The Right to Silence and The Permission to Talk: Motivational Interviewing and High Value Detainees

Frances Surmon-Böhr

University of Liverpool

Laurence Alison

University of Liverpool

Paul Christiansen

University of Liverpool

Emily Alison

University of Liverpool

Author Note

Part of the data used in this study was collected as part of a project funded by the High-Value Suspect Interrogation Group FBI-HIG Contract DJF-3900001-148419 awarded to Laurence Alison at the University of Liverpool. Statements of fact, opinion, and analysis in the study are those of the authors and do not reflect the official policy or position of the HIG or the U.S. Government. The author(s) would like to thank the regional Counter-Terrorism Units and the support from the National Counter Terrorism Police HQ.

Correspondence concerning this article should be addressed to Frances Surmon-Böhr, Centre for Critical & Major Incident Psychology, University of Liverpool, Eleanor Rathbone Building, Bedford Street South, L69 7ZA. Email: F.Surmon-Bohr@liverpool.ac.uk.

Abstract

Motivational Interviewing (MI) is an evidence-based intervention that has proved effective across diverse clinical contexts with clients ambivalent about and resistant to behavioral change. This paper argues that the principles of MI can be successfully applied to law enforcement (LE) interviews with High-Value Detainees (HVDs; i.e. terrorist suspects). Although the forms of ambivalence and resistance may differ from clinical contexts, HVDs must make the decision whether to talk or not when they are interviewed. We argue there is likely ambivalence regarding this. We theorized that four MI-consistent skills may be useful for LE interviewers: reflective listening, summaries, rolling with resistance and developing discrepancies. Using the Observing Rapport Based Techniques (ORBIT) coding manual (Alison, Alison, Elntib & Noone, 2012), 804 tapes of law enforcement interviews with 75 terrorism suspects in the U.K were analyzed. Multi-level structural equation modelling revealed that Motivational Interviewing Skills encouraged detainee engagement and subsequent information gain. It also revealed that any approach antithetical to MI had a profoundly negative impact on detainee engagement and subsequent information gain - potentially through creating reactance (a form of resistance based on motivations to regain a freedom when it is threatened). Overall, this research provides unique evidence for the use of specific skills and approaches that can increase or decrease HVD engagement and information provided.

Keywords: motivational interviewing, high-value detainees, interrogation, terrorism, rapport

This paper provides empirical support for using a humane, respectful and compassionate approach to interrogating High-Value Detainees (i.e. terrorist suspects) to encourage co-operation and disclosure of information. These findings have potential to improve methods of national security whilst promoting fair treatment of detainees.

Motivational Interviewing (MI)- an evidence-based clinical intervention originally developed for treating substance misuse is described as: (i) person-centred, using the client's own knowledge and expertise about themselves (Tudor, 2008) and (ii) goal-directive insofar as therapists intentionally target a client's ambivalence about behavioral change (Miller & Rollnick, 2013). In its original context, ambivalence refers to simultaneous motivations drawing a client towards or away from substance misuse (Miller & Rollnick, 2013). Therapists practising MI provide a directive but non-judgemental environment for clients to articulate their thoughts, feelings, and beliefs surrounding the contemplation of behavioral change. Client insight moments are never forced by therapists, since attempts to push in favour of change can create client reactance - a form of resistance in which a person is motivated to regain a freedom after it has been either lost or threatened (Brehm, 1966). Consequently, berating, rational arguments and even gentle encouragement can reinforce a client's defensive articulation of motivations to stick with the misuse pattern where, previously, they were contemplating change (Miller & Rollnick, 2013). Conversely, acknowledging a person's freedom of choice (even in the direction of continuing substance misuse) typically diminishes defensiveness and can facilitate change (Miller & Rollnick, 2013). Although it originated in the addiction domain, MI has been utilized in other areas of behavioral change that encounter resistance and ambivalence (Westra & Aviram, 2013). The efficacy of MI for targeting behavioral change has been demonstrated across many diverse contexts, with over 600 clinical trials and numerous meta-analyses and systematic reviews published (e.g., DiClemente, Corno, Graydon, Wiprovnick & Knoblach, 2017; Lawrence, Fullbrook, Somerset & Schulz, 2017).

This paper examines whether there is any support for the use of MI principles within law enforcement (LE) interviews with High-Value Detainees (HVDs i.e. terrorist suspects). The paper will argue that MI's goal-directive, non-judgmental, freedom of choice-based

underpinnings are consistent with the ethos of most European (and to some extent US based) LE interviews / interrogations of HVDs. Enshrined in LE interviews in the U.K. and many European countries are the right to silence and a non-coercive, neutral presentation of evidence in the pursuit of the truth (in the U.S., detainees have the right to silence but interrogators often adopt an accusatorial, confession-based approach where the perception of evidence against a detainee can be manipulated to encourage confessions, Miller, Redlich & Kelly, 2018). Though the psychological forms of ambivalence and reactance may differ ('Do I give up drugs?' vs. 'Do I avail myself of my right to silence?'), an MI-consistent approach is both psychologically congruent and has both the HVD's and wider society's, legal and ethical rights front and center (Alison, Alison, Noone, Elntib & Christiansen, 2013).

In a study examining why offenders choose to confess to or deny accusations against them, Kebbell, Hurren and Mazerolle (2006) reported that around half of the offenders in their sample claimed they were undecided about whether or not they would confess before they were interviewed. This challenges the often-assumed legal position that not talking (and certainly not confessing) is always the desirable end state for a legal client. Furthermore, it suggests that an interviewer's approach and behavior during the interview could significantly influence a detainee's decision to talk (or not). Indeed, according to the principles of reactance, a detainee willing to talk could decide to stop talking if overtly encouraged or pressurized to talk. Conversely, adherence to the principles of MI, should create an atmosphere conducive to communication, where the detainee can contemplate whether they intend to talk (or not) (Alison et al., 2013).

