Critically analyse two academic empirical articles from the perspective of research paradigms, research tools and ethics.

Paper 1:

Capturing quality in early childhood through environmental rating scales.

Kathy Sylva, Iram Siraj-Blatchford, Brenda Taggart, Pam Sammons, Edward Melhuish, Karen Elliot and Vasiliki Totsika.


Paper 2:

The Quality-Defining Process in Early Years Services: A Case Study.

Emily Tanner, Elaine Welsh and Jane Lewis.

Why choose these papers?

These papers were chosen as they share a key concept, that of quality in early years education and how that concept may be measured or identified. The focus of each paper is different, with paper 1 being concerned with some generalisation from quantitative analysis and with paper 2 having a more open and less rigid approach, possibly even exploratory in nature. They were chosen to illustrate how the concept of quality has moved on and that education is itself a dynamic paradigm, growing and changing with society.

Evaluative Framework.

What instruments and benchmarks do I use to evaluate these papers? Hammersly (2002) sets out an evaluative framework that looks at the case studied, the methods of data collection and analysis, the main claims and the conclusions drawn. The final test applied by Hammersly (2002) is what is the value of the knowledge the study provides. All of these will be kept in mind while analysing the papers at hand. As regards paper 2, Hammersly’s evaluative framework was written about a qualitative study and as such the issue of sample has been substituted for the cases studied in paper 1. Leininger (1994:105) also offers an evaluative framework, which will also be referred to. In addition to these frameworks, each paper will be reviewed in relation to the ESRC Framework for Research Ethics (2010) to assess the compliance with accepted ethical standards. Bryman (2004) offers several benchmarks for evaluating research. These are not codified into a specific evaluative framework, rather they are referenced throughout the book. Bryman (2004) will also be used to provide guidance and parameters in the evaluation of these papers. The use of multiple
evaluative frameworks and benchmarks allows some flexibility and objectivity in the assessment of these research papers. I will try to identify which paradigm each paper resembles most in an attempt to assess whether it displays characteristics associated with good examples of such. Prima facie, these papers both share significant common ground, namely the concept of quality within early years education, but it is noteworthy that they do not share any substantial theoretical foundation. The preceding literature review is very different for each paper and this begins to raise questions over quality of at least one paper even before I begin.

**Paper 1**

This paper explicitly looks at the relationship between the process quality characteristics of pre-school education and the subsequent developmental progress made by those students. The study is a longitudinal study and does not seek to establish a new definition of quality. The study makes extensive use of quantitative analysis and some qualitative analysis.

The literature review is substantial and leads the reader to the process of investigation used here. The extent of the literature review and the number of references adds some weight to the credibility of this paper. Hammersly (2002) asks what the value of the knowledge is and it would be fair to see the extensive material referred to in the literature review as suggesting there is at least *prima facie* value in the area of study. The specific issues raised in Sylva and Pugh (2005) and the Inter-departmental Childcare Review (2002) provide a solid and compelling rationale for this study. The study also identifies the absence of significant study of the nature they have conducted, with the only other significant study of Osborne and Milbank
Childhood Studies Essay Sample

(1987) being questioned over validity in Sylva (1994) and Sylva and Wiltshire (1993) because of procedural issues. This paper identifies the shortcoming in the existing knowledge base in respect of longitudinal outcomes from early years provision, citing Vandell, Dadisman and Gallagher (2000), Vandell (2004) and several others. This is set against the national strategy to expand early years provision DfEE/QCA (2000) that explicitly requires an expansion of early years provision. This all builds into establishing the potential value of the research at hand as required by Hammersley (2002). From this, I am inclined to see this research as valuable.

