



What really is effective in interviews with suspects? A study comparing interviewing skills against interviewing outcomes

Dave Walsh^{1*} and Ray Bull²

¹School of Law and Criminology, University of Derby, UK

²Forensic Section, School of Psychology, University of Leicester, UK

Purpose. The present study examines whether interviewing in a manner that is compatible with a recommended model of interviewing (called the 'PEACE' model) impacts on outcomes of the interviews examined in contrast to prior studies which have usually been concerned whether the conducted interviews were fair and not coercive or whether there had been positive effects of training upon subsequent interview performance.

Methods. This study, examining in detail 142 actual suspect interviews, is set in the barely researched area of social security benefit fraud, reflecting current trends in Britain concerning increasing numbers of interviews with suspects undertaken by public sector organizations and the pluralization of policing.

Results. It was found that good interviewing in each of the recommended stages that make up the PEACE model generally led to better interviews, indicating the importance that each stage contributes to overall interview quality. Further, the quality of interviews was compared against a range of interview outcomes and it was found that skilled PEACE interviewing was associated with the securing of full accounts, including confessions.

Conclusion. Given the few examples of skilled interviewing found in the study it is argued that further training of investigators is necessary to improve both interviewing performance and organizational outcomes.

The questioning of suspects in England and Wales has undergone major transformation since the mid-1980s after the introduction of legislation, namely the Police and Criminal Evidence Act (PACE) and the mandatory taping of interviews with suspects. In the early 1990s, concerns had been raised by the findings of several studies in that, whilst coercion had largely disappeared following the introduction of PACE, it had not necessarily been replaced by better interviewing methods (e.g., Baldwin, 1992, 1993; Moston, Stephenson, & Williamson, 1992). Over the next few years, however, a huge

* Correspondence should be addressed to Dave Walsh, School of Law and Criminology, University of Derby, Kedleston Road, Derby DE22 1GB, UK (e-mail: d.walsh@derby.ac.uk).

roll-out of a new model of police training took place. Further, government investigators (who were also required to interview under PACE) received training in the new model by the early part of the new century following similar initiatives to increase professional standards. Several police studies have demonstrated, with certain reservations, that the interviewing of suspects in England and Wales is being better undertaken than it was before the introduction of this new model (e.g., Bull & Cherryman, 1996; Bull & Soukara, 2009; Clarke & Milne, 2001; Griffiths, 2008). However, prior research has not examined substantially if there is any link between interviewing performance when using this new model and the investigation outcome. This is purpose of the present study.

The PEACE model

This new model, called the PEACE model, was a mnemonic acronym for the sequential phases of the model where 'P' refers to the necessary 'Preparation and Planning' before the interview. Thereafter, 'E' stands for the initial 'within' interview 'Engage and Explain' phase covering the legal requirements and the opportunity to build rapport between interviewer and suspect. Next is the 'Account' phase (later, Account, Clarify, and Challenge), pivotal to the interview, where open questions were to be put inviting an account of the events from the suspects themselves. If this questioning highlighted inconsistencies (with either the existing evidence or from within the suspect's actual account) further probing should then occur to resolve these queries. This phase would then be followed by the 'Closure' phase, being the final stage that was to be conducted in the interview itself. This is where a summary of what had been said before should be put to the suspect, allowing the suspect to either add or modify to what had been said before putting (if appropriate) any criminal charges to the suspect. The final phase, 'Evaluation', occurs after the interview, and is meant to allow self (and senior officers', judges', etc.) reflection concerning the effectiveness of the interview as well as determining if further enquiries are necessary.

An initial evaluation of the PEACE model, undertaken by McGurk, Carr, and McGurk (1993) found certain improvements in police officers' interviewing skills such as questioning and listening skills. However, a later study by Bull and Cherryman (1996) noted concerns in interviewer performance particularly in regard to (i) rapport building, (ii) empathic approaches towards suspects, (iii) summarizing, and (iv) conversation management, the latter being one of the bedrocks of the PEACE model. Whilst these two studies provide useful findings, they were conducted during the time of the implementation of the new training in England and Wales. It might be argued therefore that learned skills had not had sufficient time to be fully embedded into regular police practice.

However, Clarke and Milne (2001) undertook a study some years later. They found interviewer strengths that lay in the areas of delivering the caution, communication skills, and self-confidence. On the other hand, their study also revealed continuing weaknesses (for example, in several key tasks in the Engage and Explain phase, including building rapport). Clarke and Milne also found that police interviewers did not perform competently when challenging the suspect's account or when summarizing interviews. Following Clarke and Milne's study, there was a revised approach to training by way of a structure 'tiered' in accordance with the nature of the officer's duties with those who undertook more serious investigations receiving advanced training. More recently however, Griffiths (2008), in his baseline audit of police interviews, continued to find frailties in interviewing skills which he largely attributed to a lack of planning.

Benefit fraud investigations

An increasing number of government agencies in the UK conduct criminal investigations. However, the amount of research conducted into their actual performance remains negligible. Walsh and Milne (2008) did undertake a study of 99 real-life social security benefit fraud interviews with suspects. These were conducted by the central government department responsible for social security benefit administration, the Department for Work and Pensions whose Fraud Investigation Service (FIS) investigates the majority of such suspected criminal infractions. Walsh and Milne found that training had not significantly improved interview behaviours in many of the skills required for effective interviewing (e.g., rapport building, summarizing, and communication skills).

However, this study took place during the introduction of the PEACE training for fraud investigators and techniques learned during training arguably still remained to become established in everyday practice. Furthermore, roughly half the sample of investigators in the study still had to receive training at the time of the study. On the other hand, many of the interviews were found to cover relatively simple matters where the suspects were clearly naive of such investigative interviews appearing not to have personally experienced them before. The suspects also often seemed very recalcitrant, confessing to the crimes with little or no urging from interviewers. Many of the interviews in the study also involved cases where the evidence was incontrovertible and the size of the fraud was, at most, modest, having not been carried out with any significant amounts of ingenuity by suspects. Whilst this might be seen as a typical sample of benefit fraud cases it hardly presented, in hindsight, a true measure of the fraud investigators' skills where they are most likely to be tested (i.e., in the more challenging interviews such as those involving more complex frauds and/or resistant suspects).