The essence of MI lies in its macro-level approach to reactance rather than tactical 'tricks' to be 'deployed'. Thus, what matters is the 'spirit' or atmosphere created. This includes the creation of an accepting, empathic approach, underpinned by a partnership between therapist and client that honours client autonomy and is directed at evoking clients'

own motivations for change (Miller & Rollnick, 2013). Four key skills, characteristic of an MI style, may be particularly useful for engaging HVDs in LE interviews. These are: reflective listening, summarizing, 'rolling with' resistance and developing discrepancies (Moyers & Rollnick, 2002).

Reflective listening allows therapists to express empathy by conveying an understanding of clients' experience and ambivalence about change (Miller & Rollnick, 2013). Therapists identify the underlying meaning and feelings behind what a client has said and then present this to the client in order to check that their understanding of the client's value system is correct. A therapist may repeat back a word or phrase verbatim, or use complex reflections, such as summarizing, to add meaning or emphasis to what the client has said and/or to direct the conversation (Miller, Moyers, Ernst & Amrhein, 2003). Although it most often associated with counseling, reflective listening has been found to work effectively in a range of other settings, including hostage negotiations (Vecchi, Van Hasselt, & Romano, 2005; Voss & Raz, 2016).

When dealing with resistant clients, MI therapists can increase client engagement by avoiding argumentation and 'rolling with' resistance (Apodaca & Longabaugh, 2009). The aim is to explore and understand why the client is resistant rather than to challenge it (Moyers & Rollnick, 2002). Clients may be resistant to change (e.g., have little desire to change) or be interpersonally resistant to the therapist and/or treatment (Westra & Aviram, 2013). HVDs can of course also be resistant during interviews, employing a range of counter-interrogation tactics (CITs) to avoid co-operating (Alison et al., 2014). Many of these resemble signs of interpersonal resistance (though it is worth noting in a large-scale study of a variety of terrorist detainees, especially difficult and resistant behavior was not the norm – Alison et al., 2013). Given the potential for resistance among HVDs, an MI-consistent approach may offer LE interviewers a valuable way of dealing with HVD resistance and increasing engagement.

Aside from the interviewer's behavior, the most important factor influencing a detainee's decision to talk is the strength of evidence (Moston & Engelberg, 2011). Research has demonstrated that how and when evidence is presented can influence detainee co-operation (Hartwig, Granhag & Luke, 2014). A necessary part of LE interviews is to challenge detainees on discrepancies between their account and the available evidence (Soukara, Bull, Vrij, Turner & Cherryman, 2009). However, to date, the way in which these inconsistencies are presented interpersonally has not yet been explored. In MI, a key skill is to develop discrepancies between a client's current behavior, and goals or values important to them in order to help the client recognize that their behavior may be disadvantageous (Miller & Rollnick, 2002). To avoid client reactance, this is conducted in a non-judgmental, objective manner so that the client is able to reach this conclusion themselves. Within a LE context, this skill may be applicable to interviewers challenging detainees on discrepancies between their account and evidence. In line with the principles of MI, challenging detainees in a neutral, objective manner may lead to more information disclosed, whereas a judgmental challenge may engender resistance, leading to less information.

The following study examines the relationships between LE interviewers' use of four skills consistent with the principles of MI (reflective listening, summaries, rolling with resistance and developing discrepancies), five Global MI strategies relating to MI 'spirit' (acceptance, empathy, evocation, adaptation and autonomy), detainee engagement, and information yield. To understand the potentially deleterious effect of MI-inconsistent behaviors, we also examined approaches that represent interviewer behaviors antithetical to the four MI-consistent skills. We hypothesized that the use of skills consistent with MI would be associated with increased Global MI strategies, detainee engagement and information yield. Conversely, we predicted that MI-inconsistent behaviors would be associated with a decrease in these variables.

Method

ORBIT coding tool.

Alison et al. (2012) developed the ORBIT (Observing Rapport Based Interpersonal Techniques) coding framework to code video-taped police investigative interviews. The MI Skills element of ORBIT is based on the Motivational Interviewing Skill Code manual (MISC; Miller et al., 2003). Three specific components of ORBIT were examined in this study: (i) assessment of rapport-based skills and behaviors counter to these, drawn from the MI literature (Miller & Rollnick, 1991); (ii) a detainee engagement rating (DER), which assessed the detainee's level of engagement on an eight-point scale ranging from 1 'the detainee says nothing at any point during the session' to 8 'partial or full confession to the principal charge'; and (iii) interview yield assessment (Yield) , which assesses information of evidential significance. Yield comprises information relating to: (i) Capability - knowledge, skill or ability to engage in the offence; (ii) Opportunity - access or circumstances to commit the offence; (iii) Motive - possible reasons for comitting the offence; and (iv) PLAT- details about items/people/ locations relevant to the offence.

Rapport-based skills were coded on the following two measures: (i) MIDAS - Motivational Interviewing of Detainees Assessment of Skills (see table 1- adapted from Alison et al. 2013) and (ii) GMISC - Global Motivational Interviewing Scores, which assesses interviewers' use of five strategies: acceptance, empathy, adaptation, evocation and autonomy. All coding scales and protocols were taken directly from the ORBIT coding framework and manual. Further details of each of these coding scales, as well as how they were developed, can be found in Alison et al. (2013).

Table 1
Motivational Interviewing of Detainees: Assessment of Skills (MIDAS) coding framework

Skill	MI-Consistent Definition	MI-Inconsistent Techniques
Reflective Listening	**Accurate understanding:** demonstration that the interviewer has accurately heard and understood the detainee, using simple or complex reflections, without judgment.	**Assumptive questioning:** inaccurate or exaggerated interpretations of what the detainee has said; providing unsolicited advice; interrupting detainee; being dismissive, argumentative or accusatorial.
Summaries	**Balanced summary without judgment:** information is summarized using suspect's own words and then clarification or further detail is sought; summaries that include both positive and negative content.	**Judgmental summary**: focus is on the negative aspects of the account; summaries that introduce the interviewer's view rather than detainee's; summaries with a tone of sarcasm or disbelief.
Rapport & Resistance	**Rolling with resistance:** use of evocative prompts; statements that reflect positive and negative content; using three prompts when met with resistance, then shifting to an area of less resistance.	**Fighting resistance:** use of tactics that inhibit rapport such as threatening, ordering, use of sarcasm or judgment; warning detainee of consequences; misleading or forced questions.
Developing Discrepancies	**Neutral challenge:** inconsistencies presented to the detainee for explanation without providing excuses or passing judgment; use of detainee's own speech or specific details of forensic reports to ensure no misunderstanding; and inviting an explanation.	**Judgmental Challenge:** inconsistencies are presented in a confrontational, accusatory, or judgmental manner such as: demanding explanations, shaming or blaming; focus on police/victim perspective rather than detainee.
0=Absence	1=Mild	2=Moderate 3=Extreme

