DOES RESEARCH PROVIDE REAL ANSWERS?

Alan A MacKENZIE
Does Research provide real answers?

Reviewing papers of three research projects whose goals were to examine the effects of parental substance misuse on child psychopathology.

by A. A. MacKENZIE
7/31/2007
INTRODUCTION

There is substantial evidence of the concomitant effect of parental alcohol misuse on children through the impact of alcohol misuse on family functioning and parenting. In general, affected families perceive their environment to be less cohesive, lack ritual and routines, tend not to positively express feelings, warmth or caring (either physically or verbally), and have higher levels of unresolved conflict. Some studies (Scher et al., 1991; Hill et al., 1996; Merikangas et al., 1998; Moss et al., 2002; Clark et al., 2004; Berger & Osbourne, 2005) have attempted to demonstrate that the effects of parental substance misuse on child psychopathology appear to be cumulative and transmissible. The longer the child is exposed to parental alcohol & drug misuse, the greater the impact may be to a variety of risks and accumulation of risk factors (Brook & Tseng, 1996; Brook, Tseng, & Cohen, 1996; Stanger et al., 2004).

Disruptive behaviours by children, such as aggressiveness, hyperactivity and mental health problems, are particularly apparent in the sons of parents who both misuse alcohol (Schuckit & Smith, 1996; Clark et al., 1997). Several studies have also found alcohol consumption increases aggression (Hill & Muka, 1996; Barnow et al., 2002), leading to (a) increased marital conflict and corresponding abuse of partners and children and (b) mounting household turbulence (Eiden, Edwards & Leonard, 2004; Grekin et al., 2005). Antisocial personality disorders and mood disorders appear to be associated most strongly with alcohol and drug misuse by parents, which are also strongly correlated to neurobehavioral disinhibition in these family members (Clark et al., 2005).

This review considers four studies which focus in on the inter-relatedness of parental drug & alcohol abuse and its effects, both short- and longer-term, on children within and outside family systems. The overarching theme is the interplay between existing dependencies and psychopathologies, and the central question being asked: “Is little Johnny better off if Daddy (or Mummy) is not around?” My personal interest is to explore, where there has been integration of positivist and constructivist approaches, whether critical reviewers need be concerned with things like the conceptualisation of triangulation, abductive logic, or questions concerning the reconciliation of positivist and constructivist epistemologies.
**Davis (1997)**

Diane Davis attempted to “flesh out the bones” of earlier findings about the recovery process of alcoholic women, exploring further “what works” with women (inferring gender differences in treatment regimes). The study is severely limited by a microscopic, nonrepresentative sample population (only 10 women who attend AA groups) in a snowball sample, yet she claims to give voice to women’s experiences in areas currently of interest to researchers and practitioners, namely: (a) initiation of recovery – how is it done; (b) roles of social support networks; (c) socioeconomic influences, and; (d) single gender AA groups – do they make more a difference? Davis chose a quasi-ethnographic (constructivist) approach to “uncover partialities”, assisting “external parties in getting an insider’s view of these women and their stories”. The study’s aim was to “provide a more descriptive, richer context of the experience around the issues that can be provided by quantitative findings”. Hence, it focusses on the process more than the product, making meaningful conclusions that contribute to understanding about socio-economic and psychological factors in alcoholic recovery.

**Moss, Baron, Hardie and Vanyukov (2001)**

These co-researchers tested the hypothesis “that pre-adolescent offspring of substance dependent fathers (with Anti-Social Personality traits) would exhibit greater levels of (a) “externalization” and (b) disruptive behaviours”. Looking closely at the prevalence of specified psychiatric disorders and measuring increases in the child’s internalizing and externalizing psychopathology, the study attempted also to uncover associations between parental behaviour, personality, and the child’s resultant phenotype (i.e. substance use behaviour). Measurement techniques were backed up with prolific citations, establishing recognized diagnostic protocols.

