Does Cognitive Inflexibility Predict Violent Extremist Behaviour Intentions? A Registered Direct Replication Report of Zmigrod et al., 2019

Purpose: Research assessing violent extremist risk factors thus far largely ignored the role of cognitive processes. Zmigrod and colleagues (2019a) addressed this gap and presented first systematic evidence that lower levels of cognitive flexibility predict a higher willingness to fight and, ultimately, die for a national ingroup. This finding has important theoretical and practical implications. In order to strengthen the potential contribution of Zmigrod et al.'s work, we will conduct a registered direct replication of Study 1. Extending the original study, we further examine whether the documented relationship still holds when a self-report measure for cognitive flexibility is introduced and when analyses control for identity fusion. We also investigate if cognitive inflexibility solely predicts violent or also normative pro-group behaviour intentions.

Methods: Following Zmigrod et al. (2019a), we will administer a cross-sectional survey study. Participants ($N = 1333$) report their willingness to fight, die, and sacrifice themselves for the ingroup and complete the Remote Associates as well as Wisconsin Card Sorting tests. Afterwards, additional measures of self-reported cognitive flexibility, identity fusion, and normative pro-group behaviour are assessed.

Results: tbc

Conclusions: tbc

Keywords: cognitive flexibility, extremist attitudes, violent extremism, direct replication

Does Cognitive Inflexibility Predict Violent Extremist Behaviour Intentions?: A Registered Direct Replication Report of Zmigrod et al., 2019

A burgeoning body of research highlights numerous violent extremist risk factors that point, amongst others, to the role of exposure to radicalising actors, critical life events, and grievances (e.g., Bouhana, 2019; Jasko et al., 2017; Lösel et al., 2018; Obaidi et al., 2019; Perry et al., 2018; Vergani et al., 2018). Despite the breadth of this literature, the influence of cognitive processes - "mental functions … involved in the acquisition, storage, interpretation, manipulation, transformation, and use of knowledge" (American Psychological Association, 2020) - has, to date, been largely ignored. Zmigrod, Rentfrow, and Robbins (2019a) addressed this gap. The authors presented two cross-sectional studies, conducted in the United Kingdom and the United States, which showed that lower levels of cognitive flexibility are related to a stronger willingness to fight and die for the national ingroup.

In a field that is characterised as lacking sufficient empirical data (Schuurman, 2018), Zmigrod et al.'s (2019a) work has the potential to advance existing lines of inquiry and inform interventions to prevent and counter violent extremism (Stephens et al., 2019). Indeed, in the year since it was published the paper has been cited already 16 times (Google Scholar citations; 07.11.20). To our knowledge, however, Zmigrod et al.'s (2019a) findings have not yet been replicated by other researchers. Previous research that focused on the relationship between cognitive flexibility and ideological extremism (e.g., Van Hiel et al., 2010; Van Hiel & Mervielde, 2003) as well as two studies that assessed cognitive flexibility within terrorist samples (Alkhadher & Scull, 2019; Baele, 2017), in fact, obtained inconsistent results. To strengthen the contribution that the paper can make, and to address some of its limitations, we will conduct a registered direct replication of Zmigrod et al. (2019a), Study 1.

Following the same protocol as the original study, we aim to demonstrate a negative relationship between cognitive flexibility and willingness to fight for the national ingroup,

which, in turn, should predict intentions to die for the group (i.e., support for violent extremism). We will complement the original study by examining whether result patterns still hold when a self-report measure of cognitive flexibility is used, when identity fusion is introduced as a control variable, and when stricter exclusion criteria are applied. Moreover, we seek to assess the extent to which cognitive flexibility predicts only violent or as well normative pro-group behaviour intentions.