Note. Adapted from "Why tough tactics fail and rapport gets results: Observing rapport-based interpersonal techniques (ORBIT) to generate useful information from terrorists," by L. Alison, E. Alison, G. Noone, S. Eltnib and P. Christiansen, 2013, *Psychology, Public Policy, and Law, 19,* p. 417-418. 10.1037/a0034564. Copyright American Psychological Association.

Dataset.

Two datasets were used in this study. Dataset 1 was comprised of 563 interview tapes with 48 detainees, conducted between 2004-2010, drawn from a larger sample of coded interviews published by Alison et al. (2013; 2014). Only tapes with complete information for the variables used in the study were included. In addition, a new dataset (dataset 2) that

consisted of 241 interview tapes with 27 detainees, conducted between 2012-2017, was used. There was no missing data from this dataset. All cases were identified by agreement with the U.K.'s National Counter Terrorism Branch and were selected if the detainee had been convicted and did not have any appeals pending. Pairs of interviewers conducted each interview. All interviewers had undergone advanced interviewer training and were all assigned to counter-terrorism units across the U.K. and Ireland.

The combined data sets contained 804 audio and video recordings (mean length 40 minutes) with 75 detainees (representing 533 hours of footage), who were subsequently convicted of terrorism-related offenses. Of the 75 suspects, 50 were International (25 were Islamic State (IS) or IS-inspired), comprising 166 tapes; 24 were Al-Qaeda (AQ) or AQ-affiliated, comprising 262 tapes; one other International terrorist suspect comprising 4 tapes); 18 were paramilitary (237 tapes); and seven were right-wing terrorist suspects (135 tapes).

U.K. police interviews are usually broken into 45-minute segments based on the tapes used to record them. This provides a natural segment for coders to analyse the interaction. RBS (both Global MI scores and MI skills) were scored every 45 minutes, or at the end of the tape as the scores are intended to reflect the interaction as a whole. DER and Yield were scored at 15-minute intervals (i.e. it is scored three times in a 45-minute segment), as these variables vary more within the interview. However, mean scores across 45-minute segments were used for comparison with RBS scores.

Ethical Considerations

Due to the nature of the material being both confidential and sensitive, in addition to obtaining ethical approval from the University of Liverpool's Research Ethics committee, a strict Memorandum of Understanding (MOU) was agreed between the UK CT Senior National Co-ordinator (SNC) and the research team. Police interviews in the U.K. are the

property of the police services that conducted the interviews. Therefore, consent to use such interviews for research purposes was obtained from the National CT SNC. All researchers involved in the coding of data were vetted prior to gaining access to the material. Once vetted, the researchers were allowed access to the data which was password protected and encrypted. To ensure confidentiality, no identifiable information was recorded at any time whilst coding and coding of the material resulted in an anonymised data file. Coders followed an anonymization protocol to de-identify data by removing all elements that could be used to identify the individual(s) or their relatives, employers or household members.

Data Analysis

The data had a hierarchical structure as there were 804 interview tapes (level 1) nested within 75 detainees (level 2). As a result, multilevel structural equation modelling (MLSEM) was conducted which accounted for variance at the detainee level using STATA 14.1. Maximum likelihood estimation was used to test the hypothesized model and multiple indices of model fit were calculated to ensure that the model represented a good fit of the data. Prior to conducting the MLSEM, confirmatory factor analyses were conducted on all hypothesized latent variables (Yield, Global MI, MI-Consistent skill and MI-inconsistent behaviors). Multivariate normality was assessed using Mardia's multivariate skewness and kurtosis tests. This revealed that the variables were non-normally distributed (Mardia skewness = 31.55, χ^2 (1140, N=804) = 4244.8, p<.001; Mardia kurtosis = 377.99, χ^2 (1, N=804) = 90.30, p<.001). Consequently, Satorra-Bentler scaled χ^2 estimation with adjustment to standard errors was used to test hypothesized model fit for the latent models as it is robust to non-normality. The standardized root mean residual (SRMR) absolute fit index was also used to assess model fit, it is less affected by sample size distribution and kurtosis as it is not a simple variation of χ^2.

For this measure values under 0.08 are representative of a good model fit (Hu & Bentler, 1999).

As well as using the discrepancy function methods, two non-centrality-based indices were used to evaluate fit (Bentler, 2007). The comparative fit index (CFI), which is less sensitive to sample size than the previous measures, was used, whereby values equal to or greater than 0.95 are good fit. The root mean square error of approximation (RMSEA) was appropriate in this model due to the large df; values equal to or lower than 0.06 were used to determine a good fit (Hu & Bentler, 1999).

In describing specific relationships within the model, standard errors and confidence intervals (CI_{95}) and associated p values are reported. Unstandardized regression coefficients and their standard errors are reported.

Following this initial analysis, the hypothesized structural equation model was run as a single level and multi-level model using gsem with robust adjustment for standard errors in STATA. These models were ran using mean scores of the latent variables for yield, Global MI and MI consistent and inconsistent behaviors. This is due to computational limitations in estimating multiple latent variables and their associations with each other and the other, observed variables across multiple levels (75). To control for nesting in the data we added 'detainee' as a random intercept. Model fit indices described above cannot be computed for MLSEM, however, AIC and BIC comparative fit values were used to compare the comparative fit of the MLSEM model to the single-level model.

