Patterns of source monitoring bias in incarcerated youths with and without conduct problems

Larisa Morosan[1,2], Deborah Badoud[1, 2], George Salaminios[3], Stephan Eliez [2], Martial Van der Linden [4], Patrick Heller [5], Martin Debbané[1, 2,4]

[1] Developmental Clinical Psychology Unit, Faculty of Psychology, University of Geneva, Switzerland

[2] Developmental Imaging and Psychopathology Lab, Office Médico-Pédagogique, Department of Psychiatry, University of Geneva, Switzerland

[3] Research Department of Clinical, Educational and Health Psychology, University College London, UK

[4] Psychopathology and Cognitive Neuropsychology Unit, University of Geneva, Switzerland

[5] Department of Forensic Medicine and Psychiatry, Department of communitarian, emergency and first aid medicine and psychiatry, Department of mental health and psychiatry, University Hospital of Geneva

Corresponding Author:

Larisa Morosan

University of Geneva,

40, Bd du Pont-d'Arve,

1211 Genève

Email: Larisa.Morosan@unige.ch

Tel.: + 41 (0)22 379 06 06

Authors' e-mail adresse :

Deborah Badoud:

Deborah.Badoud@unige.ch

George Salaminios:

g.salaminios@ucl.ac.uk

Stephan Eliez: stephan.eliez@etat.ge.ch

Martial Van der Linden:

Martial.VanDerLinden@unige.ch

Patrick Heller: Patrick.Heller@hcuge.ch

Martin Debbané:

Martin.Debbane@unige.ch

Keywords : adolescence, antisocial, self monitoring, self-awareness, delinquency

Abstract

Introduction: Antisocial individuals present behaviors that violate the social norms and the rights of others. In the present study, we examine whether biases in monitoring the self-generated cognitive material might be linked to antisocial manifestations during adolescence. We further examine the association with psychopathic traits and conduct problems (CP).

Methods: Sixty-five incarcerated adolescents (IA; *M*age= 15.85, *SD*=1.30) and 88 community adolescents (CA; *M* age=15.78, *SD*=1.60) participated in our study. In the IA group, 28 adolescents presented CP (*M*age=16.06, *SD*=1.41) and 19 did not meet the diagnostic criteria for CP (*M*age=15.97, *SD*=1.20). Source monitoring was assessed through a speech-monitoring task, using items requiring different levels of cognitive effort; recognition and source-monitoring bias scores (internalizing and externalizing biases) were calculated.

Results: Between group comparisons indicate greater overall biases and different patterns of biases in the source monitoring. IA participants manifest a greater externalizing bias, whereas CA participants present a greater internalizing bias. In addition, IA with CP present different patterns of item recognition.

Conclusions: These results indicate that the two groups of adolescents present different types of source-monitoring bias for self-generated speech. Future studies may examine the developmental implications of self-monitoring biases in the perseverance of antisocial behaviors from adolescence to adulthood.

Introduction

Antisocial behavior entails a range of violations to the moral and physical integrity or the property of others, and more broadly to social norms. These manifestations lead to a variety of research topics targeting phenomena such as aggression, behavioral disorders, and delinquency. In the field of child and adolescent psychiatry, diagnoses such as conduct disorder or oppositional defiant disorder are employed to describe different types of antisocial behaviors (Moffitt, Caspi, Harrington, & Milne, 2002). Several authors group these psychiatric diagnoses under the concept of conduct problems (CP) (Hill, 2002; Schwenck et al., 2014), which are more prevalent among incarcerated youth (Köhler, Heinzen, Hinrichs, & Huchzermeier, 2009), and appear to be associated with the development of life-course persistent antisocial behavior (Moffitt, 1993; Sevecke, Kosson, & Krischer, 2009). In addition, personality researchers have demonstrated that psychopathy, defined as the lack of affectivity, deceitful interpersonal style and impulsive and irresponsible behavior, may sustain antisocial manifestations (Andershed, Kerr, Stattin, & Levander, 2002). The developmental trajectories of antisocial individuals are marked by serious personal, social, and educational difficulties, and the damage resulting from their behaviors result in important costs for the society (Morgado & Vale-Dias, 2013).

For this purpose, a large body of research focuses on the psychological processes that might underlie antisocial manifestations. Some authors propose that antisocial individuals present impairments in the monitoring of their own actions (Bernat, Nelson, Steele, Gehring, & Patrick, 2011; Brazil et al., 2009; Hall, Bernat, & Patrick, 2007; Vilà-Balló, Hdez-Lafuente, Rostan, Cunillera, & Rodriguez-Fornells, 2014). For example, a series of studies using electroencephalography methodologies indicate that antisocial individuals present lower activation of the error related negativity, an indicator of action monitoring and error detection processes (Bernat et al., 2011; Brazil et al., 2009; Hall et al., 2007; Vilà-Balló et al., 2014).

