Assessing Threats of Violence: Professional Skill or Common Sense?

Renate Geurts*

University of Gothenburg, Sweden

University of Portsmouth, England

Pär Anders Granhag

University of Gothenburg, Sweden

Norwegian Police University College, Norway

Karl Ask

University of Gothenburg, Sweden

Aldert Vrij

University of Portsmouth, England

*Corresponding author: Renate Geurts

Department of Psychology, University of Gothenburg

P.O. Box 500, 405 30 Gothenburg, Sweden

E-mail: renate.geurts@psy.gu.se

Phone: +46317864270

Word count: 4543 (excl. abstract, references, tables, figure, appendix)

Abstract

When faced with threats of violence, it is of great importance to assess the risk for actual harm to occur. Over the last decades, this task has developed into a domain of its own and professionals have specialized in *threat assessment*. However, it is yet unknown whether professional experience affects the quality of threat assessments. The present study examined how threat assessment professionals ($N = 44$), university students ($N = 44$), and laypersons ($N = 45$) assessed the risk for violence in three fictitious cases. The assessments (i.e., assigning risk values to different pieces of information) were found to be strikingly similar across the three groups. Yet, professionals agreed more with one another on their assessments, and professionals identified more relevant (empirically supported) threat cues when given the opportunity to request additional information. These results suggest that threat assessment professionals know better than non-professionals what information to look for, and hence, they may contribute most in the process of gathering information to clarify the threat. Such knowledge can help to optimize the use of expertise, which may improve the quality of threat assessments. The current findings can be of value to those who consult threat assessment professionals, as well as to the professionals themselves.

Key words: threat assessment, threat management, violence risk assessment, expert performance, decision-making

Assessing Threats of Violence: Professional Skill or Common Sense?

The 9/11 attacks and subsequent mass killings in Western societies brought about a growing awareness of risks and security, and the field of threat assessment has developed rapidly since. Threat assessment is the process of information gathering to understand and evaluate a threat of violence (Meloy, Hoffmann, Roshdi, Glaz-Ocik, & Guldimann, 2014a). Developments in this field are evidenced by a fair amount of research examining cases and classifying warning behaviours (e.g., Dietz & Martell, 2010; Hoffmann, Meloy, Guldimann, & Ermer, 2011). Such research has resulted in jargon (e.g., "intended violence", Calhoun & Weston, 2003), in guidelines on how to conduct threat assessment investigations (e.g., Fein & Vossekuil, 1998), and in threat assessment tools (e.g., White & Meloy, 2010; James, MacKenzie, & Farnham, 2014). Importantly, the growing body of knowledge has allowed practitioners to specialize, and associations have been established for so-called *threat assessment professionals* (e.g., the Association of European Threat Assessment Professional, AETAP).

Threat assessment professionals must mitigate the risk that individuals who pose threats of violence will indeed commit violence. One of their key tasks is to identify conditions and behaviours that precede planned acts of violence (Meloy et al., 2014a; Skeem & Monahan, 2011). The present study investigates how professionals and non-professionals assess (fictitious) threats of violence. The aim of this study is to identify threat assessment tasks for which professional experience is more or less critical. Such an understanding is important because allocating expertise to tasks that could be done by others would be inefficient, and relying on laypersons' knowledge when expertise is needed can be fatal.

A study on threat assessment professionals demands a brief outline of how this profession is practiced, and what professionals are expected to contribute. The methodology used in threat assessment largely resembles violence risk assessment, in which risk factors are used to estimate the likelihood of physical violence occurring in a population (Skeem & Monahan, 2011). Risk assessment and threat assessment are professional fields that share the same goal: to assess and manage risks for violence. However, threat assessors and risk assessors *typically* operate in different contexts (law enforcement vs. health care, resp.), which results in different responsibilities (operational vs. legal decision-making). For

an overview of differences and similarities in threat and risk assessment, see Meloy and colleagues (2014a). Most relevant to the present study is the fact that both fields make use of a *Structured Professional Judgment* model (SPJ; Guy, Packer, & Warnken, 2012). SPJ relies on the discretion of the professional, while providing structure to their judgment via empirically informed guidelines (Douglas, Ogloff, & Hart, 2003). SPJ is therefore considered a middle ground between two decision-making models that characterize risk assessment practices: the unstructured *clinical* model that relies on the subjective assessment of the expert and the highly structured *actuarial* model that relies on data based algorithms (Meehl, 1954). Risk assessment instruments that are developed under an SPJ approach hold a list of empirically supported risk and protective factors, these factors are operationalized, and guidelines are provided on how to score them. However, the scores do not add up to a fixed outcome, they are meant to guide the professional in their clinical judgement of risk (Douglas et al., 2003). The responsibility of professionals applying SPJ is to recognize potential risk or protective factors in an individual case, and to assess the degree to which the (combination of) factors increase or decrease the overall risk for violence (Douglas et al., 2003; Skeem & Monahan, 2011). In other words, threat assessment professionals are supposed to both *select* and *weight* information cues to indicate the risk for violence posed by the individual of concern. But how good are they in doing so?

The quality of expert performance can be defined by the ability to perform consistently with oneself and with others (i.e., reliability), and by the ability to base decisions on relevant information (i.e., biasability; Dror, 2016). Differently put, experts are presumed to identify, cluster, and weight information in the same way, and they should make their judgments relatively free of biases (Einhorn, 1974). As it remains challenging to certify expertise (Ericsson, 2008; Shanteau, Weiss, Thomas, & Pounds, 2002), the focus of this paper is on the more objective measure of *professional experience*. Specifically, we examined how professional experience affects the assessment of threats of violence.

Research suggests that superior professional performance results mainly from domain-specific knowledge (Patel & Groen, 1986; Schmidt & Rikers, 2007). Such knowledge allows professionals to quickly recognize relevant patterns and to generate scenarios, which contributes to an efficient

information search (Elstein & Schwarz, 2002). Expert physicians, for instance, were found to be significantly faster than novices producing diagnostic hypotheses (Joseph & Patel, 1990). More recently, in a study with British police officers, experienced detectives generated more investigative hypotheses and actions when presented with a missing-person case, compared to their less experienced colleagues (Fahsing & Ask, 2016).