Cognitive Flexibility and Extremism

Despite the proliferation of the literature on violent extremism in recent decades (Schuurman, 2018), implications of cognitive processes have thus far been widely neglected (for an exception see Bouhana, 2019). This is surprising, given that early work investigating predictors of *ideological extremism* focused on cognitive style, specifically cognitive flexibility. Notably, Adorno, Frenkel-Brunswik, Levinson, and Sanford (1950) showed that participants in the lower and upper quartiles on a measure of ethnocentrism were also described as less flexible in their thinking (see as well Rokeach, 1948). Cognitive flexibility is defined as the ability to adapt ways of thinking to changing environmental conditions (Dennis & Vander Wal, 2010); it is an executive function that relies on the processes of salience detection and attention, working memory, inhibition, and set switching (Dajani & Uddin, 2015). High cognitive flexibility may, for instance, be expressed by being able to adjust previously established habits to a new situation that has different demands (Moore & Malinowski, 2009) or by considering multiple aspects of an idea - such as rules - at the same time (Scott, 1962). To date, it has not been definitely concluded whether cognitive flexibility is a general executive function, which is stable over time and applies to different content, or whether it is task and domain-specific (see Deak & Wisehart, 2015).

Contemporary research supports the claim that cognitive flexibility is related to ideological extremism, in particular right-wing extremism (Zmigrod, 2020). Lower levels of cognitive flexibility have been found to predict stronger right-wing attitudes (Van Hiel et al., 2016; Van Hiel et al., 2010), political conservatism (Jost et al., 2003; Sidanius, 1985), racism (Sidanius, 1985), nationalism and authoritarianism (Zmigrod et al., 2018). Zmigrod, Rentfrow, and Robbins (2019b) also showed that those who reported being on either the extreme left or the extreme right of the political spectrum exhibited (compared to other, moderate participants) lower accuracy rates on tests that examined cognitive flexibility. That is, as documented by Ardono and colleagues (1950), not partisanship but attitude extremity was related to stronger cognitive inflexibility (*extremism theory*; Sidanius, 1985).

Despite this compelling evidence, the literature on cognitive style and extremism is far from unanimous. Van Hiel and Mervielde (2003) demonstrated in three studies a positive - rather than negative - relationship between cognitive flexibility and ideological extremism. These results are in line with context theory (Sidanius, 1976; 1985), which argues that moderates are less flexible in their cognitive style, as they have acquired their beliefs by simply conforming to the majority. Those who endorse extreme beliefs, however, would have arrived at these through active processing and incorporating controversial ideas; they should, hence, be more cognitively flexible. Evidence documenting the relationship between cognitive flexibility and violent - rather than ideological - extremism is also inconsistent. Bouhana's (2019) S5 model proposes that cognitive processes, including flexibility, contribute to individuals' susceptibility to moral change that increases the likelihood to adopt a violent extremist ideology. A study of 30 imprisoned ISIS and Al-Qaeda members in Kuwait highlighted such a positive correlation (Alkhadher & Scull, 2019). However, analyses of writings of 11 lone-actor terrorists, using tentative and contrary language as measures of cognitive flexibility, did not differ significantly from the writings of individuals who actively opposed

the use of violence (Baele, 2017). Both of these studies including terrorist samples suffer from limitations. The authors sampled on the outcome variable 'violent extremism', which reduces the variability of the dependent measure. In addition, the measures for cognitive flexibility are arguably less than ideal; the studies rely on informant-reports regarding the prisoners' difficulty to change their mindset and linguistic analysis of writings that may have been affected by audience tuning.

The Original Study: Zmigrod et al., 2019a, Study 1

Addressing these challenges, Study 1 in Zmigrod et al., 2019a, provided the first systematic empirical evidence for the relationship between cognitive flexibility and support for violent extremism. British citizens (N = 305; M_{age} = 38.02, SD_{age} = 13.51, range = 18–72; 47% female; 91% white) were recruited through the online opt-in access panel Prolific Academic. The authors operationalised cognitive flexibility as the extent to which individuals organise information in looser or more rigid semantic networks (Remote Associates Test; RAT; Mednick, 1968) and as the ability to identify, adopt, and then change rules to categorise information (Wisconsin Card Sorting Test; WCST; Grant & Berg, 1948). Support for violent extremism was assessed with two dependent variables: willingness to die on behalf of the national ingroup, as well as willingness to sacrifice one's own life to save five British people (trolley problem). Furthermore, the authors examined willingness to fight for the national ingroup and how certain participants were that they would indeed self-sacrifice. More detailed information about the measures is presented below.