These results suggest impairments in matching the expected outcome of their own actions to the actual outcome, leading to impairments in monitoring their own behaviors (Vilà-Balló et al., 2014).

A key cognitive process involved in the monitoring of one's behaviors is the ability to discriminate between different sources of information, traditionally studied within the source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993). Different types of source monitoring processes have previously been described: internal-external source monitoring, which enables one to distinguish between information generated by oneself from information generated by another person; external source monitoring, which refers to the ability to distinguish between two external sources; and internal source monitoring –distinguishing between what one imagined doing or saying from what one actually did or said (Johnson et al., 1993). Biases in the self-monitoring can arise as a result of several factors. The source-monitoring framework postulates that the amount and the clarity of sensorial signals (sensorial precision) biases towards an external attribution of the source. On the other hand, the amount and clarity of cognitive signals, such as thoughts, internal speech, imagination (cognitive precision) biases towards an internal attribution of the source of the material (Johnson et al., 1993).

The source monitoring framework can be informed by the forward model of motor control proposed by Miall and Wolpert, (1996). This model was initially developed to conceptualize the monitoring of actions, however, recent studies adapted it for the monitoring of thought content, such as internal speech (Frith, Blakemore, & Wolpert, 2000; Jones & Fernyhough, 2006). The forward model postulates that the correct attribution of the source results from a match between the predicted and the actual sensorial consequences of the action. On the other hand, a mismatch leads to biases in the attribution of the source. This mismatch might result due to interferences at different levels: in generating the prediction of the sensorial outcome

of the action or in the processing of the actual sensorial feedback of the action (Blakemore, Oakley, & Frith, 2003). Based on these two approaches, the source monitoring framework and the forward model, we can hypothesize that the impairments in the monitoring of behaviors presented by antisocial individuals might be explained by impairments in the source attribution. Considering the studies presented above describing impairments in processing the outcome of the action, we can hypothesize that the impairments in the source monitoring in antisocial individuals might be due to impairment in processing the sensorial feedback of their actions. This hypothesis comes in the continuity of several studies indicating impairments in sensorial integration in antisocial individuals (Assadi et al., 2007; Faruk, Demirel, Tayyib, & Emül, 2016; Lindberg, Tani, Stenberg, & Appelberg, 2004; Wang et al., 2016).

To the best of our knowledge, no study investigated the source monitoring of thought content in antisocial individuals. Thus, the present study focuses on self-generated speech monitoring, which represents a key component of internal source monitoring. We seek to explore the potential associations between the monitoring of self-generated speech and two crucial characteristics of antisocial individuals, CP and psychopathic traits. We focus on adolescence as a critical period for the development of antisocial tendencies (Frick & White, 2008). In addition, the investigation of source monitoring in a group of incarcerated adolescents may help identify early factors sustaining these maladaptive behaviors, and could further inform early prevention and intervention strategies.

For this purpose, we employ a task that examines the participant's capacity to discriminate between one's silently- and overtly produced speech. Previous studies indicate that the cognitive effort of the stimuli might play an important role in the monitoring of the source of the material (Debbané, Van der Linden, Glaser, & Eliez, 2010; Larøi, Van Der Linden, & Marczewski, 2004; Sugimori & Tanno, 2010). Thus, we manipulated the cognitive effort by

presenting different types of stimuli, words and non-words (Debbané et al., 2010). This task differentiates between two types of monitoring biases; the externalizing bias, which consists in reporting silently generated speech as overtly produced; and the internalizing bias, which consists in reporting overtly generated speech as silently produced. Based on previous studies investigating the monitoring of behaviors in antisocial individuals and on the postulates of the source-monitoring framework, we hypothesize that the incarcerated group will present a greater biases in the self-generated speech monitoring and that the bias will be greater for the items that require greater cognitive effort, the non-words. Furthermore, we aim to explore the relationship between source-monitoring of self-generated speech and psychopathic traits in both groups. In addition, we propose to investigate the differences in source monitoring, between the incarcerated adolescents with and without CP.

Method

Participants

Sixty-five adolescents incarcerated (IA) in an observation and detention center for youths in Geneva, Switzerland, took part in the study (*M*age=15.85, *SD*=1.30; 20 females). Eighty-eight community adolescents (CA) with no previous criminal convictions formed the comparison group (*M*age=15.78, *SD*=1.60; 30 females). The CA were recruited via advertising leaflets and by word of mouth and were tested at our research unit. The IA were individually tested at the center facility in a private room.

