INTERNATIONAL DESIGN CONFERENCE - DESIGN 2012
Dubrovnik - Croatia, May 21 - 24, 2012.

IMPROVING CONFIDENCE IN SMALLER DATA SETS THROUGH METHODOLOGY: THE DEVELOPMENT OF A CODING SCHEME

C. M. Snider, E. A. Dekoninck and S. J. Culley

Keywords: engineering design, content analysis, methodology

1. Introduction

Particularly when utilising smaller data sets, as may occur in preliminary and scoping research, it is vital to ensure confidence in the results. Although results and conclusions in such work may only be used to uncover direction for further, more rigorous research, careful development through consideration of methodology will provide a clearer understanding of the gathered data, and thus help to identify these directions with higher confidence.

This paper presents some of the most important methodological considerations that must be made in the development of a coding scheme used for content analysis; a process which occurs through the collection and analysis of verbal, written or visual data. Following, the researcher is able to describe phenomena or test issues that build upon or add to theoretical understanding [Elo et al. 2008]. Analysis itself occurs through the grouping of content within the data into categories that can then be quantified and studied according to a set of rules laid out in a coding scheme. Content analysis has had wide use across different fields and disciplines, including psychology and the social sciences.

By presenting these important factors, the paper aims to demonstrate the manner in which a robust, reliable and valid coding scheme can be developed. The way in which these considerations can be made has been demonstrated through the example of the development of a scheme designed to provide understanding of creative approaches used by designers within the later stages of the engineering design process, preliminary results of which have been presented in prior work [Snider et al. 2011]. Through the presentation of these factors, the paper aims to demonstrate the manner in which good practice can be ensured in research within engineering design; thus creating confidence in results and the directions for future work that they provide, a consideration that is particularly important when attempting to draw conclusions from smaller data sets.

2. Background of the study

The role of the coding scheme presented within this case study was to identify the creative behaviour of designers as they progress through the design process, with particular focus on the creative approaches that they may utilise and the activities in which they may utilise them (referred to in this work as *tasks*). The overall study aimed to analyse the creative behaviour of engineering designers, with focus on that within the later stages. Of particular interest were the tasks in which designers were typically creative and the overall creative approach that they took, as presented in Snider et al. [2011]. Within this paper, only the basic framework of the scheme itself and the methodology by which it was developed are presented, to display the manner by which greater confidence was ensured in the results.

2.1 The coding scheme

The coding scheme is as presented in Snider et al. [2011], and is designed to identify creative design approaches followed by designers within the later stages of the design process. This is achieved through study of the designers task activity throughout, specifically on the use of each task either to develop the information that they have about the design, the brief or the domain (called *information* tasks); or on the use of each task to develop the design itself, e.g. the physical components, layouts and materials for example (called *application* tasks).

Each task is then defined according to its initial and final state based on *information* and *application*, creating four possible options. A designer may use the information that they already have, and develop it into a broader or more developed version (I → I); they may apply the *information* that they have to create the design itself (I → A); they may take the current form of the design (*application*) and re-work it into a more developed version of itself (A → A); or they may take the current form of the design (*application*) and analyse it to develop *information* (A → I).

In addition to these task categories, the coding scheme considers creative behaviour within each task according to whether it is completed in an *expansive* or *restrained* manner. Relating to the work of [Guilford 1956], to be *expansive* refers to creativity in both divergence and convergence within the design process, through the pursuit of alternative products and technologies, or through the development and integration of new part combinations. To be *restrained* refers to a lack of creativity in either divergent or convergent processes. The coding scheme then identifies tasks as belonging to one of eight groups, as shown in Figure 1 according to their input and output states.

	Information (I)		**Application (A)**	
	Single	Translational	Translational	Single
Restraine	I → I	A → I	I → A	A → A
Expansive	I → I	A → I	I → A	A → A

Figure 1. The eight task types

2.2 Experimental methodology

This coding scheme was applied to the work of seven final-year engineering design students over the entirety of a 22 week individual project, encompassing a total of 711 distinct tasks. The project consisted of a significant part of the design process, from initial task clarification to the construction of a working, demonstrable prototype. These projects varied between designers, but all followed the same overall requirements and structure, as shown in Table 1.