The results showed negative correlations between RAT (r = -.241, p < .001) as well as WCST (r = -.216, p = .002) accuracy rates and willingness to fight for the ingroup. Moreover, RAT (r = -.207, p = .001), but not WCST (r = -.039, p = .587), accuracy rates were correlated negatively with willingness to die for the ingroup. Findings partially supported an indirect relation between cognitive flexibility and willingness to die. RAT (β = -.183, p = .005),

but not WCST (β = -.121, p = .103), accuracy rates predicted lower willingness to fight (R^2 = .145), which, in turn, was positively associated with willingness to die for the ingroup (β = .534, p = .000; R^2 = .335). Willingness to fight for the ingroup, however, was not related to participants' decisions in the trolley problem, such that an indirect effect could not be assessed for this alternative outcome variable. Individuals who indicated in the trolley dilemma that they would self-sacrifice rather than save themselves also did not differ in their RAT and WCST accuracy rates (WCST: F(1, 205) = .098, p = .754, RAT: F(1, 294) = .815, p = .367). Sub-group analysis showed that for participants who were willing to sacrifice themselves to save the ingroup, greater confidence in the decision to self-sacrifice correlated negatively with WCST (r = .333, p = .011) and RAT (r = .217, p = .034) accuracy rates.

Proposing a Direct Replication

The aforementioned findings have important theoretical and practical implications. As the review of existing literature suggests, Zmigrod et al.'s (2019a) work can serve as a much-needed stepping stone towards new lines of inquiry that incorporate cognitive style as a potential risk factor of violent extremism. The studies also introduce methods from cognitive psychology - three tasks to objectively measure cognitive flexibility - that are currently not commonly applied in research on violent extremism. Crucially, the results offer relevant insights for soft power interventions that aim to prevent and counter violent extremism (PVE and CVE). One approach to PVE and CVE emphasises the importance of cognitive resources that help build resilience and reduce the risk of being attracted to or convinced by violent extremist messages (Stephens et al., 2019). To date, cognitive complexity - overcoming "us-versus-them" and "black and white" thinking (Liht & Savage, 2013; Savage et al., 2014) - as well as critical thinking (Aly et al., 2014; Davies, 2016) are trained in educational or community settings. If evidence of the role of cognitive flexibility is strengthened, programmes

could be extended to include relevant activities, such as serious gaming (Boendermaker et al., 2017), that improve adaptive decision making and thinking (Mun et al., 2017).

It is evident that Zmigrod and colleagues' paper (2019a) has the potential to impact both theory and practice. However, some concerns about the rationale, design, and results of the original study need to be explored further. First, the authors base their work primarily on previous (inconsistent) research that has focused on ideological extremism. It is not discussed how these insights apply to the study of violent behaviour intentions, fighting and self-sacrificing for the ingroup. Moghaddam (2005) proposed that adopting a certain cognitive style is one of the final steps on the 'staircase' to committing a terrorist attack. But for someone to execute an attack, they must also engage with the morality of the terrorist group, categorise in- and outgroup members, and distance themselves from the outgroup (i.e., dehumanisation; see Tausch et al., 2011). Thus, the relationship between cognitive flexibility and self-sacrifice could be more complex - and weaker - than when considering prejudice or conservatism as outcome measures.

Second, partially inconsistent result patterns were presented in Zmigrod et al., (2019a) Study 1. As described, the RAT but not the WCST accuracy rates were related to willingness to fight and die for the ingroup. The authors pointed to the moderate correlation between both measures ($r = .21$, $p = .009$) and reported an analysis that examined the unique role of WCST accuracy rates. The latter, however, failed to show a statistically significant association ($p = .051$). Further, two of the dependent variables - a one-item self-report measure and the trolley dilemma - examine willingness to die for an ingroup. Although RAT accuracy rates are correlated with the one-item outcome measure, those who reported in the trolley dilemma that they would sacrifice their life to save ingroup members did not differ from those who would not self-sacrifice in terms of their RAT (and WCST) accuracy rates. Finally, considering findings

of Study 2, which may be seen as a replication on Study 1, not all result patterns were replicated. RAT accuracy rates correlated with the willingness to die in the UK but not the US sample.