The inclusion criteria were age (12–18 years) and fluency in French. In addition, the subjects with a history of psychotic disorders and intellectual deficiency were not included in the study. For administrative reasons, information about the reason for incarceration was available for 60 of the IA; the majority committed more than one criminal offense, including physical and verbal aggression (16.7%), drug-related crimes (35%), theft and robbery (45%), runaways and risky behaviors (33.3%), conduct difficulties (20%) and driving violations

(8.3%). In relation to availability of the detained adolescents, forty-seven IA could be screened for psychiatric problems according to DSM-IV criteria using the Kiddie-SADS Present and Lifetime Version (K-SADS-PL) semi-structured interview (Kaufman et al., 1997). Trained clinical psychologists from our team conducted the interview under the supervision of MD. Diagnostic information is reported in Table 1. On the basis of the clinical interview, two IA subgroups were created: conduct problems (CP) group, which includes the twenty-eight IA who met the criteria for CD or ODD (7 females, *M*age=15.97, *SD*=1.20), and non-CP group, which includes the nineteen IA who did not meet any of the conduct problems diagnostic criteria (7 females, *M*age=16.06, *SD*=1.41).

All the participants completed the full protocol, except for one participant from the IA group who didn't complete the Youth Self-Report. Written informed consent was obtained from all the participants and, for participants under 18 years old, also from their legal guardians. The protocol was approved by the Institutional Review Board of the Department of Psychiatry of the University of Geneva Medical School. The adolescents in both groups received monetary compensation for their participation in the study.

Instruments

Source-monitoring task

Source-monitoring was investigated using a self-generated speech-monitoring task, the word/non-word task (Debbané et al., 2010). The task consisted of two parts: a reading procedure, followed by an incidental recognition and source monitoring procedure. In the first part of the task was presented as a reading and pronunciation exercise. The participants were required to read, either aloud or silently, a series of words (low cognitive effort items) or non-words (high cognitive effort items) presented on a computer screen. They were instructed to pay special attention to their pronunciation, even when reading the items silently; they were not informed that a recognition and source monitoring procedure would follow. After two

exercise trials, making sure that the subjects understood the task, six blocks (six silent, six aloud) of eight items (eight words, eight non-words) were randomly presented. In total, each condition contained 12 items, for a total of 48 items (12 words, 12 non-words read aloud and 12 words, 12 non-words silently read). After a 10–15 minutes visuospatial filler task, the second part of the task was introduced. A recognition sheet was then handed out, containing 72 items (the 48 items read in the first part of the task, plus 12 new word and 12 new non-word items). The participants were instructed that they have to indicate which items from the recognition list had appeared in the reading phase (yes/no- recognition test), and to attribute them to a reading condition (read silently or aloud- monitoring test). By using two types of items, the task aimed to differentiate the monitoring of self-generated speech in two different cognitive effort levels, high cognitive effort (non-words) and low cognitive load (words).

For the recognition phase, signal detection theory (Stanislaw & Todorov, 1999) was used to assess the sensibility for each type of items (word and non-words) for both reading conditions (aloud and silently). The estimation of d-prime scores were calculated by subtracting the *z* score corresponding to the false alarms from the *z* score corresponding to hit rate. False recognition scores were calculated as the number of words that were not presented in the reading phase of the task (distractors), misrecognized as belonging to the reading phase. Higher d-prime scores indicate a better recognition accuracy.

In order to assess the monitoring bias, externalizing and internalizing bias scores were calculated. The externalizing bias was calculated by dividing the total score for items read silently, but identified as read aloud in the monitoring test, out of the total score of items correctly recognized as read silently. In the same way, the internalizing bias score was calculated by dividing the total score of items read overtly but identified as read silently out of the total score of items correctly recognized as read overtly. Externalizing and internalizing

bias scores were calculated overall, as well as for each item type (word and non-word) separately.

Self-report questionnaires

Externalizing (including aggressive behaviors and rule-breaking behaviors) and internalizing (including withdrawal, anxiety, depression, and somatic complains) problems in participants aged <18 years were assessed using the Youth Self-Report (YSR; Achenbach, 1991).(Achenbach, 1991) For the participants aged ≥18 but < 19 years, the Adult Self-Report (ASR; Achenbach and Rescorla, 2003)(Achenbach & Rescorla, 2003) was used. Each of the 119 items in these instruments is evaluated on a 3-point scale, with 0 corresponding to "not true", 1 to "sometimes true" and 2 to "very or often true".