The literature on risk perception further suggests that experts have a more realistic sense of the likelihood for danger to occur, whereas the general public tends to overestimate infrequent but catastrophic events (e.g., a terrorist attack; Weber, 2001). However, this difference in knowledge between experts and laymen does not necessarily influence how they weight risk cues when faced with them. Simply illustrated, an expert might assess the overall risk for a terror attack to occur as being low, but when confronted with concrete signs of an attack (e.g., someone enters the airport wearing a bomb vest) both experts and novices will judge this information as similarly dangerous. It has been theorized that risk assessment is largely affective in nature—even when approached rationally—and that affective risk assessment is equally well (or poorly) developed in experts and novices, as responding to threats is deeply rooted in human evolution (Slovic & Weber, 2002; Slovic, Finucane, Peters, & MacGregor, 2004). Studies that examined people's capability to detect suspicious behaviour showed that professionals were no better than novices in detecting deceit (Bogaard, Meijer, Vrij, & Merckelbach, 2016; DePaulo & Pfeifer, 1986), or in deciding which travellers should undergo customs inspection (Kraut & Poe, 1980). Thus, research and theory suggest that professionals may know better what information to look for, but once provided with the correct information cues, inexperienced and experienced individuals may asses this information similarly.

Based on this notion, we had no reason to expect that professional experience would affect the ability to distinguishing risk factors from neutral factors, or from protective factors. However, it was presumed that professional experience would affect the level of interrater-agreement, meaning that professionals who examined the same information would reach the same conclusion. Thus, we predicted that professionals would agree with one another on their risk assessments to a higher extent than non-

professionals would (Hypothesis 1). Furthermore, professionals were expected to know what information to look for. Therefore, we predicted that the professionals' (vs. non-professionals') search for information would be comparatively more in line with empirically supported threat cues (Hypothesis 2).

Method

Participants and Design

A total of 141 participants completed the survey. The sample consisted 45 professionals (30 men, 15 women, $M_{age} = 44.87$ years, $SD = 10.13$ years), 46 university students (12 men, 34 women, $M_{age} = 27.42$ years, $SD = 9.97$ years), and 50 laypersons (19 men, 31 women, $M_{age} = 42.48$ years, $SD = 13.01$ years). The sample size was guided by the number of professionals that could be recruited, meaning that the number of students and laypersons was matched with the number of professionals that participated. The reason for including two non-professional samples (students and laypersons) was that students may not fully represent the non-professional population, as they are studying to become professionals themselves. Working experience in threat assessment averaged 12.70 years ($SD = 8.96$) in the professional sample. One participant in the professional sample reported to have just one year of working experience in threat assessment and seven participants in the non-professional samples reported to have more than one year of working experience in threat assessment. The data of these eight individuals were excluded from further analyses. One hundred and thirty-three participants remained, of which 44 professionals (30 men, 14 women, $M_{age} = 45.23$ years, $SD = 9.95$ years), 44 students (12 men, 32 women, $M_{age} = 26.53$ years, $SD = 8.63$ years), and 45 laypersons (17 men, 28 women, $M_{age} = 42.09$ years, $SD = 12.80$ years)[1].

Professionals were approached via associations of threat assessment professionals in Europe (Association of European Threat Assessment Professionals, AETAP) and in Canada (Canadian Association of Threat Assessment Professionals, CATAP) to take part in a survey on threat assessment. They were either members of these associations, attended a conference on threat assessment in April 2016 that was organized by AETAP, or they were recommended as eligible participants by those who were

---

[1] Gender and age, when included as covariates in statistical analyses, did not change the pattern of the results and will be disregarded in the reported analyses.

initially approached. The countries in which the professionals operated were Canada ($n = 9$), The

Netherlands ($n = 8$), Germany ($n = 7$), the United States ($n = 3$), Australia ($n = 3$), worldwide ($n = 3$),

Austria, Belgium, Hong Kong, Luxembourg, New Zealand, South-Africa, Sweden, and Switzerland (all n

≤ 2). Professionals were not financially rewarded for their participation.

Students and laypersons were recruited for a study on how people make risk judgments. Students

were approached via the participant pool of a Swedish university where they had signed up for

participation in psychological research. They were compensated 50 SEK (approx. 6 dollar). Laypersons

were approached online through Amazon's Mechanical Turk (MTurk) and were paid 0.30 USD upon

completion. Participants in the laypersons sample were from the United Stated ($n = 41$) and India ($n = 9$).

All participants completed the same survey. The data collection took place from April to June 2016.

Materials

Three fictitious cases were constructed in which a person posed a potential threat of violence

towards one or several other persons. To prevent the outcomes from being the result of one particular case

or one particular domain of violence, each case reflected a different domain of violence that is commonly

encountered in the field of threat assessment. These domains were intimate partner violence (case

"Terry"), public figure violence (case "Marc"), and workplace violence (case "Frank"). The cases

consisted of 15 to 21 information cues describing i) the context in which the threat evolved and ii) the

behaviours and characteristics of the person posing the threat. Each information cue in the cases was

selected based on risk factors and protective factors that have been found to be empirically valid. For

instance, the information cue *Frank's position within the firm is uncertain* reflected the risk factor

"employment instability" (Feldman & Johnson, 1996). Protective factors are conditions and behaviours

that may reduce the risk of violence, which can also be the absence of risk factors (Borum, 2000). For

instance, the information cue *Frank joins a running group twice a week* indicated the absence of the risk

factor "social isolation" (Meloy, White, & Hart, 2013). In addition to the risk and protective factors, each

case held two pieces of neutral information. That is, information that is not empirically known to increase

or decrease the risk for violence as such (e.g., having tattoos). Two experienced threat assessment

professionals provided feedback on the cases. They recognized the cases as cases they would encounter in their profession, and hence, perceived the storylines as relevant and realistic. The cases are presented in the Appendix.

Procedure

The first page of the survey stated that the study included three descriptions of separate cases in which a person poses a potential threat. The participants were instructed to read each case and to answer the accompanying questions. Participants were informed that participating was confidential, voluntary, and for research purposes only. After agreeing on these terms, they were presented with the first case, followed by three questions. First, participants were asked to assess the overall risk for physical violence currently posed by the main character (0% = *no risk for violence*, 100% = *guaranteed risk for violence;* any number between 0 and 100 could be selected). Next, they were presented with a list of specific information cues derived from the case. They were asked to assess the extent to which each information cue increased or decreased the risk for violence in this particular case (e.g., *Frank's position within the firm is uncertain;* -4 = *very strong decrease*, 0 = *neutral*, 4 = *very strong increase*). Finally, participants were asked to list up to five additional information cues that they would request in order to make a more adequate assessment of the risk for violence for this particular case. The same procedure was repeated for each case. The cases were presented in a randomized order.