Lastly, to measure support for violent extremism, Zmigrod and colleagues used Swann and colleagues' (2009) scale, which examines willingness to fight and die for an ingroup. These two dimensions - fighting and dying - were separated such that the latter represents the outcome and the former the proposed mediator variable. Initial papers by Swann and colleagues also applied the two measures separately (e.g., Swann et al., 2009). However, in more recent publications, the two dimensions were combined into one measure, "(b)ecause the measures of willingness to fight and die are conceptually overlapping and highly correlated" (p. 828); the concept was referred to as *endorsement of extreme actions for the group* (Gomez et al., 2011; Swann et al., 2010). In their analysis, Zmigrod et al., (2019a) did not present a factor analysis of the measures capturing willingness to fight and die. Without knowing whether measures for the mediator and outcome load on the same factor, which could perhaps be expected, it is difficult to establish whether the indirect effects that were examined should indeed be computed.

The Present Research

To strengthen the conclusions presented in Zmigrod et al. (2019a) and explore the previously discussed concerns, we aim to conduct a direct registered replication of Zmigrod and colleagues' Study 1. Complementing the original study, we also introduce additional analytical steps, an alternative independent, control, and outcome variable.

More precisely, we seek to assess the factor structure of the measures of willingness to fight and die for the ingroup to confirm the suitability of the indirect effects models. In addition to the RAT and WCST, we will implement a self-report measure of cognitive flexibility, the Cognitive Flexibility Inventory, with sub-dimensions for perceived ability to control difficult situations (Control) and ability to perceive/generate multiple alternative explanations and solutions in difficult situations (Alternatives) (CFI; Dennis & Vander Wal, 2010). Previous research has found that self-report measures of cognitive flexibility achieve higher effect sizes (van Hiel et al., 2016), but may, in fact, assess perceived need for certainty or closure instead of cognitive style (Zmigrod, 2020). However, we believe that there are three reasons that justify examining the relationship between cognitive flexibility and support for violent extremism with the CFI.

First, the application of neuropsychological tests such as the WCST and RAT is less common outside the field of psychology and researchers from other disciplines may be hesitant to use these. At the same time, there is no standard measure for cognitive flexibility - the concept has been assessed with numerous tasks that investigate different cognitive subsystems (Gruner & Pittenger, 2017; Syner et al., 2015; Tchanturia et al., 2003; Zmigrod et al., 2019b) as well as with two self-report measures (CFI and the Cognitive Flexibility Scale; Martin & Rubin, 1995). It is therefore important to examine the robustness of the results with an alternative self-report tool that is likely to be applied (as well) in future studies.

Moreover, previous research has raised concerns about task impurity - more than one cognitive system is investigated - and the construct validity of the WCST and RAT. The WCST requires visual and numeric processing as well as the ability to process verbal feedback and maintain rules in the working memory (Miyake et al., 2000; Lee et al., 2014). The RAT was developed to measure creative performance (Mednik,1962) but assessment of its

convergent validity showed positive correlations with scores on working memory tasks, processing speed tasks, intelligence tasks, and grade point average, as well as, counterintuitively, low correlations with divergent thinking measures (Lee et al., 2014). With this in mind, discrepancies in results of the WCST and RAT accuracy rates may be attributed to the fact that the tests examine different cognitive subsystems and capture cognitive flexibility to varying degrees. The Alternatives sub-scale of the CFI has been found to correlate positively with the WCST (Johnco et al., 2014). By including the CFI, we can assess whether results patterns for the Alternatives sub-scale and the WCST are comparable, which would lend support to the speculation that findings of the RAT and WCST are inconsistent because the tests do not capture cognitive flexibility in the same way and degree.