Psychopathic traits were assessed using the French version of the Youth Psychopathic Inventory (YPI; Andershed *et al.*, 2002). The YPI evaluates three dimensions of psychopathy, each consisting of several subscales: an interpersonal dimension assessing grandiose, manipulative behaviors, an affective dimension assessing callous-unemotional traits, and a dimension assessing impulsive, irresponsible behavior. The 50 items of the YPI are scored on 4-point scale, from 1 corresponding to "does not apply at all" to 4 corresponding to "applies very well".

In order to assess the cognitive functioning, we used the French versions of two subtests, Vocabulary and Digit Span, of the Wechsler Intelligence Scale for Children–Fourth edition (WISC; Wechsler, 2003) (Wechsler, 2003) and, for participants ≥18 years old, the Wechsler Adult Intelligence Scale–Third edition (WAIS; Wechsler, 1997).(Wechsler, 1997) The Vocabulary subtest measures word knowledge, language development, and concept understanding, whereas the Digit Span subtest investigates the short-term memory performances.

Statistical analysis

T-test analyses were conducted for sample characteristics, such as the age, WISC/ WAIS subscales, and YPI scores. Because the groups differ on WISC/WAIS subscales scores, and to control for the potential effect of gender, both variables were entered as covariates in the following analysis. For the self-monitoring task, mixed analysis of covariance (ANCOVA) was conducted on the d-prime scores for each type of stimuli (words vs. non-words) on each reading condition (aloud vs. silently), with group (IA vs. CA) as between factor. Mixed ANCOVAs were conducted on the monitoring bias scores (externalizing vs. internalizing), for each type of stimuli (word vs. non-word) and overall, with group (IA vs. CA) as between factor. Partial correlation analysis was used to investigate the relationships between the dependent variables and the sample characteristics, using gender as covariate.

In order to further explore the effects of the CP on the monitoring bias, we conducted the same analyses to compare the subgroups of IA with without CP. Because the two subgroups did not differ in the scores of the WISC/WAIS subscales, only gender was used as covariate. Statistical analyses were carried out in SPSS, version 23 for Mac (SPSS Inc.,USA).

Results

IA vs. CA groups

Sample characteristics

Table 2 presents the results for the *t*-test analyses of the sample characteristics. In comparison with the CA group, the IA group had significantly lower scores for the Vocabulary ($t(151)=5.71$, $p<0.001$, $d=0.92$) and Digit Span ($t(151)=2.21$, $p=0.020$, $d=0.36$) subtests, significantly higher scores for the externalizing subscale in the YSR/ASR ($t(150)= -8.10$, $p<0.001$, $d=1.32$), and significantly higher scores for the impulsive, irresponsible behavior subscale of the YPI ($t(146)=-6.14$, $p<0.001$, $d=1.01$) and the callous-unemotional subscale

198 (t(146)=-2.34, p=0.020, d=0.38). The groups did not differ in the mean age (t(151)=-0.27,
199 p=0.780, d=0.04), in the mean of the internalizing subscale of the YSR/ASR (t(151)=0.90,
200 p=0.360, d=0.14).

201 *Source-monitoring task results*

202 Results of mixed ANCOVA conducted on the d-prime scores, with group (IA vs. CA) as
203 between factor, and gender and both WISC/WAIS subscales scores as covariates, reveal a
204 main effect of reading condition (F(1, 148)=15.802, p<0.001, partial η^2=0.096) and a
205 significant effect of the item type (F(1, 148)=6.046, p=0.015, partial η^2=0.039) suggesting
206 that, independently of the group, silently read items and non-words are less accurately
207 recognized. The results indicate no interaction effect and no group effect (p>0.05). Table 3
208 presents the means and the standard deviations for the d-prime scores, for each type of item in
209 both reading conditions.

210 The results of mixed ANCOVA conducted on the monitoring bias scores for each type of
211 item, with group (IA vs. CA) as between factor, revealed a main effect of item type (F(1,
212 148)= 5.003, p=0.027, partial η^2=0.033), indicating that the monitoring bias is greater for the
213 non-words. In addition, the results demonstrate a significant interaction effect between the
214 monitoring bias and the type of item (F(1, 148)= 4.85, p=0.029, partial η^2=0.032), suggesting
215 that, independently of the group, the monitoring bias affects differently the type of items. To
216 follow up this interaction effect, simple effects were analyzed, revealing that, independently
217 of group, there was a greater internalization bias for non-words than for words (F(1,
218 148)=12.628, p=0.001, partial η^2=0.064), and greater internalization bias than externalization
219 bias for non-words (F(1, 148)=5.211, p=0.024, partial η^2=0.034). The results also indicate a
220 main effect of the group (F(1, 148)=5.356, p=0.026, partial η^2=0.026), suggesting that the IA
221 group present more monitoring bias independently of the item type and bias.