Subject selection was purposive (non-random) and sampling of the female cohorts seemed insufficient. As the majority of the SD+ fathers were recruited from community treatment facilities, a risk of sampling bias exists. Statistical analyses of their findings included multivariate linear model analysis with incredible p-values (0.002< \( p <0.006 \)) with two-to-four degrees of freedom, respectively.
Of the offspring psychiatric disorders and the major depressive mood disorder chosen for examination, relative to *paternal* mental diagnoses, Chi-squared analyses revealed co-occurrence of elevated dimensional scores for all categories, but the SD+/ASP+ youth scores were significantly higher for both *internalizing* and *externalizing* behaviours, which seems counter-intuitive at first blush. Yet, these unexpected results revealed some interesting findings, pointing to *these* children becoming “higher risk” anti-social substance-abusers by adolescence and into adulthood.

**Wilens, Biederman, Bredin, Hahesy, Abrantes, Neft, Millstein and Spencer (2002)**

Some earlier & later studies argued that the association between parental substance abuse and negative child well-being is due to selection, or the antecedent characteristics of parents who abuse drugs or alcohol (Wilson, Desmond, & Verniaud, 1973; Clark, Cornelius, Wood, & Vanyukov, 2004). “*One of the most powerful methods to elicit antecedent disorder info is high-risk design*”, or so claim the researchers of this control-based study. Ninety-six families were recruited through hospital contacts and both parents & offspring were examined for levels of *psychopathology, cognitive* and *social* functioning, under varied family structures (nuclear and non-nuclear families). Six separate ‘assessment tools’ (SCID-II, KSADS-E, WISC-R, SAICA, FES and GAF) measured parent & child functioning and competency levels, also helping clinicians with their diagnoses -- whilst logistic regression models (i.e. Pearson’s Chi-squared 2-tailed tests, to a *p* ≤ 0.05) were used to test for correlative significance in all three groups (alcoholic, drug-dependent and non-SUD [control]).

Results supported earlier (similar) findings and support the above-mentioned 2001 study; in that children of parents who use drugs (opioid-dependence) have significantly higher conduct disorder rates, mood disorders, and SUD’s compared to controls. Also, more school dysfunction was noticed in the kids with SUD parents, especially the opioid-dependant ones. High risk children also showed a diminished capacity for self-regulation, via aggression, inattention and impulsivity.

Four specific hypotheses, based on two assumptions, were studied examining relationships between alcohol education, alcohol consumption and treatment-seeking attitudes in college students. 878 students were randomly selected and 813 were interviewed by the CARS (computer-aided response survey) method. The self-reporting nature of this interviewing technique may have made it possible for interviewees to underreport. Their method of stratifying the systematic random sampling is also questionable.

Although the original hypotheses assumed that alcohol education would increase student’s abilities to recognize and effectively deal with alcohol consumption & decrease associated problems, the research results revealed just the opposite pattern. In short, the authors learned that problem recognition by students with severe alcohol problems (self-admitted) did not lead to an increased willingness to seek treatment for their addiction.

RESEARCH INQUIRY PARADIGMS AND DESIGN

The design of any research study begins with the selection of a topic and a paradigm. A paradigm is essentially a worldview, a whole framework of beliefs, values and methods within which research takes place. It is within this worldview researchers conduct their work. The researcher asks what the overarching research question is; what the specific research questions are that will ensure the thorough consideration of the overarching question; which epistemological and theoretical lenses he/she will use to guide the study; which research methods will be used to answer the research questions; how the material gathered will be analysed; and the way in which the culminating findings will be presented.

Good qualitative research uses a tradition of inquiry. This means that researchers identify, study, and employ one or more traditions of inquiry. Cupchik (2001) statement points to traditions of inquiry:

*Phenomena are understood as processes which cut across the physical, social, and personal (self) worlds. Qualitative and quantitative researchers examine these phenomena, offering rich descriptive accounts or precise analyses of functional relations, respectively. It is*
assumed that both approaches to research practice face the problem of constructing "data" and are therefore subject to potential bias. While description has traditionally been viewed as preceding hypothesis testing (i.e., natural history precedes hypothesis testing), the two approaches [can be] viewed as complementary and in parallel. Qualitative methods offer an in-depth account of underlying processes and can help frame hypotheses that test specific functional relationships, while empirical findings related to processes can suggest areas which might benefit from detailed descriptive examination.