In light of these concerns, Walsh and Bull (in press) studied a further sample of, generally, more demanding benefit fraud interviews some years after the training had been implemented. In their sample of 142 actual interviews with suspects conducted by both FIS and local government investigators (who also investigate benefit frauds), it was found that interview performance was mediocre with few examples of skilled interviewing. Walsh and Bull's study concluded that interviewing skills needed to be much improved if the investigations into benefit fraud could be hailed as highly professional. One might question why such indifferent performance seems not to be a major concern for these organizations? This may be because the FIS regularly meets its set targets of sanctions (i.e., prosecutions and, where the offence is felt to be less serious, cautions, or administrative penalties; Department for Work and Pensions, 2005). Moreover, given that the vast majority of investigations are not of a really serious nature, and clearly not as grave as the most serious of police investigations, any miscarriages of justice are unlikely to attract the same level of media and public (or even, management, ministerial, or governmental) concern.

There may also be further reasons for this limited attention to the actual quality of interviews. For example, only 35% of FIS sanctions are set as a target for prosecution (and less than 20% of sanctions delivered by local government investigation teams are actually prosecuted; Department for Work and Pensions, 2005, 2008). In all other sanctioned cases, the lesser diversionary options are applied. Consequently, only a minority of cases are open to external scrutiny by virtue of their placement in, and procession through, the wider criminal justice system. This might suggest that for those cases which are prosecuted interviewing quality is being monitored. However, decisions to prosecute can employ quite separate criteria that do not necessarily measure interview performance. Further, concerning those cases that are otherwise

sanctioned, McKeever (1999) argued that the employment of diversionary measures from the criminal justice process, such as cautions and penalties, might lead to some contestable practices. His concerns lay with officials authorizing cases involving many 'perpetrators' who are naive of the criminal justice process and where there is debatable assurance that these investigations would have otherwise have led to prosecution (a legal requirement). Moreover, legal representation of suspects occurs much less often in benefit fraud interviews than police ones, further underlining that officials could influence the direction (and outcome) of many investigations regardless of their soundness. Prior research into benefit fraud interviews (e.g., Walsh & Bull, in press; Walsh & Milne, 2008) has found that legal representation was present in around only 8% of benefit fraud interviews, much less often than in police interviews, where legal representation occurred at a rate at least five times greater (Arnold, 2007; Clarke & Milne, 2001).

The findings from Walsh and Bull's (in press) study suggest that the reasons for the consistent meeting of increased targets over the past 10 years is not as a result of a widespread skilled workforce in those organizations dealing with those offences (though some skilled interviewing occasionally occurred in their study). As was found in Walsh and Milne's (2008) study, a persistent soft 'underbelly' of cases exists in benefit fraud investigations. These cases tend to culminate in diversionary disposals that do not enter the advanced stages of the criminal justice process, being exclusively authorized and delivered by officials from the investigating organization. As a result, there may well be institutional delusion in terms of actual skill levels amongst many of the organizations that investigate benefit fraud. In short, a tendency exists which measures the quality of interviewing performance through the lens of the quantity of sanctions output, which can be a quite separate measure. Simply, these outputs may not necessarily reflect skilled interviewing (Bull & Soukara, 2009; Pearse & Gudjonsson, 1996).

Walsh and Bull's findings were presented to those professionals responsible, nationally, for *delivering interview training* to benefit fraud investigators both from FIS and the local government investigation teams, and to the *inspectors* from the (now defunct) Benefit Fraud Inspectorate. There was general agreement that their findings were consistent with these professionals' views of interviewing practices out in the field. However, when these findings were discussed by the first author with *managers* and *investigators* during a series of presentations throughout England during 2007, a different (defensive) reaction was received. These were expected, given that (i) the findings, though identifying some strengths, revealed many areas of interview performance that required substantial improvement and (ii) investigators tend to overestimate their own skills' levels (Walsh & Milne, 2007). Some of the responses from these investigation personnel, it was noted, were based on their confidence of meeting their set targets and/or their belief that the PEACE model was not always a suitable model for benefit fraud interviews (Shawyer & Milne, 2009; Walsh & Bull, 2009). In response to these views, the current study examined whether the better interviews (i.e., those that followed the PEACE model more faithfully) produced higher quality interviews and outcomes.

An issue that needs further clarification is the meaning of the term 'effective', which may mean different things to different audiences concerning interviews with suspects. For instance, in the 15 years or so since the introduction of the PEACE model in England and Wales there have been no reported miscarriages of justice as a result of oppressive interviewing of suspects. This is in stark contrast to the situation that preceded its implementation. Consequently, there might be justifiable confidence that on this basis

the model is effective. An alternative measure of effectiveness is that PEACE has elevated interviewing performance. Again, successive research into police interviews suggests that, whilst certain interview behaviours remain problematical, by and large, the interviewing of suspects is being better undertaken than it was before the introduction of the PEACE model (Bull & Cherryman, 1996; Bull & Soukara, 2009; Clarke & Milne, 2001; Griffiths, 2008; Griffiths & Milne, 2006). However, the focus of the present study was to examine any relationship between effective interviewing performance under the PEACE model and the actual investigation outcome, an area where there has been insufficient research.

Any examination of what is effective in fraud interviews in terms of actual results when measured against better PEACE interviewing performance, of course, pre-supposes that the behavioural competencies that constitute this model have some relationship with the term 'quality interviewing'. Space limitations prevent a fuller discussion of this issue but, briefly, the current study relies on the relevant literature (see Bull & Milne, 2004; Griffiths, 2008; Milne & Bull, 1999; Ord & Shaw, 1999; also see Bull & Cherryman, 1996, who asked police officers what constituted good interviewing). The PEACE model is argued here to be a fairer, more ethical template for interviewing suspects that requires a higher domain of skills in obtaining an accurate account of what happened than, say, those employed in trying to trick guilty (or otherwise) suspects into confessions (see Inbau, Reid, Buckley, & Jayne, 2001, but also see Kassin, 2005; Leo, 2008).