Finally, a significant triple interaction effect between the group, the monitoring bias, and the item type ($F(1, 148)= 8.50$, $p=0.004$, partial $\eta^2=0.054$) has been found, suggesting that the interaction between the monitoring bias and the type of items was different in the two groups of participants. Simple interaction effects were analyzed, indicating that, relative to the CA group, the IA presented a significantly greater externalizing bias for non-words ($F(1, 148)=10.120$, $p=0.002$, partial $\eta^2=0.064$), and significantly greater internalizing bias for the words ($F(1, 148)=5.088$, $p=0.026$, partial $\eta^2=0.033$). In addition, the results suggested that the CA presented greater internalizing bias for non-words than for words ($F(1, 148)=16.017$, $p<0.001$, partial $\eta^2=0.098$). These results are presented in the Figure 1.

Correlation analysis

We conducted partial correlations on the source monitoring results and YPI subscales, with gender as covariate. For the d-prime scores in the recognition phase, no result exceeded the significance level ($p>0.65$). For the monitoring bias scores, in the CA group, the results did not show any significant result ($p>0.196$). In the IA group, the results reveal that the internalizing bias for non-words was negatively correlated with interpersonal problems subscale of the YPI ($r= -0.277$, $p= 0.030$). After the Bonferroni correction, no correlation reached the significance level ($p=0.004$).

CP vs. non-CP groups

Sample characteristics

The results of *t*-test analysis indicated that, relative to the non-CP group, the CP group presented higher scores for the externalizing subscale of YSR/ASR ($t(44)=3.214$, $p=0.002$, $d=0.96$). The two groups did not differ in the mean age ($t(45)=-0.237$, $p=0.813$, $d=0.07$), nor in mean scores for the Vocabulary ($t(45)=-0.698$, $p=0.489$, $d=0.20$) and for the Digit Span

($t(45)=0.507$, $p=0.615$, $d=0.15$) subtests. In addition, there was no difference between the groups in the subscales of the YPI ($p>0.375$).

Source-monitoring task results

The mixed ANCOVA conducted on the d-prime scores, with group (CP vs. non-CP) as between group factor and gender as covariate, revealed a significant effect of the reading condition ($F(1, 44)= 9.959$, $p=0.004$, partial $\eta^2=0.178$), suggesting that independently of the group, items read silently were less accurately recognized. In addition, the results suggested a triple interaction effect between the group, the reading condition, and the type of item ($F(1, 44)= 2.294$, $p=0.031$, partial $\eta^2=0.104$), suggesting that the relation between the reading condition and the item type differs across the groups. To follow up the interaction effect, simple effects were analyzed, revealing that the non-CP group showed a less accurate recognition for the words read silently in comparison with the words read overly ($F(1, 44)= 16.036$, $p<0.001$, partial $\eta^2=0.272$). The same pattern was observed for the CP group ($F(1, 44)=4.981$, $p=0.031$, partial $\eta^2=0.104$). The CP group also showed a less accurate recognition for the non-words read silently, compared to the non-words read overly ($F(1, 44)=13.373$, $p=0.001$, partial $\eta^2=0.237$). These results are presented in the Figure 2.

The results of mixed ANCOVA conducted on the monitoring bias scores for each type of item, with group (CP vs. non-CP group) as between factor and gender as covariate, did not reveal any significant effect ($p>0.145$).

Discussion

The present study investigated self-monitoring performances in a group of incarcerated adolescents (IA), in comparison to a group of community adolescents (CA). We employed a task assessing source monitoring of self-generated speech, which included stimuli of different levels of cognitive effort (words-low effort; non-words-high effort). The task yields a

recognition score (d prime) and two self-monitoring bias scores, internalizing bias score, defined as the tendency to identify overtly read items as silently read, and externalizing bias score, defined as the tendency to identify silently read items as overtly read. In light of the relevant literature, the results will be discussed in relation to the psychological and clinical characteristics of each group.

IA vs. CA groups

Firstly, no group differences were found for the d prime scores, suggesting that IA participants conserve intact recognition capacities. Regarding the source monitoring bias scores, the IA presented more overall biases in comparison to the CA group. Biases in the source monitoring might be explained by impairments in the integration of contextual information into a coherent whole and impairments in the integration of sensory information previously reported in antisocial individuals (Assadi et al., 2007; Faruk et al., 2016; Hamilton, Racer, & Newman, 2015; Lindberg et al., 2004; Wang et al., 2016).