The outworking of this can be seen in the four studies being investigated. Yu, Chin Evans and Perfetti (2003), Wilens et al. (2002) and Moss et al. (2001) all reflect a strong positivist paradigm, adopting in the latter two particularly, qualitative accounts which they used primarily to “frame hypotheses that were then tested for specific functional relationships” before resorting to more conventional British Empiricist and Enlightenment traditions of inquiry (i.e. randomized control group post-test only designs). Two of the studies (Yu, Chin Evans and Perfetti, 2003; Wilens et al., 2002) actually employed purposive (non-random) sub-population sampling techniques. Moss et al. (2001) did not mention qualitative data at all in their results; these were left to the intro and discussion sections of their report. Their study focused on testing (and grouping) psychiatric disorders (psychopathology) and scoring both internalizing and externalizing problem behaviours within each group.

All three studies held such a strong ‘scientific’ commitment to their designs, in that the outcomes were restricted to measurable parameters; failing to recognize the complexity and strictures of juxtaposing objective fact onto subjective judgments, as Cupchik (2001) again argues here:

*The juxtaposition of objective fact with subjective opinion is sometimes a struggle. Pragmatic wisdom dictates that the two domains interact: suggestions that arise in the external world meet connections in the self … In short, the boundaries between outside and inside are illusory and predicate a dichotomy between external physical fact and correlated sensory knowledge; a distinction that does not readily generalize to a social world of hearts and minds.*

*Worlds are multilayered with many levels of interacting structures ongoing simultaneously. Phenomena are physical or social events, or episodes, that take place in the world, apprehended by some or all, and
which are more clearly explained or understood when placed in appropriate contexts that bring them into sharp relief. Physical phenomena can exist without human apprehension but they only become meaningful events, in the sense of influencing action, when noticed or observed by a group of people, however small. [Psychiatric phenotypes] are contextualized phenomena which are perceived intersubjectively and defined as such. Phenomena are therefore events that unfold and recur in the flow of time and are only meaningful when understood in context; they are processes and not essences. ...phenomena are difficult to observe because they are not restricted to sense-data but involve the application of judgment. ...In order for an event to take on the quality of a salient phenomenon, it must be perceived as possessing a coherent structure and observed against [an appropriate] background. [Brackets mine]

Moss et al. (2001) attributed complex associations between paternal behaviour, personality and the child’s resultant phenotype to an aggregate of complex influences referred to as transmissible factors.

In contrast, Davis (1997) approached her research from a phenomenological (interpretivist) paradigm. Aside from a few descriptive stats about the background of her respondents, she employed a narrative thrust to theorizing à la Bochner (1997, 435) who maintains there is nothing as theoretical as a good story… “[T]here is no split between theory and story when theorizing is conceived as a social and communicative activity. In the world of social theory, we are less concerned about representation and more concerned about communication. We give up the illusions of transcendental observation in favour of the possibilities of dialogue and collaboration.” Davis chose “semi-structured in-depth interviews” to collect her data (Minichiello, 1990, 89) aiming for “in-depth and holistic understanding, in order to do justice to the complexity of social life.” (Punch, 1998, 243) The study yielded ‘multiple meaning structures through non-intellectual, embodied, dialogical conversations’ (including inquirer self-disclosure) similar to that described by Shotter (1997, 22):

[O]ur relating ourselves to and understanding of other people ... comes about in a nonintellectual, embodied, dialogical, or conversational manner, in which an embodied, temporally unfolding, responsive form of understanding... is at work. What is especially important about this dialogical form of understanding is that it is not an individual
achievement. It is an understanding developed and negotiated with others in the circumstances...

In this way, significant findings weren’t clouded by over-analysis of empirical data and, therefore, the qualitative functionality of key themes uncovered (namely how shame impacts recovery and the visions of recovery) could be contextualized. In some ways, this research holds a feminist bent, couched in a movement away from a preoccupation of the necessary connection view (causation). Punch (1998, 51-2) quoting Lincoln & Guba, says causation views relate to a need for prediction, control and power.

