Assessing Credibility Through Text: A Preliminary Analysis for Identifying Psychopathy

Jeffrey Hancock
Cornell University
jth34@cornell.edu

Michael T. Woodworth
University of British Columbia
michael.woodworth@ubc.ca

Rebecca Morrow
Cornell University
rm357@cornell.edu

Hugh McGillivray
University of British Columbia
huge116@hotmail.com

Rachel Boochever
Cornell University
rcb264@cornell.edu

Abstract
This paper reports on two studies that used computerized textual analysis to examine the discourse characteristics of positive and negative event stories told by psychopaths. In the first study, psychopathic criminal offenders were expected to display linguistic characteristics consistent with narcissistic personality disorder and to display a greater degree of psychological distancing than non-psychopathic offenders. Their language use supported these predictions, with psychopaths producing pronoun patterns consistent with narcissism and verb tense patterns consistent psychological distancing. In a follow up preliminary study, we examined whether these results could be extended to the general population and the use of the Self-Report Psychopathy (III) scale. The preliminary results are limited but encouraging. Overall the results suggest that a psychopath’s narcissistic personality and psychological distancing can be detected in language production.

1. Introduction
Assessing credibility can take on many forms, from determining the veracity of a statement to identifying trustworthiness of an individual based on their behavior. This second approach focuses on assessing whether an individual is credible or more likely to be a threat. In the present study we examine whether the language of an individual can provide important insights into their level of psychopathy, an important personality disorder that is associated with criminality, deception and manipulation. In the first study, we examine a sample of transcripts produced by convicted homicide offenders with clinically identified psychopaths. Our primary research question is to determine whether a story-telling technique that varies positively and negatively valenced biographical events can be combined with automated linguistic analysis to reliably distinguish discourse patterns between psychopaths and controls.

2. Attributes of the Psychopath
Hervey Cleckley [1] first conceptualized psychopathy in 1941 in his book The Mask of Sanity. Cleckley described psychopaths as masking their personality disorder and mimicking a normal functioning person. Today, psychopaths are defined as having an emotional deficit and antisocial behavioral tendencies. They lack remorse and guilt and have shallow affect, although not appearing to lack any intellect. Often considered great communicators, psychopaths frequently manipulate others to get what they want [2]. According to Paulhus and colleagues [3], the constructs of a psychopathic personality also overlap with the constructs of narcissism and Machiavellianism, making up the “Dark Triad” of personality. All three personality constructs correlate with disagreeableness, thus lending to the “darkness” moniker.

Psychopaths comprise about 1% of the general population [2]. Psychopaths share a number of characteristics including egocentricity, impulsiveness, shallow emotions and low empathy, pathological lying, and a willingness to violate social norms [5]. In addition to these distinctive personality characteristics, psychopaths have a high propensity for crime: in a federal offender sample, psychopaths committed more violent crimes compared with nonpsychopathic offenders (Porter & Porter, 2007). Within one year of committing violent crimes, psychopaths are more likely than nonpsychopaths to be repeat offenders for additional serious violent crimes [6].

3. Detecting Psychopaths Through Language
Despite this tendency to violence, psychopaths present themselves as normal people, wearing a “mask of sanity” [1]. Can this mask be lifted by analyzing the language of the psychopath? Recent research by our research group provides some preliminary indications that psychopaths use language in ways distinct from non-psychopaths.

Until recently, no automated language analysis programs had been employed to analyze the speech production of psychopathic individuals. Previous studies that employed human coders have suggested particular language characteristics of the psychopathic individual. For example, Porter and Woodworth [7] found that individuals scoring higher on psychopathy were more likely to exaggerate the reactivity of the homicide they committed and to omit some core detail of the incident.

Using automated language analysis programs, however, is arguably preferable considering that many of the aspects of language measured with these programs are not consciously controllable by the speaker and difficult to manually and reliably code by human raters. Psychopaths are known to be particularly skilled at manipulating, deceiving, and controlling their self-presentation, making an automated enquiry into their language production particularly useful for obtaining important insights into their behavior. Further, if they do demonstrate particular types of language patterns, it might be possible to more readily detect them in online environments where the vast majority of information is text-based.