Inter-coder Agreement

For this study, a subset of 30 tapes from the dataset were randomly selected and each tape was coded by two experienced coders to check IRR was adequate for all variables used

in the study. IRR was calculated using Intra-class correlations (ICC) using a two-way random, consistency, single measures ICC (McGraw & Wong, 1996) on the raw ordinal scores, as it allows consistency to be correlated in an additive manner rather than on absolute agreement (Hallgren, 2012). The Kappa index was also used to check categorical coding of variables that could be coded dichotomously (i.e. existence of behavior vs. absence) along with percentage agreements, bearing in mind the high sensitivity of Kappa values to peripheral methodological issues such as prevalence of one category over the other, sample size and number of ratings in each scale (Feinstein & Cicchetti, 1990).

Agreement for ICC was categorized into (<0.40) poor, (0.40- 0.59) fair, (0.60- 0.74) good, and (>0.75) excellent (Cicchetti, 1994). IRR was assessed using a two-way random, consistency, single measures ICC (McGraw & Wong, 1996), to assess the degree that coders provided consistency in their ratings of each variable. DER could only be assessed through ICC and achieved excellent agreement (ICC= .87). Agreement for kappa was categorized into (0.00- 0.20) poor, (0.21- 0.40) fair, (0.41-0.60) moderate, (0.61 to 0.80) strong, and (>.80) near complete agreement (Landis & Koch, 1977). The results of the IRR analyses are shown in Table 2 and were deemed acceptable. Although two variables (MI-Consistent Rapport & Resistance and Summaries) achieved lower ICC values at the interval level, the categorical coding of these variables was much higher, achieving fair agreement using Kappa. In addition, rater percentage agreements on these two categories remained high. This may be an indication that it is more difficult to apply subtler scaling to these two categories and they may be interpreted as mild to moderate based on small differences in interpretation.

Table 2
Interrater Reliability Scores: Motivational Interviewing of Detainees Assessment of Skills.

Measures	ICC	Kappa	Rater percentage agreement	ICC	Kappa	Rater percentage agreement
	MI-Consistent			MI-Inconsistent		
Reflective listening	.58	.52	90%	.68	.63	83%
Rapport & Resistance	.26	.35	74%	.61	.53	83%
Summaries	.26	.35	90%	.40	.41	80%
Develop Discrepancies	.41	.44	77%	.66	.53	80%
Capability	.59	.34	65%			
Opportunity	.66	.47	74%			
Motive	.68	.64	82%			
PLAT	.77	.64	82%			
Acceptance	.68	.73	90%			
Empathy	.82	.41	80%			
Adaptation	.58	.38	83%			
Evocation	.84	.84	94%			
Autonomy	.52	.47	83%			

Results

Descriptive statistics

Descriptive statistics for rapport-based skills are shown in Table 3. Global MI is scored from 1-7 and MI skills are scored from 0-3. Mean scores for MI- consistent skills were all higher than MI-inconsistent behaviors. Descriptive statistics for yield and DER are also shown in Table 3. DER is rated from 1-8 and all yield variables were scored from 0-3. Values of skewness and kurtosis ranged between the acceptable levels of -2 and 2, thus no transformations were necessary (Lewis-Beck et al., 2003).

Table 3
Means and standard deviations and Skewness and Kurtosis and standard errors of detainee engagement (DER), yield and MI skills

	Mean (SD)	Skewness (SE)	Kurtosis (SE)
MI-Consistent			
Reflective listening+	1.60 (.93)	-.17 (.09)	-.82 (.17)
Rapport & Resistance+	1.31 (1.02)	.14 (.09)	-1.13 (.17)
Summaries+	1.17 (.97)	.25 (.09)	-1.00 (.17)
Develop Discrepancies+	1.22 (.97)	.14 (.09)	-1.08 (.17)
MI-Inconsistent			
Reflective listening-	.72 (.84)	.84 (.09)	-.32 (.17)
Rapport & Resistance-	.59 (.81)	.14 (.09)	-1.13 (.17)
Summaries-	.50 (.80)	.25 (.09)	-1.00 (.17)
Develop Discrepancies-	.60 (.87)	.14 (.09)	-1.08 (.17)
DER	3.87 (2.07)	.28 (.09)	-1.33 (.17)
Capability	.63 (1.00)	1.34 (.09)	.34 (.17)
Opportunity	.58 (.94)	1.46 (.09)	.85 (.17)
Motive	.87 (1.53)	1.53 (.09)	1.18 (.17)
PLAT	.86 (1.05)	.79 (.09)	-.84 (.17)
Acceptance	4.71 (1.61)	-.40 (.09)	-.71 (.17)
Empathy	4.28 (1.55)	-.35 (.09)	-.44 (.17)
Adaptation	4.42 (1.54)	-.27 (.09)	-.50 (.17)
Evocation	4.07 (1.71)	-.24 (.09)	-.89 (.17)
Autonomy	4.63 (1.53)	-.52 (.09)	-.05 (.17)

Internal Reliability of MI Scales

Principal factor analysis (principal axis factoring) with direct oblimin rotation revealed a clear 3 factor solution for the MI variables. Sampling adequacy was good (KMO= .90) and Bartlett's test of sphericity showed sufficient correlations between items (χ^2 (78)=6438.56, p<.001). The Eigenvalues of the 3 factors were 5.96, 1.99 and 1.31 and accounted for 71% of the variance. Factor 1 comprised of the 5 global MI variables, factor 2 comprised of the four MI-consistent (MIC) skills and factor 3 comprised of the four MI-inconsistent (MIIC) behaviors. Factor loadings can be seen in table 4 (factor loadings below 0.40 were suppressed). Internal reliability was then assessed using Cronbach's alpha. All three scales showed good to excellent internal reliability as can be seen in Table 4. For the

Global MI and MIIC scales, removal of any variable would weaken the scale's internal consistency, however, if developing discrepancies was removed from the MIC scale, internal consistency of the scale would increase (α= .76).

Table 4
Factor loadings and Cronbach's alpha coefficients for Global MI skill, MI-Consistent (MIC) skill and MI-Inconsistent (MIIC) behavior.