Table 1. Project structure

Weeks 1-11	Weeks 12-22
Stage 1 Develop problem understanding	Stage 4 Develop final concept
Stage 2 Perform background research	Stage 5 Manufacture proof of principle working prototype
Stage 3 Report research and in-depth specification	Stage 6 Full report
Assessment	Assessment

As such all projects can be compared in terms of generic process, but the differences between the project briefs cannot be accounted for. For the purposes of this work this deficiency must be acknowledged and accounted for in any conclusions made; but can only truly be addressed through further experimentation in which the project brief is constant.

Coding was based on the student's logbooks, which were kept and regularly used as part of the assessment process. For this research each logbook was coded in three separate passes; the first to allow separation of individual tasks, the second to assign the type of task according to the coding scheme, and the third to identify the appropriate design stage. Coding in these separate passes allowed higher focus on each individual element of the coding scheme. Through the rules within the coding scheme the coder is able first to identify the boundaries between distinct tasks, and second individual entities within the data. These entities are based on the recordings of the designer within the logbook; as example *knowledge (K)*, when the designer records a discrete piece of information regarding the design or domain in which it exists; and *behaviour (B)*, when the designer records the working relationships between systems or components. The coder then uses their judgement to determine the task the designer was performing through determination of which entities formed the inputs and outputs from those that appeared. Finally, the coder determines the design stage at which the task occurred. The rules and definitions within the coding scheme were developed from underlying theory and development work, and are presented in detail in Snider et al. [2011].

3. The role of theory in building a coding scheme

It is important that the results and analysis produced by any coding scheme are appropriate to the purpose of the study in which it is used, to ensure the development of informative results and understanding. Hence, while coding schemes have been developed by many researchers in the past, it is likely that any new studies will require the modification of a pre-existing scheme in order to bring them in line with the specific aims of the new study, or the development of a new scheme in its entirety. For a coding scheme to be of quality it must be built on strong foundations; considerate of the fundamental theory from which it is developed and to which it will contribute, as well as of the type of data that will be analysed.

The use of theory within content analysis varies, as described by Potter and Levine-Donnerstein [1999]. A *deductive* role describes that of a scheme which is built from theory directly, in which the definition of concepts within the theory determine the areas of focus and relationships that the coding scheme identifies. The purpose of such a role is that of theory validation or extension, using the scheme to contribute to its development [Hsieh et al. 2005]. The *inductive* role then occurs when theory is initially limited, and involves the researcher forming categories according to patterns that they identify in the data; which can then be used to inform the development of new theory, typically in quite a general sense. Inductive scheme development itself occurs through examination of the data without preconception of categories into which it may fall [Kondracki et al. 2002]; through which the researcher is able to identify appropriate categories and themes into which the data can be grouped for future analysis.

This case study primarily took a deductive approach in the use of theory, but did allow an inductive approach to play a role. Initially, a study of relevant theory led to the identification of categories into which the data could fall that would potentially highlight different forms of designer behaviour; as well as highlight behaviour that promoted creative results (as opposed to the more routine). However, while providing grounding, the abstract nature of the theory led to definition of categories in this case that were too vague, creating confusion in their interpretation, implication and reliable use. This was remedied through an inductive process on one complete designer logbook, dividing data according to categories that occurred within. Through comparison of these categories with those developed from the deductive approach it was then possible to create sharper definitions, tightly grouping data into appropriate and usable categories.

4. Reliability

Reliability within a coding scheme refers to the consistency of the results produced from the data, when analysed by alternative researchers. Should a scheme be reliable, the same results will appear from the same data set regardless of who is performing the coding, and regardless of whether the data has been coded before.

Reliability is tested through the process of inter-coder testing, in which multiple researchers code the same data set, using the same rules. The results of this coding are then compared, and computed into a level of agreement that reflects the ability of the scheme rules to produce consistent results.