We also extend the direct replication by introducing a control variable that was considered in Zmigrod et al., 2019a; Study 2 as a mediator: identity fusion. Identity fusion is defined as "a visceral sense of 'oneness' with a group and its individual members that motivates personally costly, pro-group behaviors" (Swann & Buhrmester, 2015, p. 52). Previous research has consistently shown that identity fusion is related with willingness to fight and die for an ingroup (Gomez et al., 2011; Swann et al., 2009; 2010). We therefore aim to examine the extent to which cognitive flexibility predicts support for violent extremism over and beyond identity fusion. Lastly, we assess whether cognitive flexibility is solely negatively associated with violent extremist behaviour intentions or whether it predicts as well normative pro-group intentions. Although not articulated in detail, Zmigrod and colleagues (2019a) propose that lower levels of cognitive flexibility affect support for violent extremism by strengthening ideological attachment with the ingroup - operationalised through willingness to fight for the group. If this rationale held true, it would be expected that lower levels of cognitive flexibility is also related to stronger intentions to support the ingroup through normative means, such as attending demonstrations (Simon & Klandermans, 2001).

In summary, the following hypotheses are investigated, resembling those assessed in the original study:

H1: Cognitive flexibility is negatively related to support for violent extremism.

H2: Lower cognitive flexibility predicts stronger support for violent extremism mediated by a (higher) willingness to fight for one's ingroup.

H3: Lower cognitive flexibility predicts greater conviction in one's willingness to sacrifice oneself for the ingroup.

In addition, we explore:

R1: Are willingness to fight and die unique concepts?

R2: Are the aforementioned relationships (H1-3) replicated when an additional self-report measure for cognitive flexibility is introduced?

R3: Are the aforementioned relationships (H1-3) replicated when controlling for identity fusion?

R4: Are the aforementioned relationships (H1-3) replicated when considering normative support for the ingroup as an outcome variable?

Method

The pre-registration of the direct replication study as well as all data will be available online [enter link when available].

Participants. The sample size for the direct replication is determined based on an a priori power analysis. Relying on the effect sizes of the original study is not recommended when designing replication studies, as this leads to underpowered replications (Albers & Lakens, 2018; Simonsohn, 2015). We therefore determined first the smallest effect size of interest (SESOI; Lakens et al., 2018) by following Simonsohn's (2015) advise to consider the

effect size that the original study would have obtained with 33% power. The SESOI is determined as f2 = 0.016 - a small effect. Conducting an a-prior power analysis for multiple regressions - α = .05, power = .95 - then yields a target sample size of N = 1333.

British citizens who are at least 18 years old will be recruited. No quota for socio-demographic characteristics or further exclusion criteria will be defined. An attention check will be added after all measures/tests that were included in the original study were assessed. Those who fail the attention check will be excluded for exploratory analyses but not for the direct replication analyses. In the exploratory analyses, we will further exclude all participants who completed the study in less than five minutes. Missing data will not be replaced at any point.

Materials. The survey will be programmed using Qualtrics Survey Software. Participants will report their age, gender, ethnicity and highest educational attainment. In line with the original study, willingness to fight for the ingroup is assessed with a five-item measure ('I would fight someone physically threatening another British person', 'I would fight someone insulting or making fun of the United Kingdom as a whole' , 'I would help others get revenge on someone who insulted the United Kingdom', 'Hurting other people is acceptable if it means protecting the United Kingdom', 'I'd do anything to protect the United Kingdom'; 1 = *totally disagree*; 7 = *totally agree;* see Swann et al., 2009). Participants will indicate their willingness to die for the ingroup, first, by stating the extent to which they agree with the statement: 'I would sacrifice my life if it saved another group member's life' (1 = *totally disagree*; 7 = *totally agree;* see Swann et al., 2009). They will also be presented with a trolley dilemma and express how they would behave in the following situation: "Imagine that a runaway trolley is about to crush and kill 5 British people. You have the opportunity to jump from a bridge into the trolley's path and save all 5 British people. Would you: (a) let the trolley crush the 5 British people and save your own life, OR (b) save the 5 British people and

sacrifice your own life?" (Zmigrod et al., 2019a, p. 4). Following, we will examine how certain participants are of their decision in the trolley problem ('How certain are you that you would let the trolley crush the 5 British people and save your own life/save the 5 British people and sacrifice your own life?', 1 = *low certainty*, 100 = *high certainty*).