Hancock, Woodworth, & Porter [8] used text analysis tools to examine the crime narratives of psychopathic and non-psychopathic homicide offenders. The offenders were all interviewed by clinical psychologists, during which the offenders were asked to describe the day of their murder.

Psychopaths showed reliable differences relative to their non-psychopathic counterparts, such as focusing more on material needs during their narratives (food, drink, money), and making fewer references to social needs (family, religion/spirituality). Psychopaths also used more past tense and less present tense verbs in their narrative, suggesting a greater psychological and emotional detachment from the murder. Consistent with the emotional deficiencies that are the hallmark of the psychopath, their language when describing their crime was also less emotionally intense and pleasant as their scores of psychopathy increased.

Taken together, this initial study suggests that the language of the psychopath includes distinct patterns from non-psychopathic individuals, at least in the context of describing a serious crime.

4. Study 1

This initial study is important in demonstrating that language can reflect psychopathy, and that these discourse patterns can be identified using computerized language analysis. A major limitation for generalizing this work to more general credibility applications is that the study examined narratives of murder. Obviously, for most contexts for which assessing credibility is important, such as in hiring, security clearance, etc., asking people to describe the day that they murdered someone is not feasible.

In the present study we examine a more generalizable technique that involves asking individuals to describe a positive and negative story autobiographical story. This approach is highly generalizable in that it can be asked of anyone, and it can provide some ability to discriminate across the two types of emotional stories in ways that may be particularly discriminative in assessing psychopathy.

For instance, psychopaths often display characteristics consistent with the narcissistic personality, such as an aggressive-sadistic personality style, self-love, arrogance, and other-exploitation. Indeed, psychopathy is often present at the same time as, or co-morbid with, narcissistic personality disorder [9]. Stone notes that: “all psychopathic persons are at the same time narcissistic persons” (p. 162). Unlike psychopaths, however, narcissistic people feel guilt and remorse due to their negative actions whereas the psychopath does not appear to have any sense of guilt or empathy [9].

The fact that psychopaths tend to exhibit narcissism is potentially important for linguistic analysis. Psychological work examining narcissism suggests that there is a link between narcissistic personality disorder and personal pronoun use. The use of “I” in language can be interpreted as a measure of egocentrism, or self focus, because its primary function is to distinguish between the self and the other [10]. In one experiment, Raskin and Shaw [10] asked subjects to talk about a topic of their choice for 5 minutes. After the monologue subjects took the Narcissistic Personality Inventory. They found that subjects with higher narcissism scores used “I” more and “we” less than subjects with lower narcissism scores.

Given that psychopaths show a high co-morbidity with narcissism, and that narcissism is linked with higher rates of first person singular pronoun use, psychopaths should produce higher rates of personal
pronouns and lower rates of other-oriented pronouns than controls.

H1: Psychopaths will use first person singular pronouns (“I” “me”) more frequently than controls but will use first person plural pronouns (“we” “us”) less frequently than controls.

While we expect this to be true overall, how will the use of pronouns change across positive and negative stories? In general, people should use fewer first person singular and more plural when describing negative events to decrease the focus on self, while the pattern should be reversed for positive stories. We expected this pattern to be exaggerated for psychopaths given that their narcissism should make it more difficult for them to associate the self with negative events but make it more likely for them to associate the self with positive events.

H2a: Psychopaths will use relatively more first person singular but fewer plural when discussing negative stories compared to controls.

H2b: Psychopaths will use relatively fewer first person singular but more plural when discussing negative stories compared to controls.

Another important aspect of the psychopath that may be reflected in their language is the process of psychological distancing. According to [11, p. 24], psychological distance “refers to the way in which the individual equilibrates and represents information for him- or herself.” The concept of psychological distance was influenced by Piaget’s theory of the stages of cognitive development in children, as it referred to children being able to separate objects from their physical appearances [12].

Psychological distancing can be measured through text analysis. When writing personal blog entries, compared to people with low psychological distance, people with high psychological distance use longer words, less present tense, and less first person singular pronouns [13]. We observed this kind of psychological distancing in our study reported above, in which offenders retold the story of their murders. In this study, psychopaths used more past tense verbs and fewer present tense verbs to psychologically distance themselves from the murder than the controls [8].