Hammersly (2002) suggests that we need to review the data collection process as part of any evaluation. The data collection process must begin with an analysis of the respondents. As this study falls predominantly within a quantitative paradigm, it is perhaps worth beginning with a review of the sampling used. Bryman (2004:97) suggests that a sample of over 1000 is sufficient to be representative of the population as a whole and which will generate confidence levels of 95%, provided it is genuinely random. This study sampled 2857 children across 141 pre-schools in England and a further 310 children from outside the pre-school system were also sampled. The regions from which the pre-schools were selected were selected strategically, that is not randomly. The pre-schools themselves were however selected randomly from those regions. What does this tell us of the sample process? The numbers are good, and exceed that required by Bryman above. However, we do see that such was not a random process and we have to ask to what extent this damaged the study, particularly in light of the generalisations that follow. Here I favour the evaluative criteria set out by McMillan and Wergin (2002:14) who ask whether the sampling process was clear and whether it impacted the results. I think the sampling process was clear and although the selection of regions was not
random, the selection of pre-schools within that as random overcomes this. If anything, the strategic selection of regions perhaps goes someway to avoiding regional trends that may arise if they had been selected at random, the number of regions available being small and subject to unintentional bias. It would be a fair assumption that regions have different but specific cultural issues and the lack of random selection for regions is not a threat to the validity of the claims made. The opposite is possibly true given. Bryman and Cramer (1990:10) suggest that a control group is important when establishing causality, and although this is not an absolute rule, Bryman (2004:41), it is established in this paper, where there is the sample of 310 children who were not part of the pre-school system, rather they were referred to as the home group. In section 2.2 there is an acknowledgment that a particular group were over sampled to provide information on ethnic minorities, but again the extent of this and the fact that it is noted tells me that it does not threaten the validity, particularly when the claims are reviewed.

Are the measurements and means of investigation appropriate? The issues involved do lend themselves to longitudinal studies and the fact that this paper is part of a longitudinal study is a positive. While I accept the position of McMillan and Wergin (2002:15) that longitudinal studies are not always practical, this study has overcome those practical difficulties and generated some worthwhile conclusions as a result. This study specifically needs to examine early years input and the outcomes later in life if it is to achieve its aim. It does that. There is perhaps room for some cost benefit analysis. If you accept the position taken by Hazelrigg (2004:85) that there are diminishing marginal returns in respect of accuracy as your sample grows, and Bryman (2004:97) who states that a random sample of 1000 will be very representative of the population as a whole, then could the same research have
been conducted with just a random sample of children? This a budget concern though, rather than an issue with the research per se. The research instruments used are a revised version of the Early Childhood Environment Rating Scale, per Harms, Clifford and Cryer (1998) (ECERS-R). Using a previously used research instrument can allow you to use a tried and tested process and this is of value, provided you don not contaminate your data. This is echoed in Lieberman and Remedios (2007:379) where they use an existing research instrument for similar research previously conducted by Elliot and McGregor (2001). Bryman (2004:36) supports the idea that this is good practice provided no data contamination happens.

This paper also made use of in depth case-studies and multi-level modelling and this sets the study up well to triangulate the findings from a range of differing techniques. Multi-method research provides the researcher with the opportunity to see shortcomings in their work that may arise through a particular method. Further, here it goes someway to ascribing meaning to the social phenomenon observed. This is important if you accept the positions of Silverman (2001:32), Denzin and Lincoln (2005:2) and Kardorff and Steinke (2004:3) who all suggest that social research is required to understand as well as observe.

So what of the claims made? Bryman and Cramer (1990:7) set out three criteria to establish causality, an apparent relationship between the variables, the relationship to be non-spurious and for the cause to precede the effect. The educational outcomes in this paper are being measured at age 5, after the early years provision, the relationships identified and claims made all have significant statistical data to support them and there are several caveats that are dealt with in section 5. The claims all seemed relevant to the title and the abstract. Bryman (2004:104) and Lunt and Livingstone (1992:173) suggest that you cannot generalise beyond your
population, this paper makes claims about children in England specifically and this would seem to fit with this idea. The claims are set against a control group and as mentioned above, this helps internal validity. The data was dealt with through multi-level modelling, a process of analysing the same data in relation to different claims and this paper again made use of an established model for this through Rasbach et al (2000). Importantly the paper identifies the limits of the research in section 5, where it highlights the fact that the measures it was using for quality were limited and even went as far as offering alternative and more extensive models of quality in Munton, Mooney and Rowland (1995). Further, there is an acceptance of a need for context when assessing quality and an acknowledgment of the work of Dahlberg, Moss and Pence (1999) in such. This suggests balance and a degree of pragmatism in the claims being made and it should be viewed as a positive.