4.1 Coefficients of reliability

Agreement itself is measured using a variety of metrics; those typically used include percentage agreement, the number of times the researchers agreed as a percentage of the whole; Cohen's kappa [Cohen 1960], which includes the possibility of the results occurring by chance; and Krippendorff's alpha [Hayes et al. 2007], which includes the same, but is accommodating of a wider number of data types and number of coders. Of these, Cohen's kappa and Krippendorff's alpha are considered to be the better measures due to their consideration of chance; using only percentage agreement tends to produce a higher value, overestimating the ability of the scheme to produce reliable results.
Acceptable correlation values for these measures are typically taken to be above 80% [Neuendorf 2002], although a value of 70% can be taken for exploratory research [Blessing et al. 2009], [Klenke 2008] such as this that completed in this study.

4.2 Selection of coders

Care must also be taken in the choice of coders used to test reliability. A high level of prior knowledge of the scheme may skew results through coding according to what the scheme *should* produce according to its purpose, but not necessarily according to what the scheme *would* produce if used by a coder without prior knowledge. It is therefore best to use independent coders that were unrelated to scheme development, a practice that could only be partially completed in this case through the use of one independent coder, and one coder with prior knowledge.
The number of coders is also an important consideration. As the purpose of inter-coder reliability testing is to compare results between researchers it is vital that more than one is used (although occasionally coding is completed by only one, with a significant time gap between sittings). A higher number of coders will then make the scheme more robust, and should be used whenever possible.

4.3 The process of inter-coder training

As part of the process of ensuring reliability, each coder must be appropriately trained [Krippendorff 1981]. Training takes place through an iterative process of coding, assessment and re-definition designed to increase reliability within the scheme and the coders.

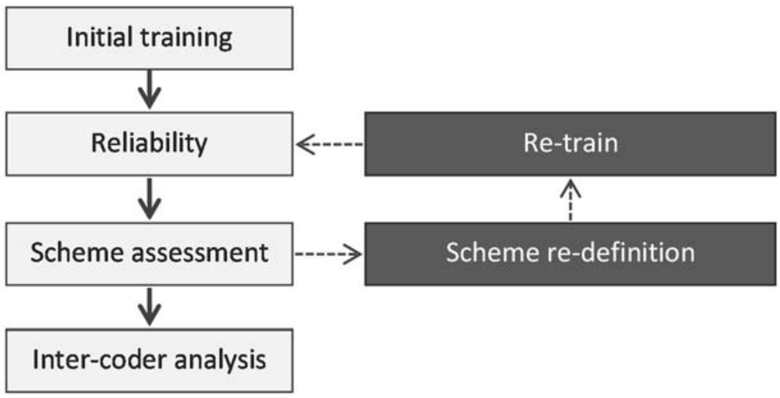


Figure 2. The training model used within this analysis, based on [Krippendorff 1981]

The training process within this case study occurred according to the process within Figure 2, and is largely based upon that recommended by Krippendorff [1981]. The initial training and reliability stages serve to introduce the trainee to the scheme in a more general sense, and to allow analysis of their interpretations of the scheme. This is then used in an assessment process, considering the accuracy of coding the trainee can provide and the various reasons by which their coding may differ from the expected, and from what is considered *correct*.

Training occurred according to the following steps:

- Familiarisation of the coding scheme with the trainee, aimed at producing an understanding of the codes and rules themselves, and how they must be applied to the data set for analysis.
- Preliminary coding of approximately 50 tasks from several logbooks; analysed for both overall reliability and any reasons for disagreement.
- Re-definition of coding terms to reduce ambiguity and the possibility of misinterpretation of data; followed by re-training of the coder.
- Preliminary coding of one entire logbook containing 158 tasks; again analysed for both overall reliability and any reasons for disagreement, which were analysed on a case by case basis.
- Further re-definition of terms to further reduce ambiguity; followed by retraining of the coder.