Will psychological distancing also be observed in psychopaths’ telling of autobiographical stories? As noted above, psychological distancing can be reflected in verb tense. If psychopaths are more able to engage in psychological distancing because of their emotional deficiencies, then they should use more past tense verbs when describing negative events but more present tense verbs during positive events relative to controls.

H3a: When describing negative events, psychopathic offenders will use more past tense and less present tense verbs relative to controls.

H3b: When describing positive events, psychopathic offenders will use less past tense (and more present tense) verbs relative to controls.

5. METHODS

5.1 Participants

The data in this study comes from interview transcripts of violent offenders in two Canadian maximum security correctional facilities: Mountain Institution in British Columbia and Springhill Institution in Nova Scotia. The transcripts were originally collected in 2000 for a previous study [7]. The selection criteria for the original study was that each subject had to have committed at least one homicide and needed to be able to provide detailed information about their crime.

The original study selected 74 offenders from Mountain Institution and 51 offenders from Springhill Institution, totaling 125 offenders. This study drew on the 51 offenders from the original study whose interviews were transcribed: 14 psychopathic and 38 non-psychopathic offenders. Psychopaths were classified as having a score of 25 or higher on the PCL-R.

5.2. Transcripts

Offenders were told that they were in a study about their memory, and they were asked to recount a positive experience, a negative experience, and their violent offense—a homicide. The present analysis focuses only on their descriptions of positive and negative experiences, and not on their description of their murder [8]. Positive experiences ranged from births, to marriages, to job tasks. Negative experiences consisted of non-homicidal crimes ranging from bank-holdups, kidnapping, and drug trafficking. Due to varying degrees of compliance with instructions, not all the prisoners provided both types of stories. There were 13 positive psychopathic stories, 11 negative psychopathic stories, 34 positive control stories, and 35 negative control stories.

Transcripts varied in the amount of detail provided. For example, one offender described a positive life event, a mechanics course, as “It was a lot of book study, a lot of on the job training,” and after probing by the interviewer to provide more detail, the offender told him “That about covers it.” Other offenders provided substantial detail, including
specific days, descriptions, and prefacing information. For example, one offender described a positive life event, meeting his father for the first time, and provided the interviewer with detailed information:

Oh, okay. My father left...or was removed from our family when I was 4. I went through a period of blaming myself for that...that I thought I had done something wrong, and he was punishing me. So, I went through a lot of my whole adult life, afterwards punishing myself. But in 1975, my grandmother...give me his address where he was. So, I went to live with him, unbeknownst to him. While - my grandmother and myself, and everyone else, we lived in [location], and my father was living up in [location] - a little remote community of about 3500 up in [location].

You know where [location] is?

5.3 Linguistic Analysis

Transcripts were analyzed using quantitative text analysis. Transcripts were analyzed for parts of speech categories using the Wmatrix corpus analysis and comparison tool [14, 15]. Wmatrix uses the Constituent Likelihood Automatic Word-tagging System, or CLAWS, to code for parts-of-speech (i.e. pronoun, verb, noun, etc.) based on surrounding linguistic context (e.g., “laugh” can be a verb or a noun depending on its surrounding context) [15]. CLAWS consistently achieves an accuracy rate of 96-97% in classifying parts-of-speech [15].

Wmatrix uses a corpus analysis approach, in which one corpus is compared to a second corpus. In this approach, one degree of freedom log likelihood ratios, calculated from contingency tables of pronouns or verb frequencies depending on hypothesis, were used to determine if the frequency of discourse patterns across the two corpora were significantly different.

A total of six corpora were prepared from the transcripts. The first one consisted of all the psychopath interviews and the second consisted of all the control interviews. The remaining four corpora split the these two overall corpora by prisoner type (psychopath vs. control) and story type (positive vs. negative): psychopaths speaking about positive events (N=13), psychopaths speaking about negative events (N=34), controls speaking about positive events (N=35), controls speaking about negative events (N=35).