Method

One-hundred and fifteen taped interviews and a further 27 transcripts of actual benefit fraud interviews with suspects conducted by either FIS or local government fraud investigators between 2004 and 2007 were examined. The tapes were randomly sampled from five FIS teams and eight local government investigation teams operating throughout England, representing a mix of rural, urban, and 'inner city' investigations (for more comprehensive details of the sample; see Walsh & Bull, in press). When requesting the sample, a stipulation was made that any forwarded interviews should not be less than 15 min in duration (so as to examine over a suitable period of time the skills actually being exhibited). It was also asked that the sample should include both those cases which failed to progress after the interview as well as those that did advance to one form of sanction or another.

An instrument of measurement for the particular purposes of this study was adapted from the rating scale used in Clarke and Milne's (2001) influential study (which in turn had evolved from prior research, e.g., Bull & Cherryman, 1995; McGurk *et al.*, 1993). Tables 1–3 reflect that a number of behavioural characteristics were assessed along an ascending five-point Likert scale where the lowest score of '1' denoted that it was felt that there was a need for further training in that behaviour and where a score of '5' reflected a 'highly skilled' performance. A 'satisfactory' score of '3' was defined as the minimum acceptable score for competent performance of the behaviour (where performance though not rated as exceptional was not weak either).

Inter-rater reliability

The skill levels along the rating scale and expected outcomes were initially assessed by the first author, possessing over 20 years of investigative experience and who was also

Table 1. Results of *t* tests: Engage and Explain phase tasks compared to (1) Engage and Explain phase quality, (2) overall interview score, and (3) interview outcomes (mean scores given with SD in brackets)

Competency	Engage and Explain phase				Interview score				Interview outcome			
	At or above PEACE standard	Below PEACE standard	<i>t</i>	<i>r</i>	At or above PEACE standard	Below PEACE standard	<i>t</i>	<i>r</i>	Comprehensive account or confession	Denials or partial admission	<i>t</i>	<i>r</i>
Delivering caution	3.31 (0.53)*	2.76 (0.64)*	− 5.50	.43	3.13 (0.53)	2.98 (0.76)	− 1.39	.12	3.14 (0.55)	2.99 (0.72)	− 1.36	.12
Checking under- standing of caution	2.30 (1.31)*	1.58 (0.95)*	− 3.67	.30	2.25 (1.16)*	1.76 (1.21)*	− 2.38	.20	2.05 (1.21)	1.93 (1.22)	− 0.60	.05
Advising legal representation	2.97 (0.82)*	2.56 (0.67)*	− 3.21	.27	2.87 (0.83)	2.70 (0.68)	− 1.33	.11	2.73 (0.80)	2.79 (0.72)	0.50	.04
Explaining interview purpose	2.58 (0.97)	2.07 (0.90)	− 3.21	.27	2.13 (0.97)*	2.46 (0.94)*	2.04	.17	2.21 (1.00)	2.38 (0.94)	1.00	.08
Advising routines and routes	1.33 (0.56)	1.34 (0.58)	− 0.10	.01	1.36 (0.58)	1.33 (0.57)	0.37	.03	1.32 (0.58)	1.36 (0.57)	0.36	.03
Explaining that interview is opportunity to give account*	1.82 (0.63)*	1.34 (0.53)*	− 4.89	.39	1.69 (0.70)	1.50 (0.60)	− 1.73	.15	1.68 (0.74)	1.51 (0.57)	− 1.50	.13
Building rapport	2.64 (0.69)*	1.77 (0.59)*	− 7.95	.56	2.39 (0.78)*	2.05 (0.73)*	− 2.70	.22	2.27 (0.75)	2.15 (0.78)	− 0.85	.07

Note. *denotes findings of significant differences where $p \leq .05$.

Table 2. Results of *t* tests: account phase tasks compared to (1) account phase quality, (2) overall interview score, and (3) interview outcomes (Mean scores given with *SD* in brackets)

Competency	Account phase				Interview score				Interview outcome			
	At or above PEACE standard	Below PEACE standard	<i>t</i>	<i>r</i>	At or above PEACE standard	Below PEACE standard	<i>t</i>	<i>r</i>	Comprehensive account or confession	Denials or partial admission	<i>t</i>	<i>r</i>
Encourages suspect to give account	3.00 (0.64)*	2.30 (0.66)*	− 6.38	.48	2.97 (0.64)*	2.34 (0.78)*	− 5.25	.40	2.85 (0.76)*	2.47 (0.72)*	− 2.97	.25
Develops topics	2.92 (0.64)*	2.07 (0.65)*	− 7.78	.55	2.88 (0.64)*	2.09 (0.77)*	− 6.52	.49	2.81 (0.72)*	2.20 (0.74)*	− 4.84	.38
Deals with difficulty	2.72 (0.70)*	1.84 (0.66)*	− 7.68	.55	2.72 (0.72)*	1.86 (0.71)*	− 7.04	.51	2.51 (0.79)*	2.07 (0.77)*	− 3.24	.27
Uses appropriate structure	2.91 (0.86)*	1.68 (0.78)*	− 8.86	.60	2.95 (0.83)*	1.69 (0.82)*	− 8.96	.61	2.71 (1.03)*	1.94 (0.90)*	− 4.66	.37
Uses appropriate questions	2.92 (0.67)*	1.88 (0.74)*	− 8.69	.60	3.03 (0.61)*	1.80 (0.72)*	− 10.97	.68	2.71 (0.88)*	2.11 (0.82)*	− 4.12	.33
Keeps interview to relevant topics	3.25 (0.56)*	2.42 (0.66)*	− 7.89	.56	3.23 (0.56)*	2.45 (0.74)*	− 7.08	.52	2.98 (0.76)	2.69 (0.71)	− 2.30	.19
Intermittently summar- izes and links	1.74 (0.78)*	1.27 (0.56)*	− 4.01	.32	1.75 (0.77)*	1.28 (0.59)*	− 3.97	.32	1.60 (0.78)	1.42 (0.64)	− 1.50	.13
Covers points to prove	2.66 (0.89)*	1.69 (0.67)*	− 7.24	.53	2.75 (0.88)*	1.69 (0.65)*	− 9.25	.62	2.69 (0.88)*	1.81 (0.74)*	− 6.39	.48
Explores motive	2.48 (1.03)*	1.62 (0.62)*	− 5.60	.43	2.53 (1.05)*	1.61 (0.72)*	− 5.84	.45	2.47 (1.05)*	1.72 (0.80)*	− 4.75	.38
Challenges appropriately	2.72 (0.67)*	1.72 (0.65)*	− 8.95	.61	2.72 (0.69)*	1.78 (0.73)*	− 7.73	.55	2.51 (0.79)*	1.99 (0.80)*	− 3.78	.31
Explores received information	2.88 (0.60)*	1.96 (0.67)*	− 8.45	.59	2.85 (0.63)*	2.00 (0.73)*	− 7.22	.52	2.76 (0.63)*	2.13 (0.77)*	− 5.05	.40