This process allowed the coder to gain sufficient relevant experience to provide a consistent and reliable result, while also highlighting and improving the coding scheme.

An important danger that must be considered when refining a scheme in this way is that of the shared understanding of the coder and the trainees, stemming from the co-interpretation and analysis of disagreements. This shared experience allows a development of understanding amongst coders, such that the scheme can only be expected to be reliable and consistent among those who were involved in its development. Clearly, this effect must be minimised. Within this research, coder training occurred on a single researcher who was not involved in the development of the scheme, and had little experience in the subject area. This enabled the above concern to be controlled to an extent; their lack of domain knowledge allowed them to remain a largely independent coder, who did not significantly influence the re-definition process beyond identification of ambiguity or areas needing attention.

Due to experimental constraints the process of training for this study occurred within five days, significantly less than the (unrealistic in this case) several month long process recommended by literature [Krippendorff 1981]. As a result, the re-definition and re-training phases of the process were completed to a less than ideal level, resulting in some ambiguity of the categories into which data should fall. Regardless, the achieved values for Cohen's kappa and Krippendorf's alpha of 0.770 and 0.768 respectively are considered suitable for exploratory research, and reflect the finding of Milne and Adler [1999] that adequate agreement can be reached with minimal training.

Table 2. Evolution of rule and code definition through the training process

Session	First	Second	Third
Cohen's kappa	0.497	0.678	0.768
Subsequent Modification	Development of stricter rules for the separation of tasks within the data. Determination of tasks themselves remains a coder judgement.	Introduction of definition of "entities" within the data that coders can use to better judge task process. Relationship between entities defines task and is reliant on coder judgement.	N/A – final inter-coder testing session

4.4 The process of inter-coder analysis

When selecting a data set to use within inter-coder analysis it is important to choose a sample that can be considered representative of the data as a whole. The tested data set should then be large (in excess of 10% of the total), and should also be randomly selected from the whole [Potter et al. 1999]. This is particularly important when coding logbooks throughout the design process, in which different stages within the data may reflect different tasks and different behaviours that must be considered.

In this case, testing occurred using a single trainee and the original researcher who developed the scheme (a typical but non-ideal process resulting in one coder with extensive prior knowledge of the scheme). The chosen data set itself was split between two logbooks; the first of which had not previously been seen by either coder (to minimise the effect of knowledge that the original researcher holds), and the second of which was chosen due to its style of recording, which had proven particularly difficult to understand. In each, logbook data was randomly selected in several contiguous sections of approximately ten tasks, taken from varying stages in the design process; with the overall total set used for reliability testing equalling 10% of the total within the analysis. Through this data set, the effect of memory and understanding of the original coder was minimised, and the scheme was tested on difficult data that was most likely to produce poor agreement.

5. Validity

When creating rules by which data is coded, it is important to provide sufficient detail to ensure that all information is coded *correctly*, according to the underlying theory and development work on which the scheme is built. There are several facets to this assurance; the rules of the scheme must produce results true to the data itself, representative of the information it contains; the data set must be representative of that which is being studied, hence allowing generalisation of the results beyond the study itself; and the results must be true to the underlying theory rather than only producing results according to the scheme. These facets can be accounted for through consideration of the form of data to be coded, justification of the data set itself, and comparison of the results to an alternative independent measure, accepted as representative of the underlying theory.

5.1 Form of data

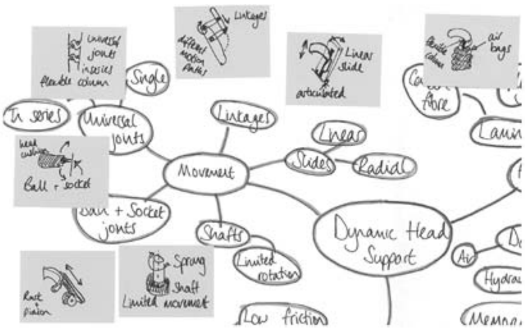
To ensure that the coding scheme identifies according to the purpose for which it was designed, it is important to determine the *type* of content that is being coded, and how any researchers may interpret it. While the data itself may remain constant, the way in which it is coded may require the researcher to pass their own judgement of how it may relate to the scheme or what it may mean in relation to the underlying theory. In cases such as this it is therefore important to create rules that provide direction to the judgements that the researcher may make, but it is equally important (in many cases) that the rules are not so rigid that judgement does not occur. In order to develop a scheme it is then important to understand the type of content that must be coded for analysis to occur (Figure 3), the requisite judgement that coders must make to identify it, and hence the extent to which the rules may steer the coding and influence the coding schemes validity.