The transcripts were analyzed in terms of pronoun and verb use, specifically the frequency of self-referencing pronouns, including first person singular (“I”, “me”) and plural (“we” “us”) pronouns, and the frequency of past and present tense verbs. Past tense verbs were defined as “were” “was” “been” “did” “done” “had,” past tenses of lexical verbs, such as “worked,” and past participles of lexical verbs, such as “given.” Present tense verbs were defined as “being, am, are, is, do, doing, does, having, has,” base forms of lexical verbs such as “give,” and –s forms of lexical verbs such as “works.”

As noted above, Wmatrix is highly accurate in classifying words into these pronoun and verb tense categories.

6. RESULTS AND DISCUSSION

6.1 Narcissism and Pronoun Analysis

6.1.1. First person singular analysis. Did the narcissistic nature of psychopaths lead them to use more language referencing themselves when retelling their stories? As expected in H1, psychopaths used more first person singular overall (freq = 1,391, relative freq = 7.00%) compared to controls (freq = 3,403, relative freq = 6.36%), LLR = 9.17 p < 0.01. This observation suggests that in general psychopaths used more first person singular than controls, consistent with the narcissistic component of psychopathy.

We were also interested in how the valence of the stories would affect first person pronoun production differentially for psychopaths compared to controls. As predicted in H2a, psychopaths used more first person singular pronouns (freq = 603, relative freq = 6.77%) during positive stories than controls (freq = 1416, relative freq = 5.84%) LLR = 8.92, p < 0.01. Failing to support H2b, when describing negative events, psychopathic stories (freq = 788, relative freq = 7.20%) did not differ significantly from control stories (freq = 1,987, relative freq = 6.78%), LLR=1.99, ns.

This pattern of results suggest that psychopaths used more first person singular on average when describing events in their lives compared to controls, but that this effect was driven by the difference observed in positive stories (see Figure 1). The results therefore support H2a, which predicted that because of their narcissistic nature psychopaths should use higher levels of self-reference when discussing positive events in an effort to associate themselves with the positive aspect of those stories. This proclivity to use more self-reference, however, disappeared for the negative stories, suggesting that
they did not want to associate themselves with the negative events.

Figure 1. 1st Person Singular Frequency by Story and Psychopathy

6.1.2. First person plural analysis. Did psychopaths use less first person plural pronouns (i.e. “we” and “us”) during their speech? Consistent with H1 and previous work on narcissism, psychopaths used less first person plural pronouns (freq = 162, relative freq = 0.82%) compared to controls overall (freq = 785, relative freq = 1.47%), LLR = 52.23, p < 0.0001.

More importantly, the valence of the story affected first person plural production differentially across psychopaths and controls. Psychopaths used more first person plural when describing negative events (freq = 108, relative freq = 1.07%) compared to positive events (freq = 54, relative freq = 0.61%), LLR = 8.95, p < 0.01. This difference did not obtain for controls, who produced statistically equivalent rates of first person plural across negative and positive stories, LLR = 2.63, ns.

These results suggest that while psychopaths use less first person plural pronouns overall when describing events in their lives, psychopaths tend to use more of these types of pronouns during negative stories compared to positive ones, unlike the controls (see Figure 2). The results supported H2a and H2b, which predicted that psychopaths would use first person plural pronouns to describe negative events in an effort to remove themselves from the event. Note that when considering the first person pronoun usage a clear pattern of narcissistic language emerges: compared with controls, psychopaths used singular self-referencing (“me” “I”) more when describing positive events in their past but used more plural self-referencing (“us” “we”) when describing negative events.

6.2. Psychological Distancing and Verb Tense Analysis

6.2.1 Past tense analysis. Our next set of analyses examined how frequently past and present tense verbs were used across the two types of stories. Consistent with the general concept of psychological distancing, in which people tend to place negative events farther back in time than positive events, offenders in the present study used more past tense during negative events (freq = 4,751, relative freq = 11.80%) than during positive events (freq = 3,303, relative freq = 9.96%), LLR = 56.36, p < 0.0001.