Cognitive flexibility will be measured with the WCST and the RAT. The WCST will be run through the license-free PsyToolkit and not Inquisit 5 by Millisecond Software, as done in the original study. Participants will see four cards on which different coloured geometric forms are presented; cards vary in the colour, type, and number of geometric forms. A fifth card is then presented, which must be matched to one of the four stimulus cards. The match may be due to cards having the same number of forms, forms of the same colour, or the same forms. Participants do not know the rule for matching in advance but are given feedback on their choices. The matching rule changes after participants have correctly used a rule 10 times. All participants will complete 128 trials. Accuracy rates will be calculated (correct trials divided by total trials). The RAT requires participants to create words that serve as a connection, or compound, for three stimulus words that are only remotely connected. For instance, the word connecting 'worm', 'shelf', and 'end' would be 'book'. We use the same problems that were used in the original study (based on Bowden & Jung-Beeman, 2003). Participants complete 20 problems for which only a single solution is correct and are given 20s for each problem.

Additional measures. We additionally will include a measure of identity fusion that was used in Study 2 of Zmigrod et al., 2019a: the Dynamic Identity Fusion Index (DIFI; Jimenez et al., 2016). Participants are asked to position a small circle representing themselves in relation to a large circle that represents the United Kingdom; the overlap between the circles reflects the extent to which the self is fused with the national identity.

We use the 20-item Cognitive Flexibility Inventory (Dennis & Vander Wal, 2010; e.g., 'I have a hard time making decisions when faced with difficult situations', ' I like to look at difficult situations from many different angles', 1 = *disagree completely*; 7 = *agree completely*) as a self-report measure of cognitive flexibility. Mean scores are calculated for the two sub-scales Control and Alternatives.

Normative pro-group behaviour is examined with four items of the 'Activism' dimension of the Activism and Radicalism Intention Scales (ARIS; 'I would donate money to an organisation that fights for my group's political and legal rights', 'I would volunteer my time working (i.e. write petitions, distribute flyers, recruit people, etc.) for an organisation that fights for my group's political and legal rights', 1 = *disagree completely*; 7 = *agree completely;* Moskalenko & McCauley, 2009); one mean score is calculated. Finally, we will introduce an attention check embedded within the items of the CFI ('This is an attention check. Please tick answer option '5' to show that you are paying attention').

Procedure. Participants will be recruited via Prolific Academic, an online opt-in access panel. They will receive a wage-based payment of approximately £5/hour upon completion of the study. Selection criteria - age and UK citizenship - will be specified directly on Prolific so that only those eligible may participate. Self-report measures of willingness to fight and die are completed first, followed by the trolley problem, the RAT, and the WCST. Afterwards, participants will report demographic information. Only then will the measures that were not used in the original study be presented. We first introduce the CFI, including the attention check, and then assess identity fusion, and normative pro-group behaviour intentions.

Planned Analyses

All analyses will be completed with R. Alternative tests/transformations may be used, and will be clearly indicated as such, if assumptions are violated. Effect sizes and confidence intervals will be reported for all analyses.

Primary Analyses Replicating the Original Study. First, accuracy rates for the RAT and WCST as well as a mean score of the five items assessing willingness to fight for the ingroup are computed. For all continuous variables, bivariate correlations, means, and standard deviations are calculated. Chronbach's alpha is assessed for the 'willingness to fight' scale. To assess Hypothesis 1, bivariate correlations between WCST and RAT accuracy rates, willingness to fight and willingness to die for the ingroup will be assessed.