Interestingly, the results suggest that the two adolescent groups present different patterns of monitoring bias, depending on the cognitive effort required by the material. Firstly, the IA group shows a greater externalizing bias for non-words, compared to the CA group. The source monitoring framework states that external attributions are more probable for stimuli with increased sensorial precision (Johnson et al., 1993). We may hypothesize that, during the reading phase, the IA used more sensorimotor information such as subvocalizations and failed to generate the kind of cognitive information that controls generated while reading the non-words. For instance, upon reading a non-word item like "TEVU", the CA group may have generated idiosyncratic cognitive information (for example, one may think, "*that is like T-View*"). In addition, accordingly to the forward model, we might explain the externalizing bias for non-words as a mismatch between the predicted and the actual sensors feedback (Blakemore et al., 2003). This mismatch could be due to impairments in the processing of the

actual feedback of an action (here the silently reading of the non-words). This explanation is in line with previous studies which indicate that antisocial individuals present impairments in processing the sensory feedback of their actions (Hall et al., 2007; Vilà-Balló et al., 2014), especially internally generated feedback (Bernat et al., 2012).

Contrary to non-words, word items were associated to a greater internalizing bias in the IA group. We believe that the words represented items that were familiar to participants, and in the IA group, familiarity may have reduced the encoding of sensory-perceptual properties of the material. The source monitoring framework suggests that weak sensory-perceptual precision engenders uncertainty about the "realness" of the items, which will therefore more likely be attributed to an internal source (Johnson et al., 1993).

Contrary to the IA group, the effect of cognitive load leads to a greater tendency to internalize overtly read non-words in the CA group. We may hypothesize that the CA group generated more cognitive operations to encode the overtly read non-words, to the detriment of sensorimotor evidence (production and sound of speech) that would have assisted in correct source monitoring.

Regarding the relationship between psychopathic traits and source monitoring capacities, our results did not reveal any association in either of the groups. This could be to the lack of discriminative power of the self-report measures investigating the psychopathic dimensions, which may be more thoroughly assessed through semi-structured interviews.

Within group analysis of the impact of conduct problems (CP) in IA group on the source monitoring abilities

The results indicate that recognition scores differed between CP and non-CP groups, depending on the item type and reading condition, the CP group showing a better recognition for the non-words read aloud than for the non-words read silently. These results may suggest that the CP group encodes items by favoring sensorimotor and perceptual information, which

may yield a recognition advantage (Johnson et al., 1993). The fact that the IA with CP recognize better the non-words read aloud than those read silently might indicate that they rely more on perceptual information in the processing of the items requiring a greater amount of cognitive effort. This result may have important implications for clinical practice, in that the source monitoring profile may divulge information to be considered within the assessment procedure and treatment plan for IA presenting CP. However, the results did not indicate any difference between the groups in monitoring bias (p>0.168).

Limitations

Some limitations of the present study should be taken into consideration. First, the assessment of psychopathic traits was performed using a self-report questionnaire, which should be complemented with a semi-structured interview measure to fully assess the links between psychopathy and source monitoring. Another limitation is that not all the adolescents completed clinical interview, thus we could not compare the two groups regarding their clinical characteristics.

Conclusion

To the best of our knowledge, this study is the first to investigate source monitoring in delinquent adolescents. Using a self-generated speech monitoring paradigm, we observe preserved recognition performances, but impairments in the source-monitoring in the IA group. More precisely, the IA showed an increased externalizing bias when monitoring cognitively effortful items, as well as increased internal biases when monitoring familiar items. In addition, we observe that conduct problems in IA individuals may worsen their recognition performances.

We propose that impairments in the source-monitoring abilities might contribute to limited self-awareness, but also to limit insight about one's own actions and their consequences. These impairments might lead to an inability to learn from their experiences and to correct

their behaviors (Vilà-Balló et al., 2014). In addition, these impairments might contribute to a more general tendency of the antisocial individual to experience their thoughts as real, manifesting an equivalence between internal and external reality (Bateman & Fonagy, 2016). These characteristics might interfere with the motivation to change and hinder psychosocial and therapeutic strategies. The present results warrant future research among IA, exploring the relations between impairments in the monitoring of self-generated material and the lack of insight about their behaviors, as well as the lack of responsibility for their actions.

Reference

Achenbach, T. M. (1991). *Manual for the Youth Self-Report and 1991 profile*. Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T. M., & Rescorla, L. A. (2003). *Manual for the ASEBA adult forms and profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.

Andershed, H., Kerr, M., Stattin, H., & Levander, S. (2002). Psychopathic traits in non-referred youths: A new assessment tool. In E. Blaauw & L. Sheridan (Eds.), *Psychopaths: Current international perspectives* (pp. 131–158). The Hague, The Netherlands: Elsevier.