We predicted that psychopaths would engage in more psychological distancing than controls given their lack of emotional involvement. Overall, psychopaths (freq = 2,114, relative freq = 10.64%) did not differ from controls (freq = 5,940, relative freq = 11.09%) in their use of past tense verbs LLR = 2.69, ns. The question of interest, however, was how the valence of the story affected past tense verb production for psychopaths compared to controls. H3a predicted that during negative events, psychopaths would use more past tense verbs than controls. This hypothesis was not supported. Psychopaths (freq = 1,301, relative freq = 12.88%) and controls (freq = 3,450, relative freq = 12.55%) produced the same rate of past tense verbs during their telling of the negative stories, LLR = 0.08, ns.

A difference did emerge, however, during positive events. As predicted, psychopaths used less past tense (freq = 813, relative freq = 9.12%) than controls (freq = 2,490, relative freq = 11%) during positive events, LLR = 8.84, p < 0.01. The results supported H3b, which predicted that during positive events, psychopathic offenders would use less past tense verbs than controls. The results suggest that offenders use less past tense overall when telling positive stories, but the effect is more pronounced for
psychopaths when describing positive events (see Figure 4).

![Past Tense Verb Usage](image1)

**Figure 3. Past Tense Verb Frequency by Story and Psychopathy**

6.2.2 *Present tense analysis.* Consistent with the general psychological distancing concept, offenders used more present tense during positive events (*freq* = 1,930, *relative freq* = 5.82%) than during negative events (*freq* = 1,687, *relative freq* = 4.19%), *LLR* = 97.59, *p* < 0.0001.

Overall, psychopaths (*freq* = 1,012, *relative freq* = 5.10%) and controls (*freq* = 2,605, *relative freq* = 4.87%) did not differ in their use of present tense verbs, *LLR* = 1.54, ns. The key question, however, was whether the valence of the story affected present tense usage differently in psychopaths and controls. H3a predicted that during negative events, psychopaths would use less present tense verbs than controls. This hypothesis was not supported. During negative events, psychopaths (*freq* = 429, *relative freq* = 3.92%) and controls (*freq* = 1,258, *relative freq* = 4.58%) produced the same rate of present tense verb usage during the retelling of their negative stories, *LLR* = 2.7, ns.

Similar to the pattern found during past tense verb usage, however, a difference emerged for positive events. As predicted, psychopaths used more present tense (*freq* = 583, *relative freq* = 6.54%) than controls (*freq* = 1,347, *relative freq* = 5.96%), *LLR* = 10.56, *p* < 0.001. The results supported H3b, which predicted that when describing positive events, psychopathic offenders would use more present tense verbs than controls. The results suggest that offenders use more present tense when telling positive stories, but the effect is more salient for psychopaths (see Figure 4).

![Present Tense Verb Usage](image2)

**Figure 4. Past Tense Verb Frequency by Story and Psychopathy**

7. Study 1 Summary

The results from Study 1 indicate that discourse characteristics for psychopaths are detectable by automated linguistic analysis in theoretically predictable ways when describing autobiographical stories. These findings extend previous work from the more limited murder descriptions to story-telling.

Examining this more generalizable discourse we observed two major effects. First, psychopaths use self-reference differently from controls. Consistent with the narcissistic aspect of psychopathy, psychopaths used more first person singular overall, but especially for positive stories, suggesting that they tend to have a greater focus on the self than controls. This difference did not occur for negative stories, presumably because the psychopaths did not want to associate themselves with the negative aspects of the story, once again with a narcissistic personality.

Second, psychopaths were able to engage in more psychological distancing, consistent with their emotional detachment. This effect, however, was not entirely consistent across our predictions. We observed psychopaths using temporally closer language for positive events, placing positive events more recent in time than controls. But we did not see the expected effect with negative events, in which we expected psychopaths to place the event farther back in time with more past but fewer present tense verbs. It is not clear why this effect was not observed and requires additional analysis.

8. Study 2

While the previous study advances our ability to generalize the linguistic analysis of psychopathic speech beyond murder narratives to positive and negative autobiographical stories, the study is still
limited to an incarcerated population. Is it possible to extend this approach to the general population?