Type	Manifest	Latent Pattern	Latent Projective
Description	Interpreted according to clear rules, directly from the content of that which is being coded	Requires judgement of the coder. Codes that which is present on the page, but requires the coder to interpret its meaning and relationship to the final codes.	Requires judgement of the coder. Codes the judgements that the coders have made of the information that is presented to them.

Figure 3. The three forms of coding [Potter et al. 1999]

The type of data coded within the coding scheme presented in this study is known as *latent pattern* [Potter et al. 1999]; in that precedence is put on the content of the data, and the belief that patterns can be found through recognising connections between the various entities the coding scheme identifies. This process requires interpretation from the coder – it is their judgement that determines whether entities are important in each task and which entity leads to another. This is in contrast to manifest content, in which the entities would be coded alone; and latent projective content, which requires a far higher level interpretation from the coder to complete. To provide an example similar to Potter and

Levine Donnerstein [1999]; when coding people, manifest content would be that which is clear, such as the colour shirt they are wearing. Latent pattern content would then require some interpretation from the researcher but would still be based on that which they can see, such as level of fitness; and latent projective content would rely heavily on the judgement of the coder, such as friendliness.
The effect of the type of data coded is that of its strong impact on the intertwining of reliability and validity. When coding manifest content, there is usually a specific correct answer for each entry. Accordingly, reliability can be ensured through the use of extensive rules, carefully constructed so that each entity can fall into only one category. Consistent results can then be produced; extensive rules will increase both reliability and validity as long as they are built accurately from the theory.
The relationship between latent content and validity is more complex. The development of extensive rules will still increase reliability between coders, but as a result the validity will decrease. This is because, with extensive rules, focus is shifted away from the difficult judgements that the coder should make onto the easier and clearer aspects; thus providing potential for the rules to cause the coding to shift away from the essence of that being coded [Potter et al. 1999]. Validity is thus lower, with the scheme classifying according to specific and explicit features of the data rather than the important coder judgments that allow understanding of the underlying theory. A trade-off exists between reliability and validity in the development of the coding scheme; enough rules must be present to produce consistent results, without so many that coder judgements are no longer needed.
Within this research, the rules developed provide ways of classifying entities within tasks, but do not control how the transformation between them is interpreted. This is left to the coder, with examples of typical transformations from other data as a guide. Validity is then not sacrificed for the sake of reliability; assuming the scheme is developed correctly from theory all coders will make similar judgements, showing scheme reliability and that it classifies in the manner for which it was designed.



A task containing a *Function* entity (dynamic head support) and several *Behaviour* entities (working principles of mechanisms). Judged by the coder to be an example of determining working principles that could complete a function, and hence taking the coding:

Entity transformation - F → B; Verb descriptor - "*Apply*"

Figure 4. Example of coding for a single task

5.2 The use of logbooks

The source of data used throughout this study was that of undergraduate engineers' logbooks, used for the entirety of a single 22 week design project. Validity then applies to the justification of the use of logbooks over other potential sources of data; demonstrating the level to which a logbook can be considered an appropriate record of the engineering design process.
Justification requires knowledge of the types of data that can be coded using content analysis. Amongst the common and effective forms of data collection is protocol analysis, in which the designer describes their thoughts and actions as they are completed. This can happen either concurrently or retrospectively; in the former the designer states what they are doing as they do it (thinking aloud protocol); in the latter the designer describes what they did after the fact. However, within this study it was both impractical for the designer to be recorded for the duration of the 22 weeks, or to regularly gather retrospective protocols directly following the period of design, a necessity to ensure accuracy and validity overall [Gero et al. 2001]. As such, protocol analysis could not be used.
In order to accurately study the design process followed in its entirety it was then necessary to utilise an alternative source of data, taken in this case from designers' logbooks; maintained exclusively for the purposes of one project. The assumption that designers logbooks are a complete, representative record of their design process is very important when considering validity; conclusions should be