**METHODOLOGY**

Methods chosen in any study will generally spring from the research paradigm followed and the research (how, what and why) questions based from it. Pragmatically, the way sampling is done and data collected depends on whether the questions, methods and data should be *prespecified* versus how much they should *unfold* (emerge) as the research develops. It also depends upon whether the simple model of research has a hypothesis or not (Punch, 1998, 23, 42).

**SAMPLING**

Three of the four studies adopted opportunitistic sampling procedures (i.e. “snowball”, clinically referred and by other sampling means not described) by selecting smaller sample populations; these sampling methods were guided by theoretical rather than probabilistic considerations. Only Yu, Chin Evans and Perfetti (2003) chose stratified (systematic) random sampling with populations sizable enough for statistical inference. This makes sense, though, as theirs was a primarily quantitative study of four hypotheses, which Punch (1998, 242) says requires “*typically larger samples than in qualitative studies, and generalization through sampling is usually important.*”

A second consideration comes out of Davis’ “snowball sampling” method. According to Cohen (1989, 67-8) tracing respondents using a snowball sampling procedure with randomized sequences…does not guarantee a random sample:

*but it seems to be the best available way to select a cross-section of substance users who are as representative as possible (Korf and Van*
Poppel, 1986; Korf, 1987a). A main problem of the snowball method is the selection of the first respondents (the zero stage). According to the methodology developed by Biernacki & Waldorf (1981; Biernacki, 1986) the main effect of a specific choice of respondents in the zero stage is its predetermination of sequent respondents.

By carefully choosing the zero stage respondents in circles other than those of alcoholics “working their AA program” Davis had hoped not to deviate from the sample population desired. She might have used an overview of the snowballs per interviewer, indicating clearly how many linear and bilinear chains were constructed in her sampling process.

Another consideration was ethnicity and gender bias (not so for Wilens et al. (2002) or Davis (1997), as the former chose to study male subjects only and the latter female subjects only). However, in Moss et al. (2001) and Yu, Chin Evans and Perfetti (2003), the sex distribution (according to Table 1 in both) was >58% male versus <14% female and 53% female versus 47% male, and the ethnicity was 77.1% and 82.2% Caucasian, respectively. Bias could arise from the possibility that addictive behaviours and aggressive traits (i.e. affective disorders) could be strongly influenced by gender.

**DATA COLLECTION**

Wilens et al. (2002) and Moss et al. (2001) similarly assessed all children and adults by structured psychiatric interviews (collecting inventories, ratings, scales, questionnaires and interview information) using five & four diagnostic tools, respectively; meeting, in some instances, rigid DSM-III (IV) -R criterion. These studies used observational data that counted the incidents (%) of particular behaviours in children. To this, Yu, Chin Evans and Perfetti (2003) added attitudinal survey data scored on a five-point Likert Scale. All but Davis (1997) used survey data to profile subjects; by contrast, she used semi-structured interviews with her respondents, giving a detailed, rich thumbnail sketch of the “typical” woman form the sample population.

As Moss et al. (2001) and Yu, Chin Evans and Perfetti (2003) appear to be mixed studies (both quantitative & qualitative); both using data collection strategies which offered rich qualitative data taken by more than one method and form. Triangulation of data along with videotape transcriptions provided enhanced verification (validity and
reliability of the data collected). The latter study also sought codified data in some instances taking away from the complexity of problem recognition in respondents. The interview protocol for obtaining variable RECOGNITION came from an open-ended question “Have you ever thought you had a problem with alcohol” (Yu, Chin Evans and Perfetti, 2003, 676) requiring more than just a “Yes” or “No” answer. In addition, the question is suspect on subconscious suggestibility grounds, as people who consume copious quantities of alcohol may have at the root of this problem low self-esteem due to stigmatization, as the following quote confirms:

“The concept "labels" relates to the so-called labeling theory that refers to... negative labels originating from the individual's self-evaluation... Heckert & Best (1997) and Fife & Wright (2000) show that labeling leads to stigmatizing, which in turn leads to low self-esteem... labels regarded as negative unfortunately have a more powerful impact than the input of positive labels on the subconscious mind of individuals.” (van Zyl, Cronjé, and Payze, 2006, 182)

By contrast, Davis (2001) used a semi-structured interviewing to collect her qualitative data, so as to allow for structure in the data to emerge during analysis.

