing in Mechanical Engineering II*, edited by S Tichkiewitch, M Tollenaere, and P Ray, 277–291. Springer.
Jakesch, M, M Zachhuber, H Leder, M Spingler, and C C Carbon. 2011. "Scenario-Based Touching: On the Influence of Top-down Processes on Tactile and Visual Appreciation." *Research in Engineering Design* 22 (3): 1–10.
Janssen, O, E Van de Vliert, and M West. 2004. "The Bright and Dark Sides of Individual and Group Innovation: A Special Issue Introduction." *Journal of Organizational Behavior* 25 (2): 129–145.
Kazdin, A E. 1998. *Research Design in Clinical Psychology*. Needham Heights, MA, USA: Allyn & Bacon.
Kim, M J, and M L Maher. 2008. "The Impact of Tangible User Interfaces on Spatial Cognition during Collaborative Design." *Design Studies* 29 (3): 222–253.
King, D W, J Casto, and H Jones. 1994. *Communication by Engineers: A Literature Review of Engineers' Information Needs, Seeking Processes, and Use*. Washington, DC: Council on Library Resources.
Kitchenham, B A. 1996. "Evaluating Software Engineering Methods and Tool Part 1: The Evaluation Context and Evaluation Methods." *ACM SIGSOFT Software Engineering Notes* 21 (1): 11–14.
Kitchenham, B A, S L Pfleeger, L M Pickard, P W Jones, D C Hoaglin, K El-Emam, and J Rosenberg. 2002. "Preliminary Guidelines for Empirical Research in Software Engineering." *IEEE Transactions on Software Engineering* 28 (8): 721–734.
Klein, H K, and M D Myers. 1999. "A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems." *MIS Quarterly*: 67–93.
Kogut, B, and H Singh. 1988. "The Effect of National Culture on the Choice of Entry Mode." *Journal of International Business Studies* 19 (3): 411–432.
Leonard, K, and M C Masatu. 2006. "Outpatient Process Quality Evaluation and the Hawthorne Effect." *Social Science & Medicine* 63 (9): 2330–2340.

Lethbridge, T C, S E Sim, and J Singer. 2005. “Studying Software Engineers: Data Collection Techniques for Software Field Studies.” *Empirical Software Engineering* 10 (3): 311–341.
Levitt, S D, and J A List. 2007. “What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?” *The Journal of Economic Perspectives* 21 (2): 153–174.
Lewis, M, and J Moultrie. 2005. “The Organisational Innovation Laboratory.” *Creativity and Innovation Management* 14 (1): 73–83.
LiveScribe. 2011. “LiveScribe: Never Miss a Word”. http://www.livescribe.com/en-us/.
Malterud, K. 2001. “Qualitative Research: Standards, Challenges, and Guidelines.” *The Lancet* 358 (9280): 483–488.
ManicTime. 2011. “ManicTime - Time Management Software [Online]”. http://www.manictime.com/.
McAlpine, H. 2010. “Improving the Management of Informal Engineering Information through Electronic Logbooks.” Edited by S J Culley. Bath: University of Bath, Dept. Mechanical engineering.
McAlpine, H, P Cash, A Storton, and S J Culley. 2011. “A Technology Selection Process for the Optimal Capture of Design Information.” In *ICoRD ’11 International Conference on Research into Design*. ICORD 11 International Conference on Research into Design. Bangalore, India.
McAlpine, H, B J Hicks, and S J Culley. 2009. “Comparing Formal and Informal Documents: Lessons for More Complete Design Records.” In . ICED 09 International Conference on Engineering Design. Stanford University, USA.
McAlpine, H, B J Hicks, G Huet, and S J Culley. 2006. “An Investigation into the Use and Content of the Engineer’s Logbook.” *Design Studies* 27 (4): 481–504.
McCandliss, B D, M Kalchman, and P Bryant. 2003. “Design Experiments and Laboratory Approaches to Learning: Steps toward Collaborative Exchange.” *Educational Researcher* 32 (1): 14–16.
Panopto. 2012. “Panopto”. www.panopto.com: Panopto Europe Ltd.
Petre, M. 2004. “How Expert Engineering Teams Use Disciplines of Innovation.” *Design Studies* 25 (5): 477–493.
Pickett, K E, and M Pearl. 2001. “Multilevel Analyses of Neighbourhood Socioeconomic Context and Health Outcomes: A Critical Review.” *Journal of Epidemiology and Community Health* 55 (2): 111–122.
Podsakoff, P M, S B MacKenzie, J-Y Lee, and N P Podsakoff. 2003. “Common Method Biases in Behavioural Research: A Critical Review of the Literature and Recommended Remedies.” *Journal of Applied Psychology* 88 (5): 879–903.
Pope, C, and N Mays. 1995. “Reaching the Parts Other Methods Cannot Reach: An Introduction to Qualitative Methods in Health and Health Services Research.” *BMJ British Medical Journal* 311 (6996): 42–45.
Prudhomme, G, F Pourroy, and K Lund. 2007. “An Empirical Study of Engineering Knowledge Dynamics in a Design Situation.” *Journal of Design Research* 6 (3): 333–358.
Puttre, M. 1991. “Product Data Management.” *Mechanical Engineering* 113 (10): 81–83.