After completing the case-related questions, participants provided their age, gender, and number of years working in the field of threat assessment. In addition, professionals were asked to indicate in which country they worked and how they were professionally involved in threat assessment (e.g., screening and security, criminal intelligence gathering). Students and laypersons were asked to indicate their country of residence and their employment status (e.g., student, employed for wages). Participants needed approximately 30 minutes to complete the survey.

Data Analyses

Coding. Two research assistants coded the information cues requested by the participants. First, the total number of requested information cues was counted per participant per case (ranging from 1 to 5).

Second, each information request was classified as '*key request*' or '*other request*'. A request was

considered a 'key request' when it matched one of the 11 key questions identified by Vossekuil, Fein, and

Berglund (2015). These key questions are supposed to cover all areas that should be inquired in order to

make a fully informed assessment about a person posing a threat. Information requests that did not fit any

of the key questions were categorized as 'other request'. The 'other requests' were then further qualified

into specific areas of inquiry. All areas of inquiry specifying 'key requests' and 'other requests' are listed

in Table 1. Two coders categorized a random 30% of the material into 'key requests' vs. 'other requests'

(85.71%, Cohen's κ = .71) and further specified the requests into one of 17 specific areas of inquiry

(74.44%, Cohen's κ = .66). One coder completed codings on the remaining material.

Recoded values. Participants assessed the information in the cases on a 9-point scale with values

ranging from -4 to +4. These values were chosen to enhance the participants' visual representations of risk

decrease (minus) and risk increase (plus). However, the scale was recoded into values ranging from 1 to 9

for data analyses.

<div align="center">

Results

</div>

Initial Analysis of Risk Assessments

To reiterate, the assignment of a risk value (1 = *very strong decrease*, 9 = *very strong increase*) to

one information cue was considered one assessment. The three cases together contained 57 information

cues. Hence, each participant made 57 assessments of specific information cues. One-sample *t*-tests

revealed that professionals assessed information cues that reflected risk factors as risk increasing (i.e., the

average rating was significantly higher than the neutral value "5"; $M_{professional}$ = 6.44, SD = 0.56, $t(43)$ =

17.14, $p < .001$, $d = 5.17$), and information cues that reflected protective factors as risk decreasing (i.e., the

average rating was significantly lower than the neutral value "5"; $M_{professional}$ = 3.90, SD = 0.74, $t(43)$ =

9.93, $p < .001$, $d = 2.99$). Information cues that reflected neither risk factors nor protective factors were

rated as neutral (i.e., the average rating was statistically similar to the neutral value "5"; $M_{professional}$ = 4.97,

SD = 0.48, $t(43)$ = 0.45, $p < .652$, $d = 0.14$). Students and laypersons made ratings similar to those of

professionals for risk factors ($M_{student}$ = 6.31, SD = 0.56, $t(43)$ = 15.37, $p < .001$, $d = 4.63$; $M_{layperson}$ = 6.05,

$SD = 0.98$, $t(44)= 7.18$, $p < .001$, $d = 2.14$) and for protective factors ($M_{student} = 4.16$, $SD = 1.06$, $t(43)=$

5.27, $p < .001$, $d = 1.59$: $M_{layperson} = 3.86$, $SD = 1.25$, $t(44)= 6.01$, $p < .001$, $d = 1.79$). However, the neutral

cues were rated as neutral by laypersons ($M_{layperson} = 4.72$, $SD = 1.27$, $t(44)= 1.44$, $p = .156$, $d = 0.43$), but

as risk decreasing by students ($M_{student} = 4.49$, $SD = 1.34$, $t(43)= 2.56$, $p = .014$, $d = 0.78$).

The mean values for each information cue are displayed in Figure 1. As shown in the figure, the

professionals' assessments were highly similar to the assessments of the non-professionals. Only two out

of 57 information cues were assessed in one direction by professionals (risk decrease) and in the other

direction by non-professionals (neutral / risk increase). Although this could have been the result of chance,

the two items (*Rebecca has avoided all contact with Terry*; *Rebecca has received a new telephone number*

and moved to another neighbourhood) reflected a similar theme, which was, contact between victim and

threatener in partner violence. Furthermore, participants in all three samples assigned the highest risk

values to information cues that reflected a communicated threat.

Besides assessing specific information cues, participants were asked to judge the *overall* risk for

violence for each case (0% = *no risk*, 100% = *guaranteed risk*). A 3 (Sample: professionals vs. students

vs. laypersons) × 3 (Case: Terry vs. Marc vs. Frank) mixed ANOVA was performed on the violent risk

ratings, with sample as the between-subjects factor. The descriptive statistics are reported in Table 2. The

analysis yielded a significant Sample × Case interaction effect, $F(4, 256) = 4.01$, $p = .004$, $\eta_p^2 = .059$.

Tests of simple effects revealed different risk assessments across samples for case "Terry", $F(2, 128) =$

10.09, $p = < .001$, $\eta_p^2 = .136$. Professionals judged the overall risk for violence lower than did students (p

$< .001$) and laypersons ($p = .001$). No difference was found between students and laypersons ($p = .329$).

Furthermore, no differences across samples were found for case "Mark", $F(2, 128) = 0.24$, $p = .784$, $\eta_p^2 =$

$.004$, or for the case "Frank" , $F(2, 128) = 1.59$, $p = .208$, $\eta_p^2 = .024$.

Hypothesis Testing

Agreement. Agreement among assessors becomes apparent in the standard deviation of their

ratings, with the lower the standard deviation, the more similar the ratings. A 3 (Sample: professionals vs.

students vs. laypersons) × 3 (Case: Terry vs. Marc vs. Frank) between-items ANOVA was conducted on

the standard deviations for the information cues. A main effect for sample occurred, $F(2, 162) = 106.67$, p

$< .001$, $\eta_p{}^2 = .568$. Planned contrasts showed that the average standard deviation within the professional

sample ($M = 1.14$, $SD = 0.21$) was significantly lower than the average standard deviation within the

student sample ($M = 1.48$, $SD = 0.26$, $p < .001$) and the layperson sample ($M = 1.75$, $SD = 0. 91$, $p < .001$).