This piece of work certainly appears to be compliant with the six principals set out by the ESRC Framework for Research Ethics (2010:3) and there is nothing to suggest that there have been any breaches of this code. In fact, as you read the methodology section, which is relatively substantial, I am left with a sense of confidence that this work has been conducted in a spirit of compliance.

**Paper**

This paper is explicitly seeking to extend the definition of quality within early years education from that set out within the official definitions of the National Childcare Standards (1998). It puts itself forward as a case-study, with the purpose of exploring what quality really means in this context. So is this paper part of the qualitative paradigm and if so, does it exhibit good practice? Denzin and Lincoln (2005:7) suggest that qualitative research is not tied to any particular method. However,
Bryman (2004:267), citing Whyte (1955) and Gans (1962), suggests that qualitative research may be associated with ethnography, observation and interviews, structured and unstructured. Denzin and Lincoln (2005:5) identify case-studies as part of the qualitative paradigm. The methodology section of this paper highlights interviews with groups of people from different stakeholder groups. The total number of interviews seems to be 53 respondents, but this is not entirely clear from the text. So, *prima facie* it would seem that this paper is part of the qualitative paradigm. However, it is possible to question this. If a case-study is designed to be genuinely exploratory, then it is worth considering Bryman (2004:273) who states that reliance on existing theories or parameters is something you associate with the quantitative paradigm, not the qualitative and definitely not an exploratory case-study. Does the use of an existing and official definition of quality in the NCS render this outside of the paradigm or the scope of an exemplifying case-study? I do not think it is that serious, but it will undoubtedly shape the findings and the usefulness of this study. It would be my position that this study would have worked better if the existing and official definition of quality had not been addressed until the concluding section, deciding then if it would fit with the findings made in the course of this research. The idea that the definitions were to be built upon and extended suggests that this study is less than exploratory and that perhaps it had a significant amount of bias built in to the data used and the conclusions made. If we accept the qualitative paradigm, Warren (2002:99) suggests that a study should have a minimum of 20-30 interviews to be meaningful. Gerson and Horowitz (2002:223) say that less than 60 interviewees cannot support meaningful conclusions. Bryman (2002:335) says that less than 50 interviewees is too small a sample. But Morse (2004a) discusses the difficulties in deciding on a number of interviewees to make a study meaningful,
arguing that context is more important than numbers. While this paper may meet some of the minimums cited above, the point raised by Morse (2004a) is pertinent and we are unable to judge the context without a transcript or summary of the interview findings. This is a significant weakness in this paper and it makes evaluation against the criteria set out by Hammersly (2002) and Leininger (1994) difficult if not impossible. Of particular concern is the very brief outline of methodology, it simply highlights the different stakeholder groups that were interviewed and whether they were interviewed over the phone or face to face. Is this sufficient to feel confident in the findings and conclusions? If you refer to Leininger (1994:105), then a good research paper will be believable and confirmable. The first being how true the claims made are and the second being how this is evidenced through repeated affirmations. In the absence of transcripts, a coded summary or similar, then the confirmability and the truth are both absent here and this raises questions over the validity of the findings. This paper therefore falls short in respect of the tests laid down by Leininger (1994). Hammersly (2002:182) specifically asks what the value of the knowledge is in an attempt to evaluate it. If the truth and confirmability of this work are questionable, then the value may also be questionable. I would argue that where truth and confirmability are questionable, then the conclusions cease to be knowledge, rather they become anecdotal or worse, speculative. Specifically I would want to know whether these interviews were structured, semi-structured or unstructured, how were the interviewees chosen and how were their responses collated and analysed before I would be prepared to affirm this as research over and above a series of anecdotes. This issue is raised by both Silverman (2001:34) and Bryman (1988:77) as a risk of weaker qualitative research and case-studies specifically. It is my position that this paper displays this
characteristic and is guilty of the anecdotalism discussed. On page 8, there are several assertions that could be read as generalisations by a reader that was not concentrating, I am unclear as to whether the author is seeking to assert generalisations or whether this is simply a record of things discovered within this case-study. This ambiguity is itself a weakness, clarity of thought is vital, especially when trying to pin down the dynamic and subjective concept of quality. Moving on from that, if the ideas referred to are generalisations, then without further detail how can we tell if there is sufficient data to support the ideas. While I accept the points of view expressed by Flick, Von Kardorff and Steinke (2004:5) cited in Denzin and Lincoln (2005:3) and Silverman (2001:32) that social research ought to ascribe meaning and understanding to social phenomenon and Lincoln and Guba (1985:316) who say that qualitative social research should focus on rich and detailed description of social events, this paper falls short. There is no rich description, there is no real meaning ascribed to the social phenomenon identified, rather we have a series of points raised by interviewees. I would be specifically concerned about how the various opinions raised in this paper were selected and whether they were representative of the majority of opinions and ideas expressed. The only tangible finding that comes from this paper is that there are different ideas of what constitutes quality in the context of early years education. This very thin finding and the specific nature of the cohort being studied leaves me questioning the value as required by Hammersly (2002) of this piece of work. The finding itself may seem reasonable to most practitioners, but there is still very little data to support such. Even if we accept the position of Bryman (2004:285), Williams (2000:215) and Mitchell (1983:207), that case studies may not always produce generalisations, the reality is that even these
more pragmatic approaches refer to the need for cogent theoretical reasoning to underpin the findings. I am not sure that such exists in this paper.