Robinson, H, J Segal, and H Sharp. 2007. "Ethnographically-Informed Empirical Studies of Software Practice." *Information and Software Technology* 49 (6): 540–551.
Robinson, M A. 2010. "An Empirical Analysis of Engineers' Information Behaviours." *Journal of the American Society for Information Science and Technology* 61 (4): 640–658.
Robson, C. 2002. *Real World Research*. Vol. 2nd. Chichester: Wiley.
Scaravetti, D, and P Sebastian. 2009. "Design Space Exploration in Embodiment Design: An Application to the Design of Aircraft Air Conditioners." *International Journal of Product Development* 9 (1): 292–307.
Seale, C. 1999. "Quality in Qualitative Research." *Qualitative Inquiry* 5 (4): 465–478.
Seale, C, and D Silverman. 1997. "Ensuring Rigour in Qualitative Research." *European Journal of Public Health* 7 (4): 379–384.
Shalley, C E, and L L Gilson. 2004. "What Leaders Need to Know: A Review of Social and Contextual Factors That Can Foster or Hinder Creativity." *The Leadership Quarterly* 15 (1): 33–53.
Shavelson, R J, D C Phillips, L Towne, and M J Feuer. 2003. "On the Science of Education Design Studies." *Educational Researcher* 32 (1): 25–28.
Shenkar, O. 2001. "Cultural Distance Revisited: Towards a More Rigorous Conceptualization and Measurement of Cultural Differences." *Journal of International Business Studies* 32 (3): 519–535.
Stewart, G L. 2006. "A Meta-Analytic Review of Relationships between Team Design Features and Team Performance." *Journal of Management* 32 (1): 29–55.
Torlind, P, N Sonalkar, M Bergstrom, E Blanco, B J Hicks, and H McAlpine. 2009. "Lessons Learned and Future Challenges for Design Observatory Research." In . ICED 09 International Conference on Engineering Design. Stanford, CA, USA.
Torlind, P, M Stenius, M Johanson, and P Jeppsson. 1999. "Collaboration Environments for Distributed Engineering Development of a Prototype System." *Computer Supported Cooperative Work in Design (CSCWD)* 99.
Veldman, J, and A Alblas. 2012. "Managing Design Variety, Process Variety and Engineering Change: A Case Study of Two Capital Good Firms." *Research in Engineering Design* 23 (4): 269–290.
Visser, W. 2009. "Design: One, but in Different Forms." *Design Studies* 30 (3): 187–223.
Wasiak, J, B J Hicks, L Newnes, A Dong, and L Burrow. 2010. "Understanding Engineering Email: The Development of a Taxonomy for Identifying and Classifying Engineering Work." *Research in Engineering Design* 21 (1): 43–64.
Wild, P J, S J Culley, C McMahon, M Darlington, and S Liu. 2005. "Starting to Audit Documents in the Engineering Domain." In *In-Use, In-Situ: Extending Field Research Methods*, edited by P Amaldi, S P Gill, B Fields, and W Wong, 36–40. London: Higher education academy information and computer sciences.
Wild, P J, C McMahon, M Darlington, S Liu, and S J Culley. 2010. "A Diary Study of Information Needs and Document Usage in the Engineering Domain." *Design Studies* 31 (1): 46–73.
Wildemuth, B M. 1993. "Post-Positivist Research: Two Examples of Methodological Pluralism." *The Library Quarterly* 63 (4): 450–468.