Assadi, S. M., Noroozian, M., Shariat, S. V., Yahyazadeh, O., Pakravannejad, M., & Aghayan, S. (2007). Neurological Soft Signs in Mentally Disordered Offenders. *Journal of Neuropsychology and Clinical Neuroscience*, *19*, 420–427.

Bateman, A., & Fonagy, P. (2016). *Mentalization-based treatment for personality disorders*. Oxford: Oxford University Press.

Bernat, E. M., Nelson, L. D., Steele, V. R., Gehring, W. J., & Patrick, C. J. (2011). Externalizing Psychopathology and Gain/Loss Feedback in a Simulated Gambling Task: Dissociable Components of Brain Response Revealed by Time-Frequency Analysis. *Journal of Abnormal Psychology*, *29*, 997–1003. http://doi.org/10.1016/j.biotechadv.2011.08.021.Secreted

Blakemore, S., Oakley, D. A., & Frith, C. D. (2003). Delusions of alien control in the normal brain. *Neuropsychologia*, *41*, 1058–1067. http://doi.org/10.1016/S0028-3932(02)00313-5

Brazil, I. A., Bruijn, E. R. A. De, Bulten, B. H., Borries, A. K. L. Von, Lankveld, J. J. D. M. Van, Buitelaar, J. K., & Verkes, R. J. (2009). Early and Late Components of Error Monitoring in Violent Offenders with Psychopathy. *Biological Psychiatry*, *65*(2), 137–143. http://doi.org/10.1016/j.biopsych.2008.08.011

Debbané, M., Van der Linden, M., Glaser, B., & Eliez, S. (2010). Monitoring of self-generated speech in adolescents with 22q11.2 deletion syndrome. *British Journal of Clinical Psychology*, *49*(3), 373–386. http://doi.org/10.1348/014466509x468223

Faruk, O., Demirel, A., Tayyib, M., & Emül, M. (2016). Neurological soft signs in antisocial men and relation with psychopathy. *Psychiatry Research*, *240*, 248–252. http://doi.org/10.1016/j.psychres.2016.04.094

Frick, P. J., & White, S. F. (2008). Research review: The importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. *Journal of Child Psychology and Psychiatry*, *49*(4), 359–375. http://doi.org/10.1111/j.1469-7610.2007.01862.x

Frith, C. D., Blakemore, S., & Wolpert, D. M. (2000). Explaining the symptoms of schizophrenia : Abnormalities in the awareness of action. *Brain Research Reviews*, (31), 357–363.

Hall, J. R., Bernat, E. M., & Patrick, C. J. (2007). Externalizing psychopathology and the error-related negativity. *Psychological Science*, *18*(4), 326–333. http://doi.org/10.1111/j.1467-9280.2007.01899.x

Hamilton, R. K. B., Racer, H. K., & Newman, J. P. (2015). Impaired integration in psychopathy: A unified theory of psychopathic dysfunction. *Psychological Review*, *122*(4), 770–791. http://doi.org/10.1037/a0039703

Hill, J. (2002). Biological, psychological and social processes in the conduct disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *43*(1), 133–164. http://doi.org/10.1111/1469-7610.00007

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. http://doi.org/10.1037/0033-2909.114.1.3

Jones, S. R., & Fernyhough, C. (2006). Thought as action : Inner speech , self-monitoring , and auditory verbal hallucinations. http://doi.org/10.1016/j.concog.2005.12.003

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., … Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*(7), 980–988. http://doi.org/10.1097/00004583-199707000-00021

Köhler, D., Heinzen, H., Hinrichs, G., & Huchzermeier, C. (2009). The prevalence of mental disorders in a German sample of male incarcerated juvenile offenders. *International Journal of Offender Therapy and Comparative Criminology*, *53*(2), 211–227.

Larøi, F., Van Der Linden, M., & Marczewski, P. (2004). The effects of emotional salience, cognitive effort and meta-cognitive beliefs on a reality monitoring task in hallucination-prone subjects. *British Journal of Clinical Psychology*, *43*, 221–233. http://doi.org/10.1348/0144665031752970

Lindberg, N., Tani, P., Stenberg, J., & Appelberg, B. (2004). Neurological soft signs in homicidal men with antisocial personality disorder. *European Psychiatry*, *19*, 433–437. http://doi.org/10.1016/j.eurpsy.2004.05.011

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Network*, *9*, 1265–1279.