In terms of using specific structured approaches for recording clinical data, Moss et al. (2001) chose to specify the best way to collect diagnostic data, by employing the diagnostic consensus review of all available information. This procedure apparently minimizes disagreements over the presence of psychiatric symptoms, similar to the argument expressed by Marta Del-Ben et al. (2004, 62): “The use of recorded interviews made by experienced professionals, observation of interviews through unidirectional mirror, and the practice of joint interviews, for example, may minimize the discrepancies between the diagnosis elaborated by the professional-to-be and the experienced professional.”

Electronic recordings also allow for gathering “data that can be analyzed later, and repeatedly, in minute detail” (Punch, 1998, 185) allowing for accurate transcribing to occur throughout the study and then by other interested parties later (i.e. for secondary analysis of qualitative interview data see Goodwin and O’Connor, 2006, 375).
DATA ANALYSIS

Common to all these studies being reviewed is interview data collected and analyzed. The analytic status of interview data can present its own set of problems, according to Punch (1998, 182), since “they are never raw, but are always situated and textual” and are influenced by “the characteristics of the interviewer, including race, class, ethnicity and gender” according to Denzin & Lincoln (1998, 353).

In the studies being reviewed, the way in which the data was analyzed most strongly reflected the researcher’s paradigmatic approach. Yu, Chin Evans and Perfetti (2003) and Moss et al. (2001) both chose to use multi-variate regression analyses; skewed continuous data in the latter were first transformed into normally distributed variables using n-Log transformations. Similarly, in-between group differences were tested by $\chi^2$ analysis, whereas for conditions demonstrating a significant linear association, post-hoc 2X2 analyses were performed.

Logistic regression models were employed by Wilens et al. (2002) for comparative analyses of psychopathology and social functioning. Any ordinal data was analysed using ordinal logistic regression models, after being adjusted with Huber’s formula (for non-independence). Pearson’s $\chi^2$ tests were employed in groups who lacked respondents meeting diagnostic criteria. Other information, like demographic and cognitive functioning data, were analysed using inferential statistics (two-tailed, with $p<0.05$) and means & standard deviations listed by SD/ASP grouping, and reported either graphically or in tabular or text formats; all tables were quite explicit and easy to read and interpret. The three studies mentioned thus far demonstrated a careful selection of statistical tests.

The three correlational studies focussed mainly on data analysis for quantitative data collected; qualitative observations were mentioned in passing in the discussion and conclusion sections of their reports. Davis’ (1997) reporting read more like a novel, so it was very difficult to extract just how she collected or analysed her data. One brief statement mentions recording of interviews, and she adds, “transcribed interviews were coded into core themes... and organized by emergent themes by the author”, to which Punch (1998, 204) admits “for analysis directed at discovering regularities in the data,
Miles and Huberman’s (1994, 4) framework for QDA, employing three concurrent streams interwoven and concurrent throughout the analysis - data reduction, data display and drawing conclusions - involving coding, memoing and propositioning, could have been employed in these particular studies to demonstrate a commitment to the validity of the quantifiable data collected (see Punch, 1998, 202-08).

**RELIABILITY, VALIDITY AND TRUSTWORTHINESS**

Reliability basically means consistency, presenting us with the idea of replicability or repeatability of results or observations. Since reliability and validity are rooted in positivist paradigm, they should be redefined for their use in any naturalistic approach. While examining reliability, validity and trustworthiness, Golafshani (2003, 601) explains:

...the terms reliability and validity are essential criterion for quality in quantitative paradigms; however, in qualitative paradigms... Lincoln and Guba (1985, 300) use "dependability" in qualitative research, which closely corresponds to the notion of "reliability" in quantitative research. To ensure reliability in qualitative research, examination of trustworthiness is crucial. ...while establishing good quality studies through reliability and validity in qualitative research, Seale (1999, 266) states that the "trustworthiness of a research report lies at the heart of issues conventionally discussed as validity and reliability".