Model	Global MI	MIC	MIIC
Acceptance	.80		
Empathy	.88		
Adaptation	.85		
Evocation	.88		
Autonomy	.76		
Reflective Listening +		.62	
Rapport & Resistance +		.55	
Summaries +		.70	
Develop Discrepancies+		.53	
Reflective Listening -			.73
Rapport & Resistance -			.87
Summaries -			.77
Develop Discrepancies-			.84
α	.93	.73	.88

Data Modelling

Measurement models. Confirmatory factor analysis was used to test the construct validity of the latent variables created for Yield, Global MI, MI-Consistent (MIC) skill and MI-Inconsistent (MIIC) techniques. Some co-variances between errors were added to the models based on modification indices and theoretical justification. The overall fit of the Yield model (which included covariance between capability and motive errors) was good on all measures (χ^2 (1, N=804) = 0.57, p=.45, SRMR= .004, CFI= 1.00, RMSEA = .01 [CL90: 0.00 to 0.10]) and all factor loadings were significant (p <.001). Likewise, the overall fit of Global MI model was good on all measures (χ^2 (3, N=804) = 3.46, p=.33, SRMR= .005, CFI= 1.00, RMSEA = .01 [CL90: 0.00 to 0.07]) and all loadings were significant (p <.001). Covariances between acceptance and evocation and acceptance and autonomy errors were

added to the model. The fit of the model for the MI-inconsistent (MIIC) Techniques (which included covariance between reflective listening and rapport and resistance errors) was good on all measures ($\chi 2$ (1, N=804) = 1.02, p=.32, SRMR= .004, CFI= 1.00, RMSEA = .03 [CL90: 0.00 to 0.10]) and all factor loadings were significant (p <.001). Lastly, the overall fit for the MI-consistent (MIC) skill model (which included covariance between develop discrepancies and summaries errors) was good on all measures (χ^2 (1, N=804) = 0.10, p=.75, SRMR= .002, CFI= 1.00, RMSEA = .01 [CL90: 0.00 to 0.07]) and all factor loadings were significant (p <.001). However, as developing discrepancies had a relatively low factor loading compared to the other three variables in the model, (B = .37 SE = .05, CL95: 0.27 to 0.47, p <.001) and its removal from the MIC scale would increase the scale's internal consistency, it was removed from the MIC latent variable.

Following CFA of each latent variable, AIC and BIC comparative fit measures were used to assess whether a single-level or multi-level model which accounted for 'detainee' was a better fit. As seen in Table 5, the multi-level models were a superior fit for all latent variable models (lower AIC and BIC values indicate a better fit).

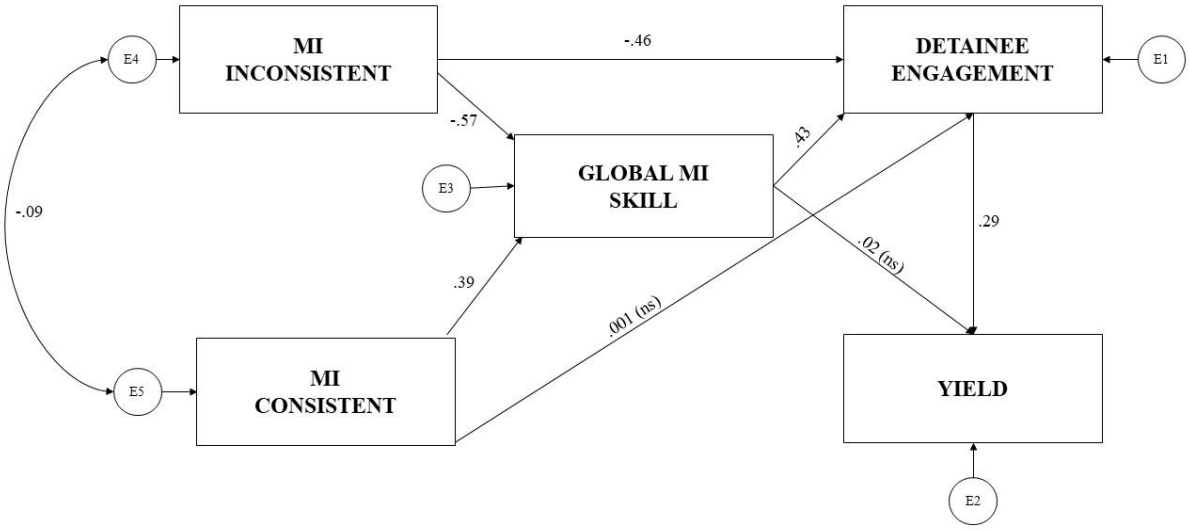
Table 5
Comparative fit values (AIC and BIC) for single and multi-level models for each latent variable.

	Single level Models		*Multi-level Models*	
	AIC	*BIC*	AIC	BIC
YIELD	7139.92	7200.88	6633.17	6712.89
MIIC	6276.41	6337.38	6123.07	6202.79
MIC	8162.26	8223.23	7857.09	7936.82
GMISC	11871.17	11950.89	11209.13	11312.3

Multilevel structural model. The dependent variable for the hypothesized structural model was interview yield. The hypothesized structural model investigated the direct and indirect effects of Global MI Skill, MI-consistent skills and MI inconsistent techniques and

level of detainee engagement (DER) on interview yield. The multi-level model that accounted for variance at the detainee level was found to be a better fit (AIC=8833.68, BIC=8941.54) for the data than the single-level model (AIC=10379.17, BIC=10463.58). In the multi-level model, 39% of the explained variance in yield was attributed to differences between detainees (level 2) and 61% attributed to differences at the interview tape level within the same detainee (level 1). For detainee engagement, 60% of variance was attributed to between detainee differences and 40% attributed to the interview tape level. For Global MI scores, 59% of variance was attributed to differences between detainees and 41% attributed to interview tape level. For MIC scores, 42% of variance was attributed to differences between detainees and 58% to differences at the tape level. For MIIC scores, 29% of variance was attributed to differences between detainees and 71% attributed to the interview tape level. Associations between all variables in the multilevel SEM are reported below (also see Figure 1).

Figure1. Hypothesized structural model for how interrelating MI skills, Global MI skills and MI-inconsistent techniques interact with one another and with detainee engagement and yield. Unstandardized parameter estimates presented are statistically significant at $p<.05$ unless otherwise indicated (ns); single level model is presented for ease of understanding.