Table 2. (Continued)

Competency	Account phase				Interview score				Interview outcome			
	At or above PEACE standard	Below PEACE standard	<i>t</i>	<i>r</i>	At or above PEACE standard	Below PEACE standard	<i>t</i>	<i>r</i>	Comprehensive account or confession	Denials or partial admission	<i>t</i>	<i>r</i>
Conversation manage- ment skills	3.00 (0.56)*	1.99 (0.59)*	− 10.40	.66	2.93 (0.61)*	2.08 (0.73)*	− 7.42	.53	2.76 (0.69)*	2.26 (0.76)*	− 3.93	.32
Cognitive interviewing skills	1.98 (0.86)*	1.26 (0.47)*	− 6.09	.46	2.03 (0.86)*	1.24 (0.48)*	− 6.43	.48	1.85 (0.85)*	1.42 (0.66)*	− 3.24	.27
Uses pauses and silences	1.42 (1.21)	1.29 (0.63)	− 0.79	.07	1.40 (1.15)	1.33 (0.84)	− 0.43	.04	1.62 (1.11)	1.20 (0.85)	− 2.49	.21
Possesses self- confidence	3.12 (0.65)*	1.86 (0.65)*	− 11.41	.70	3.09 (1.67)*	1.98 (0.80)*	− 8.71	.60	2.82 (0.82)*	2.24 (0.89)*	− 3.89	.32
Possesses open- mindedness	2.94 (0.56)*	1.92 (0.61)*	− 10.21	.66	2.87 (0.62)*	2.01 (0.74)*	− 7.24	.53	2.73 (0.71)*	2.18 (0.76)*	− 4.27	.34
Possesses flexibility	2.92 (0.71)*	1.45 (0.55)*	− 13.73	.76	2.95 (0.65)*	1.53 (0.71)*	− 12.18	.72	2.65 (0.91)*	1.82 (0.87)*	− 5.44	.42
Possesses communi- cation skills	3.05 (0.51)*	1.84 (0.57)*	− 13.01	.74	3.03 (0.52)*	1.91 (0.72)*	− 10.74	.68	2.76 (0.79)*	2.17 (0.74)*	− 4.35	.35
Displays active listening skills	3.06 (0.68)*	1.99 (0.73)*	− 8.93	.61	3.13 (0.60)*	2.00 (0.80)*	− 9.26	.62	2.84 (0.86)*	2.29 (0.84)*	− 3.74	.30

Note. *denotes findings of significant differences where $p \leq .05$.

Table 3. Results of *t* tests: Closure phase tasks compared to (1) Closure phase quality, (2) overall interview score, and (3) interview outcomes (Mean scores given with *SD* in brackets)

	Closure phase				Interview score				Interview outcome			
	At or above PEACE standard	Below PEACE standard	<i>t</i>	<i>r</i>	At or above PEACE standard	Below PEACE standard	<i>t</i>	<i>r</i>	Comprehensive account or confession	Denials or partial admission	<i>t</i>	<i>r</i>
Competency												
Supplying final summaries	1.91 (1.15)*	1.12 (0.44)*	− 3.17	.26	1.44 (0.89)*	1.08 (0.46)*	− 2.86	.24	1.33 (0.82)	1.18 (0.59)	− 1.33	.11
Reading out final caution	2.00 (1.57)*	1.24 (1.42)*	− 2.25	.19	1.69 (1.55)	1.06 (1.34)	− 2.51	.21	1.80 (1.41)*	1.01 (1.42)*	− 3.23	.27
Advising what happens next	2.77 (0.92)*	1.51 (0.81)*	− 6.54	.50	1.96 (1.08)*	1.46 (0.81)*	− 3.17	.26	1.73 (1.09)	1.67 (0.87)	− 0.35	.03

Note. *denotes finding of significant difference where $p \leq .05$.

very familiar with the relevant research literature. The second rater, an even more experienced benefit fraud investigations professional, received coaching on the rating scale by the first author. Fuller details of the intra-rater reliability are provided in Walsh and Bull (in press). Briefly, it was found that, after using Cohen's (1988) guidelines, strong correlations existed between the two independently conducted ratings regarding the overwhelming majority of the assessed competencies and the overall interview score with none of the correlations found to be weak. Thus, confidence can be expressed in the ratings attached to the behavioural dimensions. Strong statistical significance (using Spearman's rho) was also found between the two raters concerning their anticipated outcomes of the investigation following the interview ($r = .88$, $N = 140$, $p \leq .05$). Moreover, the judgment, concerning the investigation outcome, was only made by both raters once all the behavioural assessments had been conducted for each interview.