Constructivism values multiple realities that people have in their minds. Therefore, to acquire valid and reliable multiple and diverse realities, multiple methods of searching or gathering data are in order. Triangulation is a step taken by researchers to involve other investigators or peer researchers’ interpretation of the data at different times or locations. Triangulation is often defined as “a validity procedure where researchers search for convergence among multiple and different sources of information to form themes or categories in a study” (Creswell & Miller, 2000, 126). In this way, a qualitative researcher can “use investigator triangulation and consider the ideas and explanations generated by additional researchers studying the research participants” (Johnson, 1997, 284).

In the three correlational studies (Moss et al., 2001; Wilens et al., 2002; Yu, Chin Evans and Perfetti, 1997) errors of measurement in the observed variables through
improper use of clinical “tools” may pose a threat to the validity of causal judgements. The extent to which these errors in observed variables influence the validity of conclusions will depend very much on the properties of the errors of measurement. If such errors occur at random, measurement errors will reduce statistical precision by increasing estimates of standard errors and risks of Type II statistical errors. However, if errors of measurement are systematic, design validity may also be affected.

Implicit in judgements about causal relations is the assumption that these statements apply to some well defined population in which X is believed to be causally related to Y. The validity of such claims depends on the extent to which the sample population being described is a representative sample of the general population. Therefore, it seems logical that departures from random sampling assumptions will introduce threats to validity into designs. Triangulation and verification of the causal judgements on the basis of correlational evidence to relevant literature is another way of building trustworthiness. An example of this is found in Wilens et al. (2002, 47), “our findings of elevated rates of psychopathology in SUD youths... are consistent with the literature of offspring of parents with SUD.”

**FUTURE RESEARCH**

From the discussion and conclusions of these four studies, it is clear that there has been some detailed research into many aspects of the lives of parents who misuse substances and the impact on their families. The impact of, and risks associated with, parental substance misuse appear to have been well mapped. Accumulation of risk associated with certain factors, such as domestic violence, marital break-up, unemployment, deprivation etc., have also been highlighted.

Unlike these studies, other research has found no evidence of heightened risk for children stemming from parental substance misuse alone. Following on from this, a philosophical shift in the literature towards resilience is occurring and this has clear potential when applied to children, and other family members, affected by (parental) substance misuse. According to a recent scoping study released by a Scottish Advisory Group to the British Parliament (2006, 1,2):
The key gaps in the literature included the following: children's views (particularly in relation to impact, resilience factors, service needs, or views on existing service provision), fathers, siblings, mental health, rurality and ethnicity. These are key areas where future attention should be directed. This work must consider the gender and age of children. It is also clear that it is often the problems that are associated with or arise from (parental) substance misuse that can have a greater negative impact on the family than the misuse per se, and hence there is a need to view parental substance misuse as part of a far wider, multi-dimensional, picture. Finally, clear and methodologically sound attempts to measure and validate the numbers of children and families affected by substance misuse are severely lacking.

Wilens et al. (2002, 48) echo some of these same concerns in their statement “clearly, further evaluation of familial contribution for the various psychiatric and cognitive disturbances in youths is warranted.” I would add to this the need for longitudinal research studies, which focus specifically at youth-at-risk, in line with Moss et al. (2001, 277) comment “an enhanced understanding of the life-course trajectories of these high-risk youth could improve treatment approaches…”

This genre of research is about how some people’s lives touch other people, and the ripple affects it has on the overall welfare of others. As such, research of this ilk needs to redress the more serious and the far-reaching implications of salient treatments - leading researchers into the interpretivist arena in future - doing more studies similar to those of Davis (1997) which speak of the “Human Story” and the realities of human experience. Quantitative surveys cannot remain the dominant mode in which research of this nature is done; qualitative research, being much more eclectic, allows for multiple strategies and methods than quantitative studies. The range of what can count as useful data, and ways of collecting this data, is much broader in scope. Only systematic qualitative studies like that grounded theory research and phenomenology and can allow much deeper exploration of the common themes within the great diversity in the realm of psychopathology and addictions.
BIBLIOGRAPHY