Associations between DER and interview yield There was a significant positive association between DER and interview yield (*B*= .29 SE = .02, CL95: 0.26 to 0.32, *p* <.001).

Associations between Global MI skill, interview yield and DER. There was a strong, positive association between Global MI Skill and DER (*B* = .43 SE = .13, CL95: 0.17 to 0.68, *p* < .001. The direct association between Global MI skill and interview yield was not significant (*B*= .02 SE = .02, CL95: -0.01 to 0.05, *p* = .19). However, Global MI skill had a significant indirect effect on yield, mediated by increased DER (*B*=.12 SE = .04, CL95: 0.5 to 0.20, *p* <.001).

Associations between MI-inconsistent (MIIC) Techniques and Global MI Skill, DER and interview yield. As hypothesized, there was a significant negative association between MIIC skills and Global MI skill (*B*= -.57 SE = .12, CL95: -0.79 to -0.34, *p* <.001). There was also a significant negative association between MIIC skills and DER (*B*= -.46 SE = .22, CL95:-0.89 to -0.23, *p* = .039), as well as a significant negative indirect effect on DER, mediated by Global MI Skill (*B*= -.24 SE = .08, CL95: -0.39 to -0.09, *p* <.001). Additionally,

there was a negative indirect effect of MIIC on interview yield, mediated by reduced DER (*B*= -.13 SE = .07, CL95: -0.26 to -0.01, *p* = .042).

Associations between MI-consistent (MIC) skills and Global MI Skill, DER and interview yield. As hypothesized, there was a significant positive association between MIC skills and Global MI skill (*B*= .39 SE = .07, CL95: 0.24 to 0.54, *p* < .001) although not between MIC skills and DER (*B*= .001 SE = .22, CL95: -0.35 to 0.37, *p* = .99). There was however, a significant positive indirect effect on DER, mediated by Global MI Skill (*B*= .16 SE = .07, CL95: 0.03 to 0.30, *p* = .017). However, no significant indirect effect of MIC skills on interview yield was found (*B*=.001 SE = .06, CL95: -0.13 to 0.13, *p* = .99).

It is worth noting the results of the multi-level model using mean scores of the latent variables produced that same pattern of results as a single-level latent variable model.

Discussion

Careful observational coding of law enforcement interrogations with 75 convicted terrorism suspects (the largest international corpus of field data of its kind) reveals that detainee engagement (and disengagement) is impacted by interviewer behavior. In legal terms, this broadly relates to the concept of detainees exercising their right to silence, whilst also respecting, where they wish to, their right to talk. Whilst results reveal that Motivational Interviewing skills encourage engagement, even more pronounced was the finding that any approach antithetical to MI (accusation, assumption and confrontation) had a profoundly negative impact on detainee engagement. By accounting for individual differences between detainees, the results reveal that these effects are found regardless of how co-operative (or not) a detainee may be generally. This suggests that a detainee previously willing to talk is likely to be discouraged from doing so by an interviewer working too hard to convince them

to talk (e.g., through accusatory statements, pre-judging their answers, confronting them too forcefully with evidence or rational cornering). These approaches appear to cause reactance and thus encourage a detainee to avail him/herself of the right to silence, whereas previously, they may have been considering their right to speak.

In previous studies of MI, a client's contemplation of behavioral change (indicated by articulating their reason for change) is the main predictor of a 'successful outcome' (i.e. reducing substance misuse) (Apodaca & Longabaugh, 2009). In this sample of HVDs, contemplating engagement ('Should I talk or not talk to this interviewer?') determined whether they would reveal information. The results suggest that many of the HVDs do contemplate engagement rather than the often-held legal assumption that it is always in their best interest to say nothing. Additionally, the results showed the use of MI skills and commitment to creating an accepting, empathic atmosphere was associated with increased engagement. Interviewing officers and attorneys /solicitors need to consistently bear in mind that it is not their job to convince the individual either way what is in their best interest. As enshrined in UK and US Law, it is a detainee's *choice* whether to avail themselves of their right to silence and their right to speak. Even gentle persuasion on the part of the interviewing officer to speak could push an individual into a decision to not speak. Equally, expert legal representatives recognise that their advice is just that – advice – and not an instruction to not speak.

Assumptive questioning, judgemental summaries, and accusatory challenges (behaviors inconsistent with MI) caused detainees to disengage and stop talking. This supports MI research that shows how therapist MI-inconsistent behaviors are associated with higher levels of resistance, lower client engagement and worse outcomes (Apodaca & Longabaugh, 2009). It also supports investigative interviewing research that has found accusatory, confrontational approaches (e.g., disallowing denials and asserting authority) are

associated with increased resistance and decreased co-operation (Kelly, Miller & Redlich, 2016). The current study did not examine whether particular interrogator and/or detainee characteristics predicted the use of MI-inconsistent behaviors but this may be an interesting area to explore.

In therapeutic settings, the use of MI-inconsistent behaviors can lead to particularly negative outcomes (i.e., increased substance misuse) with angry and/or highly reactive patients (Karno & Longabaugh, 2004; 2005). Reactance theory posits that a person's reaction to a loss of freedom will be greater the more important it is perceived to be and when several freedoms are threatened. As all detainees in the sample had been arrested, were being held in police custody at the point of interview, and were having to be interviewed, it is likely that many of them will have been experiencing reactance in response to their loss of freedoms. Consequently, MI-inconsistent approaches directed a pressurizing or persuading detainees to talk, may have increased detainee reactance further, and thus reinforced their motivation to resist (their way of regaining some freedom of choice). For ambivalent detainees, such approaches may have removed any doubt they had about whether to co-operate or not, strengthening their resolve not to speak. Since controlling interviewer behavior can arouse reactance, officers must work hard to avoid such techniques and abstain from language that builds the illusion of limiting choice (Place & Meloy, 2018). This may prove a difficult task for interrogators if their well-established interviewing style incorporates MI-Inconsistent behaviors such as controlling or accusatory language. Research in the therapeutic literature has shown whilst therapists from other counseling backgrounds can successfully learn MI, it is considerably harder for them to stop using MI-inconsistent behaviors (e.g. directing, persuading, confronting) (Miller & Mount, 2001). As such, it is suggested that interrogator training should first and foremost focus on identifying and removing MI-inconsistent behaviors from interrogators repertoires.