Hypotheses

It was considered that the level of skills demonstrated in the interviews would be linked to the 'overall' level of interview skills *and*, further, that there would be an association between the level of 'overall' interview skills and the outcome of the interview. These outcomes were defined along a range of five categories beginning with interviews that result in 'no comment' outcomes. Next are those interviews where 'denials' were accepted though not fully tested by the interviewer. Thirdly, are those interviews that yield merely a 'partial admission' where it remains highly debatable how much offending had occurred or whether the case has real grounds to proceed due to the lack of exploration concerning the suspect's partly acknowledged offending. Fourthly, those interviews where admitted wrongdoing had been fully examined and explored or, where denials were received, these were tested as much as one could reasonably have expected, were designated as 'comprehensive accounts'. The final category concerns interviews that culminate with a 'confession' being accepted without either investigator or the suspect going into much detail. The sample of tapes was examined in order to find (a) what, if any, association existed between interviewer performance in the first four phases of the PEACE model (i.e., (i) Preparation and Planning, (ii) Engage and Explain, (iii) Account, Clarify and Challenge, and (iv) Closure) and the 'overall' interview assessments, (b) which particular behavioural competencies were most strongly related to the quality of performance in the Engage and Explain, Account, and Closure phases, and (c) whether any association existed between interview performance and interview outcome. Not all the phases of the PEACE model were assessed in every interview due to their brevity. For example, though the vast majority of interviews were 15 min or more in length, a small minority of the sample were considerably shorter (e.g., one was of 4 min duration) and this clearly did not allow a full opportunity to measure performance.

Results

In this section the findings will be presented separately in each of the first four phases of the PEACE model. Firstly, there will be a comparison of the assessed behaviours in each interview phase with the assessed overall skill levels in that phase. The t tests were conducted with the two groups of interviewers being defined as those performing at

PEACE standard or above (i.e., those who were rated at either '3' (i.e., 'satisfactory'), '4', or '5' (assessed as 'skilled' or 'highly skilled', respectively) against those assessed at below the expected benchmark for competent PEACE interviewing (i.e., those interviewers assessed as 'needing training', rated at '1' on the rating scale, or were 'not quite adequate' in their abilities, and rated at '2'). Secondly, using the same demarcations of performance, the assessed behaviours will be matched against the overall quality of the interview.

Finally, in each of the first four phases of the model, the findings will compare skill levels in each of the interview phases against the interview outcome. Again, *t* tests were used with the categorization of the two groups partitioned between differing sets of outcomes denoted by quality criteria. For example, either eliciting a comprehensive account or a fairly obtained confession (whether or not lacking supportive detail) are argued to be more qualitative outcomes than 'no comment', 'denials', or 'partial admissions'. It is contested that these first two outcomes represent situations where either (i) confessions have been fairly obtained that leave little or no doubt about their veracity (whether either a confession or a denial occurred) or (ii) satisfaction exists that the interview has been as thorough as it should have been under an ethical interviewing framework. These two outcomes should also be seen as interview conclusions that are more desirable towards organizational targets providing a credible belief that those who confessed did so voluntarily because they were guilty. These outcomes would also provide confidence that those interviewing possessed higher order investigative interviewing skills so that, even in cases where denials were received, it was evident that those denials had been robustly, but fairly, tested. On the other hand, those three identified outcomes categorized as possessing less quality (and therefore, desirability) were those involving (i) non-responsive and totally silent interviewees in a 'no comment' interviewee (which, incidentally, did not occur in this study), (ii) 'denials' that were not robustly tested and therefore it still remained open to conjecture whether or not those denials represented any veracity, and (iii) 'partial admissions' which were not extensively examined or tested but clearly treated as full and frank ones by interviewers whereas actually they were not.

Preparation and Planning

What constitutes good examples of planning? It might be argued that because this task, by its very nature, is a pre-interview activity, measurement as to how well it has been attended to could not be comprehensively made entirely by examination of the interview alone. However, it is contended that an accurate representation of planning skills can be ascertained when experienced assessors examine how much (and, in turn, how well) the interviewer (i) possesses an understanding of the suspected offence(s) and the legal points to be proved, (ii) considers any potential defence mitigation (legitimate or otherwise) and has prepared contingency responses to these, (iii) ensures evidence and exhibits are available and has paid regard to their order of presentation, and (iv) considers the quality of the tactics and strategies employed by the interviewer when questioning the suspect.

After examining the sample, it was found that associations existed between the better preparation and planning phase skills and 'overall' interview quality. An independent sample *t* test found a significant difference in preparation and planning skills between those interviewers performing overall to at least PEACE standard ($M = 3.15$, $SD = 0.84$) and those who fell below this standard ($M = 2.18$, $SD = 0.86$;

$t(137) = 6.68$, $p < .01$, $r = .50$). This is a significant finding. Moreover, once the interview outcomes were grouped into the two categories as defined earlier (i.e., the more qualitative and those felt to be less desirable), a further significant difference was found between the former group ($M = 2.95$, $SD = 0.95$) and the latter ($M = 2.35$, $SD = 0.91$; $t(137) = 3.76$, $p < .01$, $r = .30$). These findings suggest that good preparation positively affects both overall interview quality and, subsequently (although somewhat more tentatively in view of the smaller effect size), the result of the interview itself.

Engage and Explain

To assess which of the particular competencies undertaken in this phase (listed in Table 1) had more of an impact on the rated quality of the Engage and Explain phase, a multiple regression analysis was conducted. A significant model emerged ($F(6, 134) = 42.2$, $p < .0005$; $r = .66$) reflecting that almost two-thirds of the variance in the rating for this phase was accounted for by the behavioural competencies required to undertake it. Significant beta values were found for the following variables; (i) building rapport (where, $\beta = 0.49$, $p \leq .0005$), (ii) explaining the interview purpose ($\beta = 0.27$, $p \leq .0005$), and (iii) checking the suspects' understanding of the caution ($\beta = 0.22$, $p \leq .0005$). Multi-collinearity was not found between the independent variables, more indicative of the discrete nature of many of the Engage and Explain phase's competencies.