Post hoc test further revealed that the difference between the student sample and the layperson sample was

significant too ($p < .001$). The main effect of case was not statically significant, $F(2, 162) = 0.52$, $p = .598$,

$\eta_p{}^2 = .006$, neither was the Sample x Case interaction effect, $F(4, 162) = 1.51$, $p = .200$, $\eta_p{}^2 = .036$. Thus,

the results indicate that the consensus among the professionals was systematically (i.e., regardless of the

case) higher than the consensus among the non-professionals. This supports Hypothesis 1.

       The above analysis does not account for where on the 9-point scale the spread was located. Spread

on the lower (1-3), central (4-6), or higher part (7-9) of the scale was weighted equally. However, when

examining agreement using this scale, the spread on the central part may be particularly relevant because

it contains both a risk decreasing value (4) and a risk increasing value (6). More specific, agreement on the

direction of risk (i.e., increase or decrease) can be considered more critical than agreement on the

magnitude of risk (e.g., strong increase or very strong increase). To test for differences between samples

in agreement on direction, a chi-square test was performed on the number of confidence intervals within

each sample that included the central value of the scale ("5"). Confidence intervals containing the middle

value imply that some participants in the sample rated the information as risk decreasing (values below 5),

whereas others in the sample rated the same information as risk increasing (values above 5). The six

neutral information cues were excluded from the analysis because these cues were expected to obtain

ratings close to the value of 5. The analysis was conducted on the assessments of the remaining 51

information cues. A significant difference across samples was found with regard to the number of

confidence intervals containing the value 5, $\chi^2 (2, N = 153) = 8.70$, $p = .013$, $\varphi = .24$. Thus, in further

support of Hypothesis 1, professionals rarely disagreed with each other on whether information should be

assessed as risk increasing or risk decreasing ($N = 3$, 5.9%), whereas such disagreement among laypersons

occurred in approximately one third of the assessments made ($N = 14$, 27.5%, $p < .05$). The frequency of

disagreement among students ($N = 8$, 15.7%) was not found significantly different from the other two

samples.

   Information search. After assessing the risk for violence, participants were asked to list up to

five additional information cues that they considered relevant for making a more adequate assessment.

The requested information cues were classified into 'key requests' and 'other requests', and further

divided into different areas of inquiry (see the Method section). ANOVA's were conducted for each case

on (a) the total number of information requests, (b) the number of 'key requests', (c) the number of 'other

requests', and (d) the number of different areas that were inquired. The latter measure served as an

indication for the variety of the requested information. Descriptive and inferential statistics are presented

in Table 3. Planned contrasts confirmed Hypothesis 2. Professionals requested significantly more key

information than students (all cases $p < .001$) and laypersons did (all cases $p < .001$). Furthermore,

requests that were made by professionals covered significantly more different areas of inquiry compared

to request made by students (all cases $p < .001$) or laypersons (all cases $p \leq .001$). Post-hoc tests showed

no differences between students and laypersons on any of the measures. Thus, when given the opportunity

to gather additional data, professionals requested comparatively more relevant information and from a

comparatively wider variety of domains.

<p style="text-align:center">Discussion</p>

   The results demonstrate that threat assessment professionals agreed more with one another on

their risk assessments compared to non-professionals. Furthermore, professionals requested comparatively

more relevant information (i.e., information that was consistent with empirical threat cues) and they based

their risk assessment on information that covered more domains. Both the relevance and the range of the

information gathered relates to the quality of threat assessments (Fein & Vossekuil, 1998). On a more

general level, empirically supported decision-making as well as inter-rater agreement reflect quality

standards in expert performance (Dror, 2016; Einhorn, 1974). Hence, the results suggest that domain-

specific experience advances threat assessment practices. With that said, professionals' beliefs about what

cues were (and were not) alarming were highly similar to non-professionals' beliefs, suggesting that the actual interpretation of information in threat cases may be rather intuitive.

The current findings fit well in the state of science on expert performance. Early theories proposed that expert knowledge is more critical for cue selection (i.e., what information should be looked for) than for cue weighting (i.e., how should the information be interpreted; Dawes, 1979; Einhorn, 1974). Advanced performers, ranging from chess grandmasters to cardiologists, and intelligence analysts, have been found to be particularly good at extracting a meaningful gist from larger sets of data (De Groot, 1978; Corbin, Reyna, Weldon, & Brainerd, 2015; Reyna & Lloyd, 2006). In other words, their experience directs their attention to critical information. When interpreting this information, however, studies have shown that professionals, just like laypersons, often rely on common sense (Masip, Herrero, Garrido, & Barba, 2011; Stolper et al., 2011). Drawing more specifically on literature on risk perception, it has been theorized that people rely on two different systems when they assess risk in financial, health, and safety domains: an affective and cognitive system (Slovic et al., 2004). The cognitive system contains, among others, knowledge on probabilities of danger to occur and is better developed in experts than in novices (Weber, 2001), whereas the affective system, containing emotional reactions to risk, is inherently present and equally developed in all humans. Although the two systems operate simultaneously, the affective response to risk is assumed to be dominant (Damasio, 2001; Slovic et al., 2004). This theory might explain why common sense notions on risk can overrule expert knowledge when assessing threats of violence. However, the role of affect in professional threat assessment yet needs to be investigated.

The consensus found among professionals is to some extent reassuring. It suggests that threat assessment professionals tend to reach similar conclusions when assessing the same case. This sheds a positive light on threat assessment practices, as research on other legal domains reveals that disagreement among professionals is not uncommon. For instance, forensic experts were found to produce inconsistent interpretations with regard to DNA evidence (Dror & Hampikian, 2011), fingerprints (Ulery, Hicklin, Buscaglia, & Roberts, 2012), and footwear (Majamaa & Ytti, 1996). Although such inconsistencies are worrisome, it should be acknowledged that consistent interpretations are not necessarily accurate. Whether

the professionals' judgments for the current study were accurate or not could not be established given the fictitious cases used. Yet, their judgments matched the risk and protective factors that are known from the literature, which speaks to the quality of their assessments.