drawn only from data that is representative of its purpose, otherwise the results can be as interpreted as valid in terms of the data set, but not in terms of the wider context in which content analysis occurs.
Engineers logbooks are a good record of the process followed, in terms of the chronology of recordings within [McAlpine et al. 2006], and due to the reliance of undergraduates on hand-drawn representations [Sobek 2002]. However, while logbooks capture a large amount of the expansive idea generation process [Currano et al. 2009], they will not necessarily capture all tasks that occur. For example, while initial dimensioning tasks may be drawn, the logbook will not capture any evolution of these dimensions that occurred during any computer-aided design process. In attempt to counter potential incompleteness, the 7 studied logbooks were collected from a possible 17 and were chosen specifically for their apparent completeness and legibility in recording style, judged through quantity of data and evidence of recording throughout the project. In a similar manner this process also attempted to discriminate against logbooks that may have been written after the task occurred.
Although considered and minimised within the selection process of logbooks, a further consideration is that of the ambiguity of the data recorded in logbooks itself. The designer is unconstrained in their writing and recording process, and as such the data within the logbooks was often difficult to interpret. Careful identification of features within the data through the rules and codes used greatly ease this difficulty, but due to the nature of the use of logbooks themselves and the work within it is likely that both reliability and validity suffer to some degree. While not possible in this case, identifying markings within the logbook in tandem with a time dimension or other more direct observation would help to rectify the problem, increasing the scores for inter-coder reliability and validity of the results.
Accordingly, it is apparent that an engineer's logbook provides a good account of the design process but may be missing some important data. Although a suitable source of data for preliminary work, detailed analysis will require a more robust data set, which can be stated to be representative of the design process with greater confidence.

5.3 External validity

A further form of validity requires evidence that the scheme presents results that are representative of their intention; in this case that it does measure creativity within the designers' approach, and not some other characteristic. To complete this task the result must be compared to an independent measure that is accepted and proven; correlation will exist should the coding scheme be valid. In this study a creative style test similar to the KAI test [Kirton 1976] fulfilled this purpose. Correlation between those determined as more creative by the test and those determined as more creative by the coding scheme would suggest that the coding scheme is indeed measuring creativity.

5.4 Selection of a creativity test

As this coding scheme aimed primarily to study creative behaviour within the design process the most appropriate measure would also be one concerned with creativity. Within the psychological literature these are numerous developed tests from which to select, each of which measures slightly different characteristics and interprets in a slightly different way. For the purposes of this coding scheme at its current stage of development the test that was chosen was of creative style, similar to that developed by Kirton [1976] (*Kirton Adaption-Innovation test*, KAI). The KAI test was developed with the aim of identifying between two contrasting personality traits; that of the *innovator*, who strives to do things differently; and the *adaptor*, who strives to do things better. It is a relatively quick and simple test, the result of which is a classification of a person according to a spectrum from *innovator* to *adaptor*; the closer the person is to either end the stronger the characteristics of each present. When compared to the definition of creative used within this research, that of Howard et al. [2008], the characteristics of the innovator style (scores higher on the spectrum) bear closer resemblance. This view is corroborated by research, which has found a positive correlation between KAI score and other accepted creativity scales [Isaksen et al. 1988], showing a link between higher KAI scores and higher levels of creativity.

Table 3. Some behaviour descriptions of adaptors and innovators, from Kirton [1976]

Adaptor	Innovator
Characterised by precision, reliability, efficiency, methodicalness, prudence, discipline and conformity.	Seen as undisciplined, thinking tangentially, approaching tasks from unsuspected angles.
Seeks solutions to problems in tried and understood ways.	Queries problems' concomitant assumptions; manipulates problems.
Is an authority within given structures.	Tends to take control in unstructured situations.