Testing Hypothesis 2, we replicate the path models of the original study. The independent variables are WCST and RAT accuracy rates; the mediator is willingness to fight for the group, and the dependent variables are willingness to die for the national ingroup and chosen behaviour in trolley dilemma. Residual covariances are allowed between all independent and dependent variables, respectively. All analyses will control for age, gender, and educational attainment, and residual covariances are allowed between these demographic variables. Fit of two models will be compared. Model 1 includes direct paths between willingness to die/trolley dilemma behaviour, willingness to fight for the group, as well as WCST and RAT accuracy rates. Model 2 specifies no direct path between the independent and dependent variables; the association between WCST and RAT accuracy rates and willingness to die/trolley dilemma behaviour is to be fully mediated by willingness to fight for the group. Lastly, to investigate Hypothesis 3, a correlation between confidence in their decision to self-sacrifice and WCST and RAT accuracy rates will be calculated. This correlation is then repeated in a sub-group of people who indicated in the trolley dilemma their willingness to self-sacrifice. In addition, a multivariate analysis of variance (MANOVA) will be conducted with

the factor being decision to self-sacrifice in the trolley dilemma and the dependent variables being WCST and RAT accuracy rates.

Additional Analyses. We will further compute a confirmatory factor analysis (CFA) of the items assessing willingness to fight and die for the ingroup to determine whether these represent two correlated, yet distinct, factors. Confirmatory factor analysis will also be completed for the CFI to confirm the two sub-scales, Control and Alternatives. In both CFA's errors of variables that load on the same factor are set to be correlated. To complement the analyses assessing Hypothesis 1, we will also compute a Pearson correlation between the Control and Alternatives scores of the CFI and willingness to fight and die for the group as well as point biserial correlations with willingness to self-sacrifice in the trolley dilemma. We will determine whether the relationship between the CFI scores and the dependent measures replicates associations with WCST and RAT accuracy rates (see below how we judge success of a replication).

Adding to the analysis of Hypothesis 2, we will extend the aforementioned two models such that Control and Alternatives scores are included as manifest, independent variables. Residual covariances are allowed between the two CFI scores, WCST, and RAT accuracy rates. In Model 1, additional direct paths between willingness to die/trolley dilemma behaviour, the willingness to fight for the group, as well as self-reported cognitive flexibility scores are specified. In Model 2, no direct path between the CFI scores and dependent variables is modelled; the association between self-reported cognitive flexibility and willingness to die/trolley dilemma behaviour is to be fully mediated by the willingness to fight for the group. Further to demographic variables, analyses will control for identity fusion. In Model 1 and 2, direct paths between identity fusion and the two outcome variables are proposed; residual correlations are allowed between identity fusion and willingness to fight. We explore if Model 2 provides a better fit than Model 1, whether overall result patterns for the CFI and RAT as

well as WCST are comparable, as well as whether Model 2 with additional independent and control variables replicates the findings of the original study. Following, we assess Model 1 and 2 as described in the original study as well as the aforementioned extended Model 1 and 2 by considering normative pro-group behaviour as the outcome variable. We assess the extent to which result patterns pertaining to violent extremist behaviour intentions are replicated for the new dependent measure.

Hypothesis 3 is also tested by including the self-reported cognitive flexibility scores, Control and Alternatives, in the correlation analysis and as dependent variables in the MANOVA. We assess if the findings of the objective measures are replicated. Lastly, as a robustness test, all aforementioned primary and additional analyses will be conducted with a sample where those who failed the attention test and those who completed the study in less than 5 minutes were excluded. We explore whether all results patterns hold when these stricter exclusion criteria are applied.

Summary of Differences Between the Original and Replication Study

The direct replication differs from the original study only insofar as we use PsyToolkit to administer the WCST. All additional measures for the independent variable - the CFI - as well as a control variable, identity fusion, the attention check, and a measure for normative pro-group behaviour intentions are assessed after the protocol for the direct replication is completed.

Determining Success of the Replication

Success of replication of result patterns - both a direct replication of the original study as well as replications of findings relying on different measures - is determined based on the following criteria. For Hypothesis 1 and 3, we, firstly, rely on the *p*-value ($< .05$) to determine whether results that were identified as statistically significant in the original analysis are

also statistically significant in the replication. We also confirm that the direction of the relationship is in the same direction. In addition, we inspect the 95% CI of the correlation coefficients of the replication analyses to determine whether the effect identified in the original study falls within this interval (Nosek et al., 2015). To test replications of Hypothesis 2, we inspect the X^2 test, the RMSEA including its confidence interval (acceptable score: .05 to .08), SRMR (acceptable score: .05 to .10), and comparative fit index (acceptable score: .95 to .97) and determine whether an acceptable fit is achieved.