Appendix: Summary of the Overall Contextual Factors

Company		
Focus	**Measures**	**Metric or description**
Technical Environment	Technical layout	The layout and distribution of equipment
	Resources available	The equipment available to the team
	Overall layout	Office and site plan
	Number of co-workers	Size of working team
	Distribution of co-workers	Geographic distribution of team
	Types of co-workers	Background and role of team members
Social	Funding/income sources	Source of income for project and company
	Market pressures	Competitors and main market forces
	Environmental drivers	Environmental and legislative forces
	Other pressures	Any other forces affecting the company
	The overall number and breakdown of employees	Overall N^o, role composition and geographic distribution of workforce
	Hypothesis specific employees	N^o, role composition and geographic distribution of people under investigation
Cultural	Main aim(s) and scope	Company and team aim
	Values/mission statement(s)	Company and team priorities and future vision
	Expertise	Composition and distribution of workforce roles and expertise
	Company culture	Using Hofsetede et al's (1990) measures
	Focus and level/type of engineering/design	The complexity and domain of the design work being undertaken
	Past projects	N^o of related past projects
	Significant partners	Sister, parent or subsidiaries
	Significant partners role in management	Level at which there is contact with and influence by partners
Historical	The turnover of the company	The current size and historical growth
	The age of the company	The number of years in the market segment

Participant		
Focus	**Measures**	**Metric or description**
Technical Environment	Specific technical features of the participants work station	Operating system and specialist software/tools available
	Distribution of time across work areas	Nominal time working in office, workshop, home etc.
	Use of resources	Resources available to the participant e.g. whiteboard, notepad, phone, bookshelves etc.
Social	Socioeconomic status	Age, role, highest level of education, gross individual annual income, level of property ownership
	Area-based measure of socioeconomic status	Using e.g. ACORN www.caci.co.uk/acorn-classification.aspx
Cultural	Nationality and national heritage	Current nationality, previous nationalities or time in other countries
	Cultural distance measures	Using Hofsetede et al's (2010) measures
Historical	Formal education: subjects and grades and focus	Subjects and grades, focus of any larger projects e.g. Masters thesis

	Professional qualifications	Subjects and levels (where relevant)
	Professional experience over six months	Role, duration, description of company
	Development within the current professional framework	Participation in development schemes in the company e.g. leader development or technical training

Appendix: Code Definitions

Group	Code	Definition
Situation	Individual	No real time interaction with any other individual or group
	Group	Real time interaction with one or more other individuals
	Synchronous	No delays between communications
	Asynchronous	Significant delays (longer than a few seconds) between communications
	Co-located	Working in the same location at the time of an interaction
	Distributed	Working in different locations at the time of an interaction
Environment	Location	The specific location of the participant in their main work site
Focus	Design process stage	The stage at which an interaction is taking place within the associated project – see Hales (1987) for stage definitions
	People	The subject of an interaction includes: personnel, personal, managing people, customers
	Product	The subject of an interaction includes: prototypes, design documents, project management
	Process	The subject of an interaction includes: resources/time allocation, scheduling, stage gate management
Problem solving	Goal setting	Identifying where the design is and where it needs progressing to
	Constraining	Imposing boundaries with requirements and desirables
	Exploring	Discussing possibilities and ideas invoking suggestions
	Solving	Involves searching, gathering, creating, developing solutions
	Evaluating	Judging the quality, value and importance of something
	Decision making	Considering key factors from evaluation and possible compromises to form decisions
	Reflection	Reflecting upon a design decision or process already adopted or occurred
	Debating	Discussing opposing views
Information exchange	Recognising need	Recognising a problem or deficit
	Seeking	Finding information
	Requesting	Direct requests to another party to provide information
	Interpretation	Assigning meaning or value to information
	Validation	Checking the authenticity or value of information
	Informing	Using information to inform one or more people
	Clarifying	Using information specifically to resolve issues or clarity problems
	Confirming	Using information specifically to affirm or confirm a issue or point
Management exchange	Managing	Specifically arranging, directing or instructing with regards to people, product or process
Audiovisual	Audio only	Only using audio input or output
	Visual only	Only using visual inputs or outputs
	Audiovisual	Using both audio and visual inputs or outputs
Documentation	Formal	Provides a specific context and measure with a structure or a focus such that individuals exposed to it may infer the same knowledge from it (Hicks et al. 2002)

	Informal	This encompasses any unstructured information (Hicks et al. 2002)
Physical	Environment	Physical objects not directly related to the design
	Tools	Design tools used with respect to the design (Schon 1984)
	Design representations	Objects related to the specific design under discussion – prototypes, visualisations, mock-ups etc
Type of exchange	Opinion	Giving or receiving opinions: includes evaluation, analysis, expression of feeling or wish
	Orientation	Giving or receiving orientation or scene setting: includes information, repetition, confirmation
	Suggestion	Giving or receiving direction or proposed possibilities: includes direction, possible modes of action
Understanding	Agree/disagree	The participant shows passive acceptance/rejection, understands, concurs, complies/formality, withholds resources
Personal	Antagonism/ solidarity	Giving or receiving support/criticism: increases/decreases others status, gives help or rewards others/asserts or defends self
	Tension/ tension release	The participants jokes, laughs, shows satisfaction/asks for help, withdraws