Limitations and Future Research

The study is limited in ways that might be addressed in future research. First, there is no guarantee that the information cues in the cases are directly tied to one particular factor, whether risk, neutral, or protective. For instance, suffering from diabetes was included as a neutral factor since diabetes has never been proven to relate to the enactment of threats. However, a participant could have reasoned that suffering from a chronic disease is a stressful condition and should thus be assessed as information that is risk increasing. Such reasoning would be in line with research showing that personal stressors may indeed indicate a higher risk for violence (Meloy et al., 2013). As the current study did not tap into the participants' reasoning behind their assessments, it remains unknown if (and if so, how) participants connected the information cues to possible risk factors.

Another potential concern is that the cues in the cases may have been self-evident, meaning that their interpretation was straightforward. For example, the cue *Marc is highly frustrated that the Minister has not replied*. It is quite clear that frustration is risk increasing rather than risk decreasing. If the information given was less ambiguous than the information threat assessors typically face in real life, this may form an alternative explanation for the similarities found between the professionals and non-professionals. With that said, measures were taken to mitigate the risk of biased materials. The cases were perceived as realistic by two experienced threat assessment professionals and the cases contained risk factors known from the literature on threat assessment. Future studies could address potential limitations in the materials by extensive pre-testing and/or by studying threats assessors in a naturalistic setting. Naturalistic Decision-Making (NDM) research has provided valuable insights on how to improve professional performances (Klein, 2008). The domain of threat assessment suits the NDM characteristics well, as threat assessors—just like criminal investigators—typically operate in an uncertain and dynamic environment in which the stakes are high and time is limited (Ask & Alison, 2010).

A brief note is warranted on a possible bias in the professional sample. Professionals may have been more motivated than non-professionals as the study concerned their own field of expertise. This could explain the higher number of information requests on behalf of the professionals. Yet, motivation does not explain the higher number of *relevant* information requests made by the professionals. The latter may still be best explained by their professional experience.

The present study solely focussed on threat assessment but the responsibilities of threat assessment professionals typically also involve *threat management* (Meloy, Hart, & Hoffmann, 2014b). Threat management refers to violence prevention and concerns questions such as: What action should be taken to mitigate the threat? Who should be involved in such action? And, how should the case be monitored? Future research might profit from examining how professional experience affects threat management tasks. Few sentences on how current study could have added an interview with profs

Taken together, the current study did not mirror the full dynamics and complexity of threat assessment work. However, the controlled and systematic set-up of the study holds great value because it allows for normative conclusions (i.e., "Outcome A is more/better than outcome B") rather than mere descriptive outcomes. Such a semi-experimental approach is novel to the field of threat assessment and it may contribute to a stronger scientific base.

Contribution to the Field

The current findings indicate that threat assessment professionals may contribute most during the *information gathering* stage. Professionals may advance the information search by setting the right scope, filtering diagnostic cues from less relevant information, and recognizing what intelligence is still missing. The field of threat assessment can benefit from such knowledge in three important ways. First, those who consult threat assessment professionals can make more deliberate decisions when allocating responsibilities and resources. Second, the professionals can better explain to their clients what may be expected of them. Third, and most important, having the right person on the right job improves the quality of threat assessments, which should ultimately result in a safer society.

Acknowledgement

References

Ask, K., & Alison, L. (2010). Investigators' decision making. In P. A. Granhag (Ed.), *Forensic psychology in context: Nordic and international perspectives* (pp. 35-55). Cullompton, UK: Willan.

Bogaard, G., Meijer, E. H., Vrij, A., & Merckelbach, H. (2016). Strong, but wrong: lay people's and police officers' beliefs about verbal and nonverbal cues to deception. *PloS* ONE, *11*. doi.org/10.1371/journal.pone.0156615

Borum, R. (2000). Assessing violence risk among youth. *Journal of Clinical Psychology*, *56*, 1263 - 1288. doi.org/10.1002/1097-4679(200010)56:10<1263::AID-JCLP3>3.0.CO;2-D

Calhoun, F. S., & Weston, S. W. (2003). *Contemporary threat management: A practical guide for identifying, assessing, and managing individuals of violent intent*. San Diego, CA: Specialized Training Services.

Corbin, J. C., Reyna, V. F., Weldon, R. B., & Brainerd, C. J. (2015). How reasoning, judgment, and decision making are colored by gist-based intuition: A fuzzy-trace theory approach. *Journal of Applied Research in Memory and Cognition*, *4*, 344-355. doi.org/10.1016/j.jarmac.2015.09.001

Damasio, A. R. (2001). Emotion and the human brain. *Annals of the New York Academy of Sciences*, *935*, 101-106. doi.org/10.1111/j.1749-6632.2001.tb03475.x

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*, 571-582. doi.org/10.1037/0003-066X.34.7.571

De Groot, A. D. (1978). *Thought and choice in chess* (2nd ed.). The Hague: Mounton Publishers.

DePaulo, B. M., & Pfeifer, R. L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology*, *16*, 249-267. doi.org/10.1111/j.1559 1816.1986.tb01138.x

Dietz, P., & Martell, D. A. (2010). Commentary: approaching and stalking public figures—a prerequisite to attack. *Journal of the American Academy of Psychiatry and the Law Online*, *38*, 341-348.

Douglas, K. S., Ogloff, J. R., & Hart, S. D. (2003). Evaluation of a model of violence risk

assessment among forensic psychiatric patients. *Psychiatric Services*, *54*, 1372-1379.

doi.org/10.1176/appi.ps.54.10.1372

Dror, I. E. (2016). A hierarchy of expert performance. *Journal of Applied Research in Memory

and Cognition*, 5, 121-127. doi.org/10.1016/j.jarmac.2016.03.001

Dror, I. E., & Hampikian, G. (2011). Subjectivity and bias in forensic DNA mixture interpretation.

Science & Justice, *51*, 204-208. doi.org/10.1016/j.scijus.2011.08.004

Einhorn, H. J. (1974). Expert judgment: Some necessary conditions and an example. *Journal of

Applied Psychology, 59*, 562-571. doi.org/10.1037/h0037164

Elstein, A. S., & Schwarz, A. (2002). Clinical problem solving and diagnostic decision making:

selective review of the cognitive literature. *British Medical Journal*, *324*, 729-732.

doi.org/10.1136/bmj.324.7339.729

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: a general overview.

Academic Emergency Medicine, *15*, 988-994. doi.org/10.1111/j.15532712.2008.00227.x

Fahsing, I., & Ask, K. (2016). The making of an expert detective: the role of experience in English and

Norwegian police officers' investigative decision-making. *Psychology, Crime & Law*, *22*, 203

223. doi.org/10.1080/1068316X.2015.1077249

Fein, R. A., & Vossekuil, B. (1998). *Protective intelligence and threat assessment

investigations: A guide for state and local law enforcement officials*. Washington, DC: US

Department of Justice, Office of Justice Programs, National Institute of Justice.