Moffitt, T. E. (1993). Adolesence-limited and life-course-persistent antisocial behavior:a developmnetal taxonomy. *Psychological Review*, *100*, 674–701. Retrieved from http://graphics.tx.ovid.com/ovftpdfs/FPDDNCDCJGMHPK00/fs047/ovft/live/gv024/00006832/00006832-199310000-00006.pdf

Moffitt, T. E., Caspi, A., Harrington, H., & Milne, B. J. (2002). Males on the life-course-persistent and adolescence-limited antisocial pathways: follow-up at age 26 years. *Development and Psychopathology*, *14*(1), 179–207. http://doi.org/10.1017/S0954579402001104

Morgado, A. M., & Vale-Dias, M. D. L. (2013). The antisocial phenomenon in adolescence: What is literature telling us? *Aggression and Violent Behavior*, *18*(4), 436–443. http://doi.org/10.1016/j.avb.2013.05.004

Schwenck, C., Gensthaler, A., Romanos, M., Freitag, C. M., Schneider, W., & Taurines, R. (2014). Emotion recognition in girls with conduct problems. *European Child & Adolescent Psychiatry*, *23*(1), 13–22. http://doi.org/10.1007/s00787-013-0416-8

Sevecke, K., Kosson, D., & Krischer, M. (2009). The Relationship Between Attention Deficit Hyperactivity Disorder, Conduct Disorder, and Psychopythy in Adolescent Male and Female Detainees. *Behavioral Sciences & the Law*, *28*(2), 577–598. http://doi.org/10.1002/bsl

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. http://doi.org/10.3758/BF03207704

Sugimori, E., & Tanno, Y. (2010). The effects of cognitive activity and perceptual details on speech source monitoring. *British Journal of Psychology*, *101*, 777–90. http://doi.org/10.1348/000712610X485727

Vilà-Balló, A., Hdez-Lafuente, P., Rostan, C., Cunillera, T., & Rodriguez-Fornells, A. (2014). Neurophysiological correlates of error monitoring and inhibitory processing in juvenile violent offenders. *Biological Psychology*, *102*(1), 141–152. http://doi.org/10.1016/j.biopsycho.2014.07.021

Wang, X., Cai, L., Li, L., Yang, Y., Yao, S., & Zhu, X. (2016). Neurological soft signs in Chinese adolescents with antisocial personality traits. *Psychiatry Research*, *243*, 143–146. http://doi.org/10.1016/j.psychres.2016.04.023

Wechsler, D. (1997). *Manual of the Wechsler Adult Intelligence Scale–third edition (WAIS-III)*. New York, NY: Psychological Corporation.

Wechsler, D. (2003). *WISC-IV technical and interpretive manual*. San Antonio, TX: Psychological Corporation.

Table 1. Diagnostic information for the incarcerated group of adolescents.

K-SADS-PL diagnostics	**% of the group**
Substance abuse	10.6 %
Conduct disorder	14.8 %
Conduct disorder and substance abuse	34 %
Anxiety disorder	8.5 %
Conduct disorder, substance abuse and other diagnosis (MDD, ADHD)	10.6 %
No diagnosis	21.2 %

K-SADS-PL- Kiddie-SADS Present and Lifetime Version; ADHD- Attention-deficit/hyperactivity disorder, MDD- Major depressive disorder

Table 2. Sample characteristics result

WISC Wechsler Intelligence Scale for Children, *WAIS* Wechsler Adult Intelligence Scale,

	Community adolescents		*Incarcerated adolescents*	
	Mean	*SD*	*Mean*	*SD*
Age (years)	15.78	1.60	15.85	1.30
WISC/WAIS (Vocabulary)	10.90 **	3.53	7.78**	3.06
WISC/ WAIS (Digit Span)	9.04 *	2.76	8.09*	2.44
YSR (externalizing)	56.70 **	9.57	68.54 **	7.87
YSR (internalizing)	54.12	10.22	52.60	10.06
YPI (impulsive-irresponsible subscale)	23.83**	5.38	29.25**	5.19
YPI (CU subscale)	29.48*	5.56	32.07*	7.93
YPI (interpersonal problems subscale)	28.12	8.74	31.09	10.27

YSR Youth Self Report; *YPI* Youth psychopathic Inventory

**p<0.01

*p<0.05

Table 3. Means and standard deviations for the d prime scores in the recognition phase, of both groups of adolescents

		Community adolescents		*Incarcerated adolescents*	
Reading condition	*Item Type*	*Mean*	*SD*	*Mean*	*SD*
Overtly	Words	2.06	0.65	1.78	0.70
	Non-Words	1.77	0.71	1.55	0.63
Silently	Words	1.27	0.63	1.22	0.53
	Non-Words	1.42	0.74	1.16	0.65