Until recently, not much research has been done in community samples because of the lack of a valid measure for large populations. The Psychopathy Checklist-Revised (PCL-R) is considered the “gold standard” of psychopathy measures”, but is only applicable to criminal or other in-depth cases because of its need for clinical interviews and subsequent life history related information [3]. However, the Self Report Psychopathy Scale III (SRP III) has proved to be valid along the four-factor structure of the PCL-R, and has been considered a valid and reliable self-report measure [4].

While the percentage of psychopaths in criminal populations is much higher than in the general population, that is not to say there are no high-functioning, non-criminal psychopaths. For example, Babiak, Neumann, & Hare [16] found that corporate professionals scoring high on the PCL-R were viewed as good communicators, strategic thinkers, and innovative by their immediate bosses.

Is it possible, then, to detect psychopathy in non-criminal populations using the autobiographical technique described in Study 1? To examine this question we asked students to complete the SRP-III and provide two stories, one positive and one negative. Here we report our preliminary analysis of this data.

9. METHODS

9.1 Participants

The participants in the study were 188 undergraduate students (140 Female) at the University of British Columbia Okanagan. The participants received course credit for compensation.

9.2 Procedure and Materials

After providing informed consent, participants were asked to write a short positive story and a short negative story in an online web form. After completing the stories, the participants completed the 64 item SRP III. Once they had completed this information they were asked a number of other demographic questions note reported here. Participants were debriefed and awarded credit at the end of the survey.

The 64 questions on the SRP III were ordered into the four facets of psychopathy to which they relate: callous affect, erratic lifestyle, interpersonal manipulation, and criminal tendency. Each participant was given a score for each facet and an overall score, giving each participant five scores.

The text from the positive and negative stories were analyzed using Linguistic Inquiry and Word Count (LIWC), a text analysis software program. LIWC identifies and categorizes words based on linguistic dimensions, psychological constructs, personal concern categories, and paralinguistic dimensions, among other output variables. LIWC outputs the percent of words that fall into each of its categories, including psychological (e.g., affect, cognitive words, etc.) and linguistic (e.g., pronouns, verbs, etc.). LIWC has been widely validated and reported in hundreds of studies.

10. PRELIMINARY RESULTS AND DISCUSSION

For this preliminary analysis we conducted correlations between the SRP scores and discourse categories identified in our previous studies, including pronouns, verb tense, affect, and material versus social categories of language (see [8]).

Here we briefly report the initial correlations with the various aspects of psychopathy provided by the SRP-III.

Overall there were very few correlations between the total SRP-III scores and the language categories examined. People scoring high on the SRP-III used fewer anxiety related words (r = -.90), suggesting that the emotional deficiency for psychopathy may be reflected in the lack of negative emotion language use. We did not observe the narcissism related pronoun correlations we expected, with the correlation between 1st person singular or plural not reaching significance.

When we separated the texts by story valence, however, we saw several predicted correlations for negative stories. First, the use of 1st person plural (“we”) was positively correlated with SRP-III scores (r = .19), consistent with our observations in Study 1 in which psychopaths used more plural self-reference when referring to negative events to disassociate the self with the event. Second, as expected, psychopathy was negatively associated with negative emotion words (r = -.1), especially anxiety terms (r = -.19), consistent with the emotional deficiencies noted above. No other correlations, including for verb tense, were observed, although the directions of the correlations were as anticipated. We did not observe any correlations for positive stories.

11. General Conclusions

The data from these two studies suggest some promising evidence that psychopathy can be detected in language using automated techniques. These observations, while preliminary, offer a potentially
important perspective for assessing credibility in a world increasingly dominated by text.

The first study revealed large effects between psychopaths and controls, which should perhaps be expected given that the study involved homicidal psychopaths. The effects were much more modest when taken out of the criminal population. Nonetheless, the data for study 2 require additional analysis that take more advantage of the valence of the stories. For instance, study 1 results suggest it will be important to look at difference scores between the positive and negative scores rather than simple correlations. Further, the SRP-III identifies four factors of psychopathy that should relate to different kinds of language. These analyses are now underway.

Despite these limitations, the present work suggests that the ability to discriminate credibility-related constructs, such as psychopathy, by analyzing text with automated tools is a promising approach that can advance both theory related to psychopathy and application development related to assessing credibility.

12. REFERENCES