Feldman, T., & Johnson, P. (1996). Workplace violence: A new form of lethal aggression. In Hall, H. V.

(Ed.), *Lethal violence 2000: A sourcebook on fetal domestic, acquaintance and stronger

aggression* (pp. 311—338). Kamuela, HI: Pacific Institute for the Study of Conflict and

Aggression.

Guy, L. S., Packer, I. K., & Warnken, W. (2012). Assessing risk of violence using structured

professional judgment guidelines. *Journal of Forensic Psychology Practice*, *12*,

270-283. doi.org/10.1080/15228932.2012.674471

Hoffmann, J., Meloy, J. R., Guldimann, A., & Ermer, A. (2011). Attacks on German public figures, 1968
2004: Warning behaviours, potentially lethal and non-lethal acts, psychiatric status, and
motivations. *Behavioral Sciences & the Law*, *29*, 155-179. doi.org/10.1002/bsl.979

James, D. V., MacKenzie, R. D., & Farnham, F. R. (2014). *Communications Threat Assessment
Protocol.* London: Theseus LLP.

Joseph, G. M., & Patel, V. L. (1990). Domain knowledge and hypothesis generation in diagnostic
reasoning. *Medical Decision Making, 10,* 31-46. doi.org/10.1177/0272989X9001000107

Klein, G. (2008). Naturalistic decision making. *Human Factors: The Journal of the Human Factors and
Ergonomics Society*, *50*, 456-460. doi.org/10.1518/001872008X288385

Kraut, R. E., & Poe, D. B. (1980). Behavioral roots of person perception: The deception
judgments of customs inspectors and laymen. *Journal of Personality and Social
Psychology*, *39*, 784-798. doi.org/10.1037/0022-3514.39.5.784

Majamaa, H., & Ytti, A. (1996). A survey of the conclusions drawn of similar footwear cases in
various crime laboratories. *Forensic Science International*, *82*, 109–120. doi.org/10.1016/0379
0738(96)01972-X

Masip, J., Herrero, C., Garrido, E., & Barba, A. (2011). Is the behaviour analysis interview just
common sense? *Applied Cognitive Psychology*, *25*, 593-604. doi.org/10.1002/acp.1728

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of
the evidence.* Minneapolis: University of Minnesota Press. doi.org/10.1037/11281-000

Meloy, J. R., Hart, S. D., & Hoffmann, J. (2014b). Threat assessment and threat management.
In J. R. Meloy & J. Hoffmann (Eds.), *International handbook of threat assessment*
(pp. 3-17). New York, NY: Oxford University Press.

Meloy, J. R., Hoffmann, J., Roshdi, K., Glaz-Ocik, J., & Guldimann, A. (2014a). Warning
behaviours and their configurations across various domains of targeted violence. In J.
R. Meloy & J. Hoffmann (Eds.), *International handbook of threat assessment* (pp.39-53). New

York, NY: Oxford University Press.

Meloy, J. R., White, S. G., & Hart, S. (2013). Workplace assessment of targeted violence risk: The development and reliability of the WAVR-21. *Journal of Forensic Sciences*, *58*, 1353-1358. doi.org/10.1111/1556-4029.12196

Norman, G. (2005). Research in clinical reasoning: past history and current trends. *Medical Education, 39*, 418-427. doi.org/10.1111/j.1365-2929.2005.02127.x

Patel, V. L., & Groen, G. J. (1986). Knowledge based solution strategies in medical reasoning. *Cognitive Science*, *10*, 91-116. doi.org/10.1207/s15516709cog1001_4

Reyna, V. F., & Lloyd, F. J. (2006). Physician decision making and cardiac risk: Effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *Journal of Experimental Psychology: Applied, 12*, 179-195. doi.org/10.1037/1076-898x.12.3.179

Schmidt, H. G., & Rikers, R. M. (2007). How expertise develops in medicine: knowledge encapsulation and illness script formation. *Medical Education*, *41*, 1133-1139. doi.org/10.1111/j.1365-2923.2007.02915.x

Shanteau, J., Weiss, D. J., Thomas, R. P., & Pounds, J. C. (2002). Performance-based assessment of expertise: How to decide if someone is an expert or not. *European Journal of Operational Research*, *136*, 253-263. doi.org/10.1016/S0377-2217(01)00113-8

Skeem, J. L., & Monahan, J. (2011). Current directions in violence risk assessment. *Current Directions in Psychological Science*, *20*, 38-42. doi.org/10.1177/0963721410397271

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis*, *24*, 311-322. doi.org/10.1111/j.0272-4332.2004.00433.x

Slovic, P., & Weber, E. U. (2002, April). *Perception of risk posed by extreme events*. Paper prepared for discussion at the conference Risk Management Strategies in an Uncertain World, New York.

Stolper, E., Van de Wiel, M., Van Royen, P., Van Bokhoven, M., Van der Weijden, T., & Dinant,

G. J. (2011). Gut feelings as a third track in general practitioners' diagnostic reasoning. *Journal of General Internal Medicine*, *26*, 197-203. doi.org/10.1007/s11606-0101524-5

Ulery, B. T., Hicklin, R. A., Buscaglia, J., & Roberts, M. A. (2012). Repeatability and reproducibility of decisions by latent fingerprint examiners. *PloS ONE*, *7*, e32800. doi.org/10.1371/journal.pone.0032800

Vossekuil, B., Fein, R. A., & Berglund, J. M. (2015). Threat assessment: Assessing the risk of targeted violence. *Journal of Threat Assessment and Management*, *2*, 243. doi.org/10.1037/tam0000055

Weber, E. U. (2001). Decision and choice: Risk, empirical studies. In N. J. Smelser & B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences*. (pp. 13347-13351). Amsterdam, NL: Elsevier. doi.org/10.1016/b0-08-043076-7/00634-3

White, S. G., & Meloy, J. R. (2010). The WAVR-21: A structured professional guide for the workplace assessment of violence risk. *San Diego: Specialized Training Services*.