Further analysis of the scores in this phase found that after conducting t tests (see Table 1) there were significant differences concerning all but two of the behavioural competencies between those interviews where the Engage and Explain phase was assessed at or above PEACE standard and those assessed less favourably. Though it may appear on the face of it that several of these tasks undertaken at this stage of the interview are inextricably linked, Table 1 shows that the competency, by some margin, that yielded the largest significant difference when examining both interview quality and outcomes (and the largest effect size) was rapport building. This emphasizes how crucial this task is to the Engage and Explain phase of the interview.

Account

Of those skills assessed in this phase (see Table 2), good communication skills and demonstrating flexibility in the interview were significantly found to be the main predictors of quality in the Account phase (where, in regard to communication skills ($F(19, 119) = 27.1$, $p \leq .0005$, $r = .81$; $\beta = 0.333$, $p \leq .0005$) and for flexibility ($\beta = 0.245$, $p \leq .0005$)). However, these findings reflect that no single variable was individually strong, and multi-collinearity, though not excessive, nevertheless, was located amongst many of the behavioural dimensions. The findings displayed in Table 2 also reveal that, after conducting t tests, significant differences and stronger effect sizes between behaviours and the quality of the phase were found in all areas of performance particularly those concerning (i) conversation management skills, (ii) a confident demeanour, (iii) open mindedness, (iv) an ability to be flexible, and (v) skilled communication.

The latter two behavioural competencies, along with skilled questioning styles were notable as possessing significant differences and strong effect sizes, when examining the performance levels in those tasks and comparing them to the overall interview

quality (see Table 2). Table 2 also shows that several other significant associations were also found amongst other performed tasks (albeit with more modest effect sizes). Further examinations undertaken revealed that 90% ($N = 66$) of those interviewers that were assessed as possessing insufficient skills in the Account phase were also poorer interviewers overall. Moreover, 82% ($N = 53$) of those rated as 'satisfactory' or 'skilled' in this particular interview phase were also those whose "overall" interviewing skills were assessed as at least at PEACE standard. Aggregated, these findings may well underline the contention that whilst all phases of the PEACE model have importance, the skills required in the Account stage are more vital in achieving high quality interviewing.

When turning to the actual interview outcomes, 63% ($N = 25$) of those rated as 'satisfactory' or 'skilled' interviewers in the Account phase, either obtained a comprehensive account or a full confession, whereas of those who needed further training, 88% ($N = 22$) received denials or only partial admissions. The various competencies were also examined to determine whether any were more prominent in relation to the differing outcomes (as defined in the Engage and Explain phase above). From Table 2, it can be seen that many tasks involved in the Account phase were performed at significantly higher levels when obtaining the 'better' outcomes, again providing currency to the belief that the tasks, effectively carried out in this phase, may be highly influential in deciding the interview's outcome.

Closure

Table 3 reflects an overall lack of quality in this interview stage. When combined with the regular omission of key tasks that should be undertaken by investigators in the Closure phase this rendered multiple regressions as rather meaningless in trying to establish which of the tasks were likely to predict a better Closure phase and an 'overall' interview quality. However, Table 3 shows significant differences when comparing assessed behaviours in the better quality closure phase performances though our findings should be judged on a provisional basis in light of the effect sizes attributable to the infrequency that acceptable skills actually occurred in this phase. Table 3 indicates, however, that there was a significantly better exhibition of skills in the better quality interviews relating to two of the tasks in this phase of the interview, however a cautious note is again exercised in light of the limited effect sizes.

The mean scores in the current study, when associating Closure skills to interview outcomes, continue to illustrate the widespread poor performance in this interview phase. Table 3 shows that there were no meaningful differences found in the level of displayed skills in the tasks required to be performed in this phase when examining them against the two sets of interview outcomes.

Overall

Figure 1 illustrates that effective demonstration of the skills demanded by the PEACE model is indeed actually related to the more effective interview outcomes. Sixty-three percent ($N = 38$) of those assessed as at least at PEACE standard in their 'overall' interviewing ability secured comprehensive accounts or full confessions. None of the 'skilled' interviewers obtained denials without also obtaining a comprehensive account from suspects, re-affirming that amongst the essential elements of a good interviewers' repertoire is the ability to elicit and fully test suspects' accounts of events. On the other hand, 92% ($N = 22$) of the weakest interviewers failed to obtain either a comprehensive

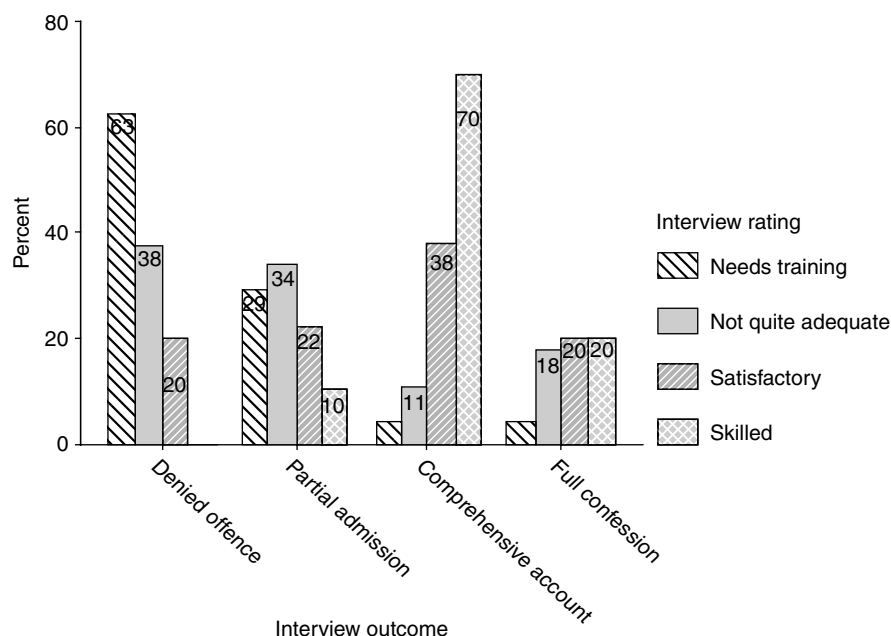


Figure 1. Association between the PEACE interviewing skills and interview outcome.