In support of this paper, I like the model of a quality continuum and I think the idea of pluralism within the definition of quality is eminently sensible if not essential. However, I do not think this piece of research establishes either of these concepts with any degree of certainty. Ethically, the research appears to conform to the guidelines set out in ESRC Framework for Research Ethics (2010), but again the lack of information surrounding methodology means it is impossible to comment on the extent of compliance or otherwise.

**Conclusion**

These papers both examine the concept of quality within early years, however that is where their similarities end. Their theoretical foundations, their methodology and their purpose are all quite different. It is clearer if we deal with each paper separately. In paper 1 the literature review and context for the study is substantial and it establishes the value of the study very well indeed. The sampling does raise a question over whether the strategic selection of regions damaged the validity of claims made, but my assessment is that the strategic sample of regions coupled with random sampling therein probably added to the validity rather than detracted. This was a fair departure from some of the orthodoxy on this matter. The paper is multi-method and the analysis is both quantitative and qualitative and this provides some assurance that the findings are valid and credible. The claims made and the route to them seem to comply with the benchmarks of quality used by Bryman (2004), Hammersly (2002) and Leininger (1994). There is an acceptance of the limits of the study, and because these are addressed they cease to be weaknesses and become
strengths. There are no apparent breaches of ethical protocols. Overall, Paper 1 is a strong piece of well executed research. It adds to the knowledge base in an area that probably needs development and so has real and tangible value. There is perhaps more reliance on statistical data for the claims made and the qualitative data is perhaps under used, but this is a very subjective point. As regards paper 2, does it do what it sets out to do and is this of value? If paper 2 is to explore, then it ought to set aside the pre-existing definitions and explore the different stakeholders’ notions of quality first, comparing and contrasting with official definitions in the conclusion. The use of the official definition as a starting-point, does little more than fetter this study and pre-determine the findings. I rely on Bryman (2004:24) to support the idea that the reliance on existing theory or definitions is part of quantitative analysis. If that is the case, then the research is not truly exploratory and the sample size it has falls a long way short of acceptable per Bryman and Cramer (1990:8-10). This paper is neither exploratory nor is it supported by enough existing theory to be anything else. The absence of detail in respect of the methodology leaves the trust in the findings damaged. The findings themselves are little more than a collection of what some respondents have said. This paper displays those things that characterise weak case-study research as identified by Silverman (2001:34) and Bryman (1988:77). If you accept that point, then the tests used by Leininger (1994) and Hammersly (2002) cannot be passed. If the reader cannot trust the research findings then there is not truth or confirmability and therefore the value of the research is questionable. This paper did not add to the knowledge base, it did not establish the rationale for the study and the claims made were not predicated on solid foundations as far as the reader was concerned. The lack of information supplied also made it difficult to assess the compliance with the code of ethics required. Overall, paper 2
lacked a clear direction, clear method and the process was ambiguous with the findings being unsupported and of questionable value.

References:


London: Strategy Unit


Whyte, W. F. (1955) *Street corner society : the social structure of an Italian slum.* Chicago, Univ. of Chicago P.