Table 1

Areas of Inquiry Used to Classify Requested Information

Key requests (Vossekuil et al., 2015)

---

1. What are the subject's motives and goals?
2. What has the subject communicated about his or her intentions to anyone (target, law enforcement, family, friends, colleagues, associates, diary/journal)?
3. Is there evidence that the subject has engaged in attack-related behaviours?
4. Is there evidence that the subject has engaged in menacing, harassing, and/or stalking type behavior?
5. Does the subject have a history of mental illness involving command hallucinations, delusional ideas, feelings of persecution, and so forth, with indications that the subject has acted on those beliefs?
6. How organized is the subject? Does the subject have the ability to plan and execute a violent action?
7. How organized is the subject? Does the subject have the ability to plan and execute a violent action?
8. Is what the subject says consistent with his or her actions?
9. Does the subject see violence as acceptable, desirable, or the only way to solve problems?
10. What concerns do those who know the subject have about the subject's behavior?
11. What factors in the subject's life and/or environment could change and thereby increase the subject's risk of attacking? What factors could change and thereby decrease the risk posed?

Other requests

---

1. Requests about upbringing
2. Requests about criminal records that do not concern violence or stalking
3. Requests about drugs, alcohol and medication that do not concern mental illness
4. Requests about protective factors that do not concern the future
5. Requests about finances without a specified reason for the request
6. Requests that do not fit any of the previous categories

Table 2

Mean Percentages for Violence Risk per Case Judged by Professionals, Students, and Laypersons

	Professionals	Students	Laypersons
Case "Marc"	43.05% (20.17)	46.11% (20.75)	43.56% (24.54)
Case "Terry"	57.83% (23.84)	76.07% (17.29)	71.98% (17.53)
Case "Frank"	48.62% (23.87)	47.73% (22.25)	55.44% (20.89)

Note. Mean percentages are based on a rating scale ranging from 0% (*no risk for violence*) to 100% (*guaranteed risk for violence*). The numbers in parentheses represent standard deviations.

Table 3

Means and Inferential Statistics for the Analyses of the Requested Information

Measure	Professionals	Students	Laypersons	F	p	η_p^2
Case "Marc"						
No. of information pieces requested	3.66 (1.55)	2.16 (1.60)	2.56 (1.67)	10.27	.000	.136
'key requests'	2.68 (1.63)	1.18 (1.08)	1.62 (1.13)	15.42	.000	.192
'other requests'	0.98 (1.00)	0.95 (1.18)	0.93 (1.23)	0.02	.984	.000
No. of areas inquired	3.23 (1.46)	1.86 (1.31)	2.13 (1.34)	12.24	.000	.158
Case "Terry"						
No. of information pieces requested	3.91 (1.54)	2.11 (1.70)	2.51 (1.53)	14.68	.000	.185
'key requests'	2.64 (1.63)	1.07 (1.19)	1.22 (0.97)	22.70	.000	.260
'other requests'	1.27 (1.11)	1.05 (1.16)	1.29 (1.49)	0.25	.779	.004
No. of areas inquired	2.95 (1.40)	1.64 (1.33)	1.91 (1.02)	12.46	.000	.162
Case "Frank"						
No. of information pieces requested	3.82 (1.72)	2.27 (1.48)	2.49 (1.62)	11.93	.000	.155
'key requests'	2.64 (1.54)	1.34 (1.12)	1.56 (1.16)	12.84	.000	.165
'other requests'	1.18 (1.24)	0.98 (1.30)	0.93 (1.16)	0.51	.602	.008
No. of areas inquired	3.11 (1.56)	2.00 (1.24)	2.09 (1.26)	9.15	.000	.123

Note. The numbers in parentheses represent standard deviations.
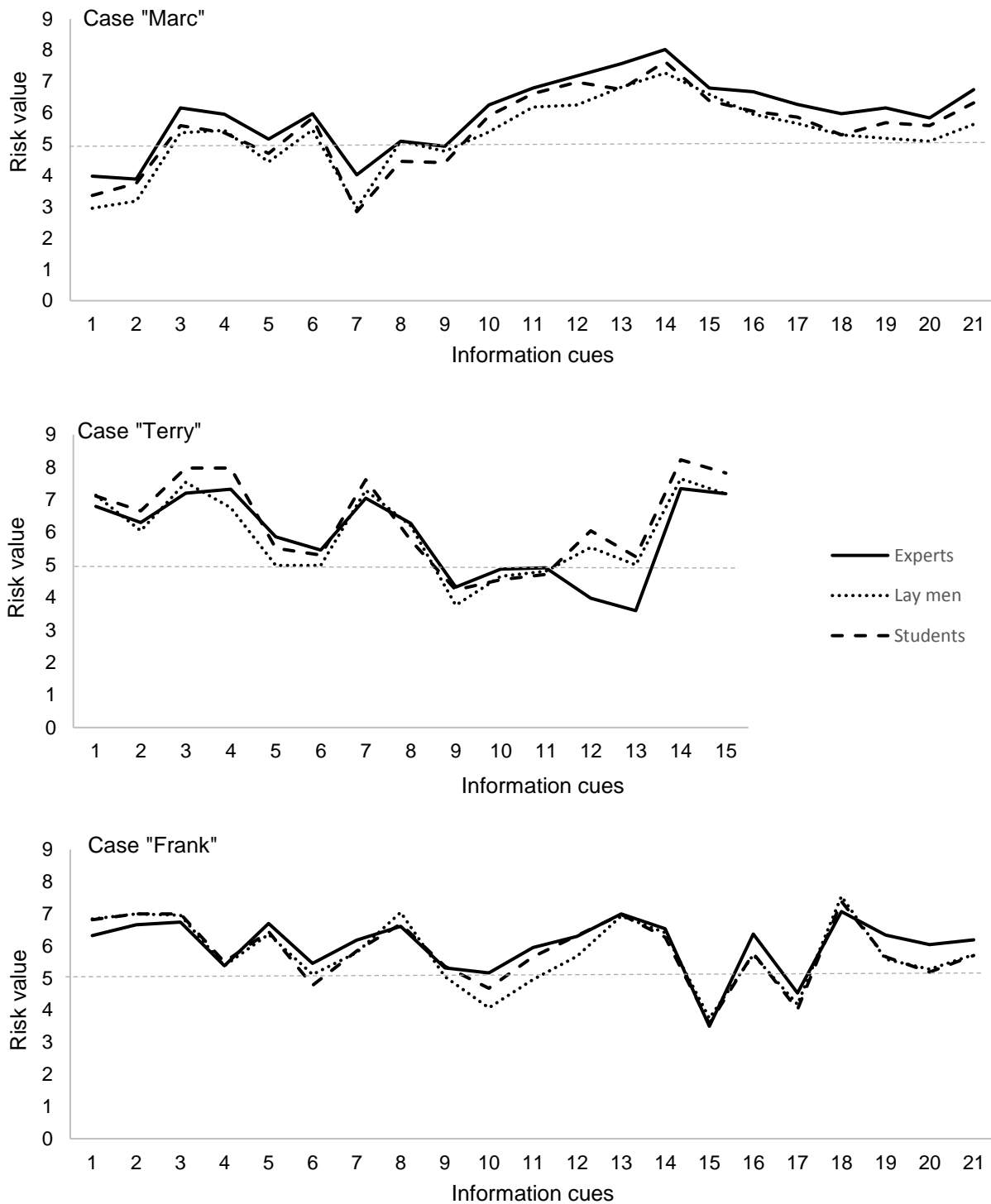
Figure 1. Risk assessments (1 = *very strong risk decrease*, 9 = *very strong risk increase*) that professionals, students, and laypersons made for the information cues presented per case.