account or a confession. An association between the quality of the interview and its outcome was found to exist with those who attained PEACE standard and above in their overall interviewing skills ($M = 2.79$, $SD = 0.75$) significantly more likely to obtain the more desirable outcomes than those who fell below this standard ($M = 2.00$, $SD = 0.75$; $t(137) = 6.09$, $p < .01$, $r = .46$).

Discussion

The research sought to find out whether adhering to the PEACE model would lead to 'better' outcomes (i.e., either obtaining a comprehensive account or a confession). The finding that there was a relationship between higher skill levels and the more positive outcomes creates an opportunity for senior civil servants and local government managers to obtain higher yields from the staffing resources they assign to combating fraud. Alternatively, if (politically?) they feel that 'sufficient' results are currently being obtained, then the findings from the present study will identify those factors which will enable these managers to maintain the same amount of results but with lesser resource expenditure via a reduced workforce! At the very least, the findings from the current study have identified those skills that are needed to integrate quality into interviews with suspects.

The study consistently found that those interviewers able to perform well in interviews, particularly in the Preparation and Planning, and Account phases, are better interviewers 'overall' and are substantially more likely to obtain a confession or yield a comprehensive account than those less able performers.

Performing at satisfactory levels alone in the Account phase is insufficient to influence interview outcomes emphasizing that the higher skilled interviewers are more

likely to be successful in obtaining either comprehensive accounts or full confessions. Moreover, those interviews resulting in denials were associated with (although not necessarily caused by) poorer interviewing skills. This may well suggest that the ineffectiveness in results is associated with the ineffectiveness in interviewing skills. The current study found that only a few of those needing training actually gained a comprehensive account from suspects. It is argued that the guilt presumption that Walsh and Milne (2008) found amongst investigators persists which leads to insufficient or unskilled planning ahead of interviews which in turn creates problems when suspects are resistant in interviews. Furthermore, despite contrary evidence found in other studies (e.g., Shawyer & Milne, 2009; Walsh & Bull, in press; Walsh & Milne, 2008), investigators tend not to see themselves as lacking in interviewing skills (Walsh & Milne, 2007). This failure to recognise accurately their own shortfalls in their interviewing skills found in these other studies may well, therefore, limit any opportunity for immediate improvement. Where skilled interviewing occurred in the current study, it was found that those 'skilled' interviewers were seven times more likely to yield fulsome details from suspects when compared to the 'poorer' interviewers. Quite simply, skilled planning leads to skilled interviewing in the Account phase which in turn leads to better outcomes.

The findings from the current study also suggest that there appears to be no single indicator in the Account phase that influences eventual interview outcomes but that there is an interplay between competencies that seems to be influential. Bull and Soukara (2009) also found that no one single tactic was dominant in this phase of the interview. Perhaps, given the range of skills and tactics that need to be employed by the investigator in the Account phase, this is how it should be.

This does mean, however, that the most skilled of interviewers are those that are able to perform the range of tasks in this phase rather than just a selection of them. Since some of these tasks are demanding then it should be not be so surprising that those 'poorer' interviewers flounder and lose confidence either in the face of denials or when the fraud is more complex. These concerns also emerge elsewhere in this study. For example, whilst an accusatory approach by interviewers was rarely found, there was a tendency to find submissive, unassertive styles with (i) little preparation and planning (again) and (ii) repeated questioning in the vain hope that such a tactic might cause the suspect to confess. These findings are reminiscent of the police studies of Baldwin (1992, 1993) and Moston *et al.* (1992) who also found aimless questioning, ineptitude and submissiveness by the police in the face of denials. Furthermore in the present study, when investigators did receive vague acknowledgements of guilt there tended to be a rush to finish the interview, closing it hastily in the hope that what they had was enough to achieve a sanction.

Surprisingly, half the suspects in the sample confessed though the interviewers were judged to be less than satisfactory. This suggests that the findings of Bull and Soukara (2009), Hilgendorf and Irving (1981), and Pearse and Gudjonsson (1996) may well be accurate in that some suspects enter the interview room having decided to confess and will carry out this decision irrespective of the investigator's performance. Consequently, this further indicates that interview standards cannot be measured by outputs alone when these can occur from such voluntary pre-disposed confessions. However, this eagerness of suspects to confess may have a side-effect of allowing both less able investigators, and the organizations that employ them, to delude themselves about their own effectiveness in fighting fraud. The present study indicates that it would be wrong to suggest that the continued meeting of current targets means that there is already a

high level of interviewing skills. Many other variables lead to the meeting of targets not least these many 'softer' cases where suspects confess with little prompting (Walsh & Milne, 2008). In short, many investigators are thought to be 'picking low hanging fruit' when meeting targets with the more demanding cases eluding them.

The findings from this study may be seen as early steps on a longer road of continuing research and are to be treated somewhat tentatively as effect sizes, at times, were admittedly unexceptional due to the infrequency of skilled performance across the range of assessed competencies. However, this study does demonstrate that many benefit fraud investigators are found wanting in those types of interviews that require extra skills in dealing with certain suspects (i.e., those which involve suspects who maintain their innocence, whether a true position or not, or those interviews involving the more complex benefit fraud irregularities). These findings echo Griffiths' (2008) recent research into police interviewing where he noted that those more demanding tasks (such as topic development, summarizing, and challenging suspects) were generally poorly performed. In view of how rarely skilled interview behaviours were actually present in the study's weaker interviews, it is questioned as to whether the less able investigators in this study are aware of the higher level of interviewing skills that are required, or indeed are capable of delivering them without much needed further training.