In contrast, the use of MI–consistent skills were associated with increased detainee engagement. Specifically, reflective listening, balanced summaries and rolling with resistance, contributed to creating a non-judgmental, supportive atmosphere, in which interviewers genuinely sought to understand detainees' perspectives and respected their right to choose to talk or not. It was in the presence of this atmosphere of communication that detainees chose to engage with the interviewers and provide information. This supports research which suggests that MI's success can be attributed to its macro-level approach - known as MI 'spirit'- based on collaboration between therapist and client, honouring client autonomy and evoking clients' own motivations for change (Miller & Rollnick, 2013; Copeland, McNamara, Kelson & Simpson, 2015). It also supports the view that evoking 'MI Spirit' is aided by employing the use of specific skills (Resnicow & McMaster, 2012). In mastering these, interviewers remained neutral and open to hearing detainees' versions of events, rather than pre-judging them and assuming a pre formed version of events. This created an atmosphere conducive to communication and encouraged engagement without external pressure. In doing so, interviewers adhered to their central goal of collecting information neutrally and as a search for truth.

It should be noted that the interrater reliability scores for two of the MI-consistent scales (summaries and rolling with resistance) achieved lower scores at the interval level which could have influenced the results. However, categorical coding (i.e. presence/absence) of these variables achieved fair agreement and the rater percentage agreements were high (both above 70%). Future research should be mindful that it may be more difficult to apply subtler scaling to these variables.

Interestingly, interviewer use of developing discrepancies appeared to operate differently to the use of other MI-consistent skills. In MI, developing discrepancies between the client's values and current behavior is vital in encouraging behavioral change (Westra &

Aviram, 2013). However, within a LE context and the current sample of individuals convicted for terrorism offences (i.e., eventually deemed guilty in a court of law), discrepancies mostly existed between the detainee's account and evidence. Hence, even where interviewers developed discrepancies in a non-judgmental, neutral manner, these individuals were more likely to choose to disengage. The same neutral challenges may have had a very different effect on innocent detainees since they may have perceived the exact same option to tell the truth and explain the inconsistency as an opportunity to engage and clarify rather than, as here, shut down. In subsequent research we intend to test this hypothesis.

Multilevel analysis allowed us to understand how much variance in each variable was accounted for by individual differences between detainees (level 2) and differences between interviews with the same detainee (level1). The results revealed that for detainee engagement, yield, Global MI, and MI-consistent scores, variance was fairly equally spread between both levels. This suggests that while there are individual differences in detainees that contribute substantial variance to these measures, what happens across interviews is equally as important. Interestingly, variance in interviewer MI-inconsistent behavior was predominantly attributed to differences between interviews rather than differences between detainees. This suggests that interviewers exhibit these behaviors across interviews with a range of detainees. Future research should explore what factors within an interview are associated with interviewers displaying MI-inconsistent behaviors, in order to stop them doing so. Additionally, future research could explore other level 2 predictors that may explain some of the variance between detainees (e.g. the terrorist organisation/group a detainee is affiliated with and thus, how much (if any) counter-interrogation training they have received; Alison et al., 2014).

This study supports the use of neutral, objective, and compassionate approaches to interviewing detainees. However, there are many contexts around the world where detainees (especially HVDs) are still held and treated inhumanely. By providing empirical support for an ethical, objective and compassionate approach we hope to encourage interviewers around the world to move away from coercive practices. Regarding the use of MI in this context, we must re-emphasize that the essence of MI lies in its macro level approach (i.e. 'spirit') based on an egalitarian relationship between interviewer and client/detainee and honouring client/detainee autonomy (Miller & Rollnick, 2013). It is not just a set of tactics or tricks to be used on someone (Arkowitz & Miller, 2008). MI strategies require a context in which detainees' rights and autonomy are respected. Thus, we suggest that using MI techniques alongside coercion, persuasion or manipulation (i.e. inconsistent with the ethos of MI) is ethically dubious, and is not in the spirit of MI (i.e. as soon as such influence 'tricks' are used alongside it, it ceases to be MI).

Although this study was based on a sample of HVDs, we predict that similar results would be found with other suspect populations. Crucially, our results reveal the detrimental use of behaviors counter to the ethos of MI (e.g., pressurising, confronting and judging) and that these increase resistance and reduce engagement. Hence, we echo Alison et al.'s (2013) assertion that, although it may not always be possible to engage a highly resistant detainee, using accusatory, pressurising techniques always makes things worse. In highlighting this, we hope to encourage law enforcement interviewers to first and foremost eliminate techniques that disengage detainees (i.e. remove behaviors antithetical to MI) and thereafter seek to adopt a set of behaviors that are more positively inclined to generate and display objectivity, compassion and empathy.

References

Alison, L., Alison, E., Elntib, S., & Noone, G. (2012). Educing information through interpersonally competent interviewing: An assessment tool for coding interviews with high value suspects. Coding manual. University of Liverpool Internal Document.