Appendix

Case "Terry"

After a one-year relationship, Rebecca (29 years old) decided to break up with her boyfriend Terry (38 years old). Terry has been very bitter about this sudden ending. He felt betrayed and expected that Rebecca would change her mind and come back to him. One month ago, when Rebecca came by his house to pick up her belongings, Terry told her that she better think twice about the breakup if her life is worth something to her (14). Rebecca went to the police and reported that she has been threatened by her ex-boyfriend. So far, Rebecca has avoided all contact with Terry (12). She has received a new telephone number and moved to another neighborhood (13). Rebecca has a new boyfriend and Terry recently found out about this (15).

The police took up the case and gathered some more information. It appeared that when Terry was a child, he often bullied other children and ruined their property (1). Terry was expelled from school at the age of 16 (2) and started working in a logistics company. At the age of 19, he served 5 years in prison for bank robbery (3). Back then Terry stated that he did not feel bad for the bank employees who were present during the robbery, as it was not their money he had taken (4). Terry left his latest job as a construction worker because he could not get along with his boss (8). He is now thinking of setting up his own protein drink business (9). Terry collects exotic reptiles at home, such as lizards and chameleons (10). His back, chest and arms are covered with tattoos of different types of reptiles (11). Furthermore, Terry has had many short-term relationships with women (5). He has four children with three of his ex-girlfriends. None of the children are Rebecca's (6). When Terry goes out with friends, he sometimes uses drugs, such as XTC or speed (7).

Case "Marc"

Marc (24 years old) studies chemistry at the university and is in his final year. He has always been an excellent student (1). Two years ago, Marc read an article claiming that the pharmaceutical industry is

contaminated by a few influential companies with malevolent intent. The article triggered Marc's interest, even though he had never used substances or medication other than homeopathic cures (2). Since then, Marc has delved into the topic (10) and is more and more convinced that the pharmaceutical industry is a corrupt business (11). Marc believes that the government should take responsibility by ending collaboration with big pharmaceutics companies. Marc has written several letters to the Minister of Health and Social Affairs. After all the hours he has spent helping the Minister to make the right decision, Marc is highly frustrated that the Minister has not replied (12). One week ago, during a late evening round, a security guard found Marc in one of the university labs mixing various powders and liquids (13). When the guard asked him what he was doing, Marc answered: "If the Minister does not want to listen to reason, I will show him how poisonous their medication can be" (14). The security guard informed the university about this incident and Marc was temporarily suspended (15).

The university consulted an expert to investigate the case. This expert discovered that Marc's most recent letters to the Minister were accompanied by pictures and videotaped messages (16), in which he outlined his case against the pharmaceutical industry (17). In one video, Marc made a presentation for the National Pharmaceutical Association (NPA) (18). The NPA will have their annual board meeting one month from now (19). During this meeting, they will discuss new medicines and treatments. The meeting will take place 120 kilometers from the city Marc lives in (20). The Minister will attend this meeting (21). Further information from Marc's acquaintances reveals that Marc was brought up by his grandparents (5), as his parents could not take care of him (6). Marc's clothing style is described by others as "gothic" (8) and he likes to listen to heavy metal bands such as *Led Zeppelin* (9). Marc never had many friends (3) and prefers to spend time alone (4). Yet, people described him as a friendly and good hearted guy (7).

Case "Frank"

Frank (52 years old) works as an administrator for a big law firm. He started in this job 30 years ago when he was a law student. Back then he hoped to eventually work there as lawyer; much to his disappointment,

he has never been promoted (8). Nevertheless, Frank did not resign from his job (9) and often tells

outsiders about his company and the high-profile cases they handle (10). Three weeks ago, Frank's

manager called him into his office. As Frank's 30 year anniversary at the firm was coming up (11), Frank

expected that they would talk about the arrangements for the party (12). Instead, the manager discussed

the upcoming re-organization of the company and forced redundancies. The manager told Frank that his

position with the firm is uncertain (13). Frank has called in sick every day since then (14). Yesterday, a

note was taped on the entrance of the firm stating: "I hate you all and one day soon you will hate me too"

(18).

The firm consulted their internal Security & Integrity Team (SIT) and mentioned Frank as a

subject of concern. SIT gathered the following additional information. Frank has type 1 diabetes and

therefore needs daily insulin doses (6). Furthermore, Frank's acquaintances state that he is often angry (1)

and feels that he is being wronged (2). From Frank's perspective, people are either for or against him (3).

Frank and his wife divorced 8 years ago (4) and he has suffered from depressive episodes ever since (5).

Frank and his ex-wife bought an expensive house together. After the divorce, Frank was able to keep the

house, but a large share of his monthly income goes to the mortgage payment (7). Frank normally joins a

running group twice a week (15), however, he has not shown up in the last three weeks (16). Once per

year, Frank travels abroad to attend a three-day science fiction event (17).

While Frank's case is under examination, the management of the law firm has called for an

internal meeting with the entire staff (19). The new management arrangements will be explained and the

exact number of redundancies will be announced. The meeting will be held four days from now (20).

Frank has been invited to attend (21).

Note. The numbers in parentheses indicate the specific pieces of information that were extracted for risk

assessment.