When compared with the experimental results; several high, positive and significant correlations (Table 4) were found between the creative test and the creative behaviour displayed by the designers as measured by the coding scheme, showing that the scheme is valid in its identification of high levels of creative behaviour within designers and their process.

Table 4. Correlations between coded variables and the creative style test

Variable 1	Variable 2	Correlation	Significance (P < …)
Translational task creative behaviour (later stage)	Creative style test	0.844	0.00851
Latter stage creative behaviour	Creative style test	0.806	0.0143
Overall creative behaviour	Creative style test	0.657	0.0543

6. Discussion

Depending on the purpose for which the coding scheme is intended it may be developed in a variety of ways. In this case a primarily *deductive* method was taken, in which the formation of the scheme was heavily influenced by the underlying theory already present in the field. Thus the scheme has a strong base, well accepted within the literature, and is highly likely to produce useful results. Additional features of the data were highlighted using an *inductive* method, used to complement and provide additional understanding and definition within the categories developed from theory.

When developing the coding scheme rules, care must be taken in order to be considerate of the type of data to be coded. Depending on the form, the rules can be of varying levels of detail and explicitness without sacrificing reliability and validity; the further towards the latent projective end of the spectrum, the more the scheme must allow the coders to make their own judgements. As the data gathered within this study are of the *latent pattern* form, the rules within the scheme state the way in which the coder should identify various entities that are manifest within the data but do not state how they should interpret the relationships within. While the relationships possible are well defined within the coding scheme, it is the judgement of the coder that determines which applies based on the entities that are present in each task.

The important considerations when choosing the type of data to be coded are its accuracy as a data set, and its comparability to a realistic setting. For the purposes of this experiment the most practical option was that of a logbook study, allowing roughly chronological and complete analysis of a significant amount of work completed over a significant period of time. There is, however, some uncertainty in the determination of the data set as complete. Although past research states that the logbooks should be largely suitable, the possibility of work being completed in some other medium and not recorded, or of the designer formally recording work subsequent to its completion and hence out of chronological sequence are important considerations. In order to remedy this uncertainty it would be necessary to perform study that validates the content of the logbooks through the use of an alternative research technique. This process is currently on-going.

Without comparison to an accepted independent measure, it is impossible to claim that the results are representative of the purpose of the coding scheme and can be further studied or applied. Validation of this sort occurred through comparison with a creative style test similar to that of the Kirton Adaption-Innovation; and achieved high, significant correlations between creative behaviour as judged by the coding scheme and as judged by the test.
To ensure the results are repeatable and representative of the data from which they were taken and the coding scheme from which they were developed it is important to perform inter-coder analysis. Although adequate coefficients of agreement were produced (in excess of 0.75), this aspect could still be improved. Amongst the influences on reliability was that stemming from the form of data itself and the inherent ambiguity. While presenting a good record of the process, the unstructured form of marks on a page with only a general sense of the chronology of their application makes judgement difficult. While agreement was still sufficient in this case, further work will reduce ambiguity through higher detail of the time dimension attached to recording in logbooks.
Furthermore, using the original researcher as a coder is less than ideal; their detailed understanding and interpretation of the scheme and underlying theory may lead to different results beyond the scheme rules, and so are not completely representative. An attempt to balance this was made through the second coder, who was largely inexperienced and hence relied heavily on the scheme rules.

7. Conclusions

By presenting some of the important factors that must be considered when designing a coding scheme for use within content analysis, this paper has demonstrated part of the method by which good methodological practice can be achieved, thus helping to provide confidence in future research directions. Such considerations are of particular importance when performing smaller studies, which often do not provide the large data set required for analytical rigour.
Although designed and completed as a part of preliminary research, through understanding of how reliability and validity can be improved it has been possible to develop a scheme that produces good results, effective in the purpose for which it was designed, and with strong correlations to an independent, accepted measure.
Thus the results produced as part of this content analysis can be trusted, providing a strong grounding for the work completed to date and reliably describing future directions that it will take.