Several studies have noted that training in the PEACE model is considered to have had beneficial effects upon police interviews (Bull & Cherryman, 1996; Bull & Soukara, 2009; Clarke & Milne, 2001; Griffiths & Milne, 2006; McGurk *et al.*, 1993). However, Walsh and Milne (2008), in terms of benefit fraud interviews, found only occasional improvement. Moreover, Walsh and Bull (in press) also found far too many interviews to be below a level of acceptable competence, suggesting that re-training of many investigators was necessary.

This study reinforces the view that any organizational satisfaction based on achievement of targets is disguising the shortfalls in PEACE interviewing skills that many investigators possess when confronted with resistant suspects or when they are dealing with the more sophisticated benefit fraud interviews. Additionally, this study has shown that adherence to the PEACE model, undertaken skilfully, does have a significant influence on the more desirable interview outcomes. This means that justice is not served well if the number of 'easy' cases that are solved is the key to target achievement and tackling the benefit fraud problem. Neither should case complexity nor resistance in interviews be allowed to be pathways for benefit fraudsters to evade their criminal culpability.

References

- Arnold, L. (2007). *Observing suspect strategy used during police interviews*. Paper presented at the British Psychological Society's Division of Forensic Psychology Annual Conference, University of York, July 24–26.
- Baldwin, J. (1992). *Preparing records of taped interview*. London: HMSO.
- Baldwin, J. (1993). Police interview techniques: Establishing truth or proof? *British Journal of Criminology*, 33, 325–351.
- Bull, R., & Cherryman, J. (1996). *Helping to identify skills gaps in specialist investigative interviewing*. London: Home Office.
- Bull, R., & Milne, R. (2004). Attempts to improve police interviewing of suspects. In G. D. Lassiter (Ed.), *Interrogations, confessions and entrapment* (pp. 181–196). New York: Kluwer/Plenum.

- Bull, R., & Soukara, S. (2009). A set of studies of what really happens in police interviews. In G. D. Lassiter & C. Meissner (Eds.), *Interrogations and confessions: Current research, practice and policy*. Washington, DC: American Psychological Association.
- Clarke, C., & Milne, R. (2001). *National evaluation of the PEACE investigative interviewing course*. Police Research Award Scheme Report No. PRAS/149. London: Home Office.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Department for Work and Pensions (2005). *Reducing fraud in the benefit system: Achievement and ambitions*. London: Crown Copyright.
- Department for Work and Pensions (2008). *Housing benefit and council tax benefit claims administration quarterly performance data*, Retrieved from http://www.dwp.gov.uk/asd/asd1/hb_ctb/performance/Security_0607.xls
- Griffiths A. (2008). *An examination into the efficacy of police advanced investigative interview training?* Unpublished PhD thesis, University of Portsmouth.
- Griffiths, A., & Milne, R. (2006). Will it all end in tiers? Police interviews with suspects in Britain. In T. Williamson (Ed.), *Investigative interviewing: Rights, research and regulation* (pp. 167–189). Cullompton: Willan.
- Hilgendorf, E. L., & Irving, B. (1981). A decision-making model of confessions. In M. A. Lloyd-Bostock (Ed.), *Psychology in legal contexts: Applications and limitations* (pp. 67–84). London: MacMillan.
- Inbau, F. E., Reid, J., Buckley, J. P. & Jayne, B. (2001). *Criminal investigations and confessions*. Gaithersburg: Aspen.
- Kassin, S. M. (2005). On the psychology of confessions. Does innocence put innocents at risk? *American Psychologist*, 60, 251–258.
- Leo, R. A. (2008). *Police interrogation and American justice*. New York: Harvard University Press.
- McGurk, B. J., Carr, M. J., & McGurk D. (1993). Investigative interviewing courses for police officers: An evaluation. Police Research Group Paper No. 4, London: Home Office.
- McKeever, G. (1999). Fighting fraud: An evaluation of the government's social security fraud strategy. *Journal of Social Welfare and Family Law*, 21(4), 357–371.
- Milne, R., & Bull, R. (1999). *Investigative interviewing: Psychology and practice*. Chichester: Wiley.
- Moston, S., Stephenson, G. M., & Williamson, T. (1992). The effects of case characteristics on suspect behaviour during police questioning. *British Journal of Criminology*, 32, 23–40.
- Ord, B. & Shaw, G. (1999). *Investigative interviewing explained: The operational guide to practical interviewing skills*. Woking: The New Police Bookshop.
- Pearse, J., & Gudjonsson, G. (1996). Police interviewing techniques at two South London police stations. *Psychology, Crime and Law*, 3, 63–74.
- Shawyer, A., & Milne, R. (2009). *Investigative interviewing: Investigation, counter-fraud and behaviour*. Paper presented at the 2nd International Investigative Interviewing Research Group Conference, University of Teesside, April 14–16, 2009.
- Walsh, D. W., & Bull, R. H. (2009). *PEACE in our time? Does evaluation of practice by investigation professionals effectively occur?* Manuscript submitted for publication.
- Walsh, D. W., & Bull, R. H. (in press). Interviewing suspects of fraud: An analysis of interviewing skills. *The Journal of Psychiatry and Law*.
- Walsh, D. W., & Milne, R. (2007). Giving PEACE a chance. *Public Administration*, 85(3), 525–540.
- Walsh, D. W., & Milne, R. (2008). Keeping the P.E.A.C.E.? An analysis of the taped interview performance of benefit fraud investigators within the DWP. *Legal and Criminological Psychology*, 13, 39–57. doi:10.1348/135532506X157179

Received 11 September 2008; revised version received 3 May 2009