Alison, L. J., Alison, E., Noone, G., Elntib, S., & Christiansen, P. (2013). Why tough tactics fail and rapport gets results: Observing Rapport-Based Interpersonal Techniques (ORBIT) to generate useful information from terrorists. *Psychology, Public Policy, and Law*, *19*(4), 411-431. doi:10.1037/a0034564

Alison, L., Alison, E., Noone, G., Elntib, S., Waring, S., & Christiansen, P. (2014). Whatever you say, say nothing: Individual differences in counter interrogation tactics amongst a field sample of right wing, AQ inspired and paramilitary terrorists. *Personality and Individual Differences*, *68*, 170-175. https://doi.org/10.1016/j.paid.2014.04.031

Apodaca, T. R., & Longabaugh, R. (2009). Mechanisms of change in motivational interviewing: a review and preliminary evaluation of the evidence. *Addiction*, *104*(5), 705-715. https://doi.org/10.1111/j.1360-0443.2009.02527.x

Bentler, P. M. (2007). On tests and indices for evaluating structural models. *Personality and Individual differences*, *42*(5), 825-829. https://doi.org/10.1016/j.paid.2006.09.024

Brehm, J. W. (1966). *A theory of psychological reactance*. Oxford, England: Academic Press

Copeland, L., McNamara, R., Kelson, M., & Simpson, S. (2015). Mechanisms of change within motivational interviewing in relation to health behaviors outcomes: a systematic review. *Patient education and counseling*, *98*(4), 401-411. https://doi.org/10.1016/j.pec.2014.11.022

DiClemente, C. C., Corno, C. M., Graydon, M. M., Wiprovnick, A. E., & Knoblach, D. J. (2017). Motivational interviewing, enhancement, and brief interventions over the last

decade: A review of reviews of efficacy and effectiveness. *Psychology of Addictive Behaviors*, *31*(8), 862-87. doi:10.1037/adb0000318

Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low kappa: I. The problems of two paradoxes. *Journal of clinical epidemiology*, *43*(6), 543-549. https://doi.org/10.1016/0895-4356(90)90158-L

Hartwig, M., Granhag, P. A., & Luke, T. (2014). Strategic use of evidence during investigative interviews: The state of the science. In: Raskin, D.C., Honts, C.R., Kircher, J.C. (Eds.), *Credibility Assessment: Scientific Research and Applications* (pp. 1-36). Academic Press. https://doi.org/10.1016/B978-0-12-394433-7.00001-4

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology*, *8*(1), 23. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3402032/

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modelling: a multidisciplinary journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Karno, M. P., & Longabaugh, R. (2004). What do we know? Process analysis and the search for a better understanding of Project MATCH's anger-by-treatment matching effect. *Journal of studies on alcohol*, *65*(4), 501-512. https://doi.org/10.15288/jsa.2004.65.501

Karno, M. P., & Longabaugh, R. (2005). An examination of how therapist directiveness interacts with patient anger and reactance to predict alcohol use. *Journal of Studies on Alcohol*, *66*(6), 825-832. https://doi.org/10.15288/jsa.2005.66.825

Kebbell, M., Hurren, E. J. and Mazerolle, P. (2006). An investigation into the effective and ethical interviewing of suspected sex offenders. *Trends & Issues in Crime & Criminal Justice, 327,* 1-6. Retrieved from https://aic.gov.au/publications/tandi/tandi327

Kelly, C. E., Miller, J. C., & Redlich, A. D. (2016). The dynamic nature of interrogation. *Law and human behavior*, *40*(3), 295-309. http://dx.doi.org/10.1037/lhb0000172

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 33(1), 159-174. Retrieved from https://www.jstor.org/stable/2529310

Lawrence, P., Fulbrook, P., Somerset, S., & Schulz, P. (2017). Motivational interviewing to enhance treatment attendance in mental health settings: a systematic review and meta-analysis. *Journal of Psychiatric Mental Health Nursing, 24,* 699–718. doi:10.1111/jpm.1242

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological methods*, *1*(1), 30-46.

Miller, J. C., Redlich, A. D., & Kelly, C. E. (2018). Accusatorial and information-gathering interview and interrogation methods: a multi-country comparison. *Psychology, Crime & Law*, *24*(9), 935-956. https://doi.org/10.1080/1068316X.2018.1467909

Miller, W. R., & Rollnick, S. (1991). *Motivational interviewing: Preparing people to change addictive behavior.* New York, NY: Guilford Press.

Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing: Preparing people for change (2nd ed.).* New York, NY: Guilford Press.

Miller, W. R., & Rollnick, S. (2013). *Motivational interviewing: Helping people change (3rd ed.)*. New York, NY: Guilford Press.

Miller, W. R., & Mount, K. A. (2001). A small study of training in motivational interviewing: Does one workshop change clinician and client behavior?. *Behavioural and Cognitive Psychotherapy*, *29*(4), 457-471. https://doi.org/10.1017/S1352465801004064

Miller, W. R., Moyers, T. B., Ernst, D., & Amrhein, P. (2003). Manual for the Motivational Interviewing Skills Code (MISC) v. 2.0. Retrieved from http://casaa.unm.edu/codinginst.html.

Moston, S., & Engelberg, T. (2011). The effects of evidence on the outcome of interviews with criminal suspects. *Police Practice and Research*, *12*(6), 518-526. https://doi.org/10.1080/15614263.2011.563963

Moyers, T. B., & Rollnick, S. (2002). A motivational interviewing perspective on resistance in psychotherapy. *Journal of clinical psychology*, *58*(2), 185-193.

Place, C. J., & Meloy, J. R. (2018). Overcoming Resistance in Clinical and Forensic Interviews. *International Journal of Forensic Mental Health*, 17 (4), 362-376. https://doi.org/10.1080/14999013.2018.1485189

Resnicow, K., & McMaster, F. (2012). Motivational Interviewing: moving from why to how with autonomy support. *International Journal of Behavioral Nutrition and Physical Activity*, *9*(1), 19. https://doi.org/10.1186/1479-5868-9-19

Soukara, S., Bull, R., Vrij, A., Turner, M., & Cherryman, J. (2009). What really happens in police interviews of suspects? Tactics and confessions. *Psychology, Crime & Law*, *15*(6), 493-506. https://doi.org/10.1080/10683160802201827

Tudor, K. (Ed.). (2008). *Brief person-centred therapies*. Beverly Hills, CA: Sage Pubilcations.

Vecchi, G. M., Van Hasselt, V. B., & Romano, S. J. (2005). Crisis (hostage) negotiation: Current strategies and issues in high-risk conflict resolution. *Aggression and Violent Behavior*, *10*(5), 533-551. doi:10.1016/j.avb.2004.10.001

Voss, C., & Raz, T. (2016). *Never split the difference: Negotiating as if your life depended on it*. New York: Harper Business.

Westra, H. A., & Aviram, A. (2013). Core skills in motivational interviewing. *Psychotherapy*, *50*(3), 273- 278. doi:10.1037/a0032409