Acknowledgement

The work reported in this paper has been undertaken with support from the Engineering and Physical Sciences Research Council’s (EPSRC) Innovative Manufacturing Research Centre at The University of Bath (grant reference EP/E00184X/1).

References

Blessing, L. and Chakrabarti, A., "DRM, a Design Research Methodology", Springer, 2009.

Cohen, J., "A coefficient of agreement for nominal scales", Educational and psychological measurement, Vol.20, No.1, 1960, pp 37-46.

Currano, R. and Leifer, L., "Understanding idealogging: The use and perception of logbooks within a capstone engineering design course", ICED'09: International Conference on Engineering Design, Stanford, CA, USA, 2009.

Elo, S. and Kyngäs, H., "The qualitative content analysis process", Journal of advanced nursing, Vol.62, No.1, 2008, pp 107-115.

Gero, J. S. and Tang, H. H., "The differences between retrospective and concurrent protocols in revealing the process-oriented aspects of the design process", Design Studies, Vol.22, No.3, 2001, pp 283-295.

Guilford, J. P., "The structure of intellect", Psychological Bulletin, Vol.53, No.4, 1956, pp 267.

Hayes, A. F. and Krippendorff, K., "Answering the call for a standard reliability measure for coding data", Communication Methods and Measures, Vol.1, No.1, 2007, pp 77-89.

Howard, T. J., Culley, S. J. and Dekoninck, E. A., "Describing the creative design process by the integration of engineering design and cognitive psychology literature", Design Studies, Vol.29, No.2, 2008, pp 21.

Hsieh, H. F. and Shannon, S. E., "Three approaches to qualitative content analysis", Qualitative health research, Vol.15, No.9, 2005, pp 1277.

Isaksen, S. G. and Puccio, G. J., "Adaption-innovation and the Torrance Tests of Creative Thinking: The level-style issue revisited", Psychological reports, Vol.63, No.2, 1988, pp 659-670.
Kirton, M., "Adaptors and innovators: A description and measure", Journal of applied psychology, Vol.61, No.5, 1976, pp 622.
Klenke, K., "Qualitative research in the study of leadership", Emerald, 2008.
Kondracki, N. L., Wellman, N. S. and Amundson, D. R., "Content Analysis: Review of Methods and Their Applications in Nutrition Eduction", Journal of Nutrition Education & Behavior, Vol.34, No.4, 2002, pp 224.
Krippendorff, K., "Content analysis: An introduction to its methodology", Sage, 1981.
McAlpine, H., Hicks, B. J., Huet, G. and Culley, S. J., "An investigation into the use and content of the engineer's logbook", Design Studies, Vol.27, No.4, 2006, pp 481-504.
Milne, M. J. and Adler, R. W., "Exploring the reliability of social and environmental disclosures content analysis", Accounting, Auditing & Accountability Journal, Vol.12, No.2, 1999, pp 237-256.
Neuendorf, K. A., "The content analysis guidebook", Sage Publications, 2002.
Potter, W. J. and Levine Donnerstein, D., "Rethinking validity and reliability in content analysis", Journal of Applied Communication Research, Vol.27, No.3, 1999, pp
Snider, C. M., Dekoninck, E. A. and Culley, S. J., "Studying the appearance and effect of creativity within the latter stages of the product development process", DESIRE'11: The 2nd International Conference on Creativity and Innovation in Design, Eindhoven, Netherlands, 2011.
Sobek, D. K., "Representation in design: data from engineering journals", ASEE/IEEE: 32nd Frontiers in Education Conference, Boston, MA, 2002.

Chris Snider
Postgraduate Researcher
University of Bath, Department of Mechanical Engineering
Claverton Down, Bath, UK
Telephone: +44(0)1225 386131
Email: C.M.Snider@bath.ac.uk