References

Alkhadher, O., & Scull, N. C. (2019). Demographic variables predicting ISIS and Daesh armed political violence. *Crime, Law and Social Change*, *72*(2), 183-194. Available at: https://link.springer.com/article/10.1007/s10611-018-9808-5

Aly, A., Taylor, E., & Karnovsky, S. (2014). Moral disengagement and building resilience to violent extremism: An education intervention. *Studies in Conflict & Terrorism*, *37*(4), 369-385. DOI: 10.1080/1057610X.2014.879379

American Psychological Association (2020). Cognitive process. Available at: https://dictionary.apa.org/cognitive-process

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. Harper and Rowe: New York.

Atran, S. (2016). The devoted actor: unconditional commitment and intractable conflict across cultures. *Current Anthropology*, *57*(S13), 192-203. Available at: https://www.journals.uchicago.edu/doi/full/10.1086/685495

Baele, S. J. (2017). Lone-actor terrorists' emotions and cognition: An evaluation beyond stereotypes. *Political Psychology*, *38*(3), 449-468. DOI: 10.1111/pops.12365

Becker, J. C., & Tausch, N. (2015). A dynamic model of engagement in normative and non-normative collective action: Psychological antecedents, consequences, and barriers. *European Review of Social Psychology*, *26*(1), 43-92. DOI: 10.1080/10463283.2015.1094265

Boendermaker, W. J., Peeters, M., Prins, P. J., & Wiers, R. W. (2017). Using serious games to (re) train cognition in adolescents. In: Ma M., Oikonomou A. (eds). *Serious games and edutainment applications* (pp. 307-321). Springer, Cham.

Bouhana, N. (2019). The moral ecology of extremism: A systemic perspective. *UK Commission for Countering Terrorism*.

Buitenweg, J. I., van de Ven, R. M., Prinssen, S., Murre, J. M., & Ridderinkhof, K. R. (2017). Cognitive flexibility training: A large-scale multimodal adaptive active-control intervention study in healthy older adults. *Frontiers in Human Neuroscience*, *11*, 529. Available at: https://www.frontiersin.org/articles/10.3389/fnhum.2017.00529/full

Davies, L. (2016). Security, extremism and education: Safeguarding or surveillance?. *British Journal of Educational Studies*, *64*(1), 1-19. DOI: 10.1080/00071005.2015.1107022

Dennis, J. P., & Vander Wal, J. S. (2010). The cognitive flexibility inventory: Instrument development and estimates of reliability and validity. *Cognitive Therapy and Research*, *34*(3), 241-253. Available at: https://link.springer.com/article/10.1007/s10608-009-9276-4

De Vries, M., Prins, P. J., Schmand, B. A., & Geurts, H. M. (2015). Working memory and cognitive flexibility-training for children with an autism spectrum disorder: A randomized controlled trial. *Journal of Child Psychology and Psychiatry*, *56*(5), 566-576. DOI: 10.1111/jcpp.12324

Gill, P., & Corner, E. (2017). There and back again: The study of mental disorder and terrorist involvement. *American Psychologist*, *72*(3), 231-241. DOI: 10.1037/amp0000090

Gill, P., Horgan, J., & Deckert, P. (2014). Bombing alone: Tracing the motivations and antecedent behaviors of lone-actor terrorists. *Journal of Forensic Sciences*, *59*(2), 425-435. DOI: 10.1111/1556-4029.12312

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*(4), 404. DOI: 10.1037/h0059831

Jasko, K., LaFree, G., & Kruglanski, A. (2017). Quest for significance and violent extremism: The case of domestic radicalization. *Political Psychology*, *38*(5), 815-831. DOI: 0.1111/pops.12376

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*(3), 339-375. DOI: 10.1037/0033-2909.129.3.339

Klandermans, P. G. (2014). Identity politics and politicized identities: Identity processes and the dynamics of protest. *Political Psychology*, *35*(1), 1-22. DOI: 10.1111/pops.12167

Liht, J., & Savage, S. (2013). Preventing violent extremism through value complexity: Being Muslim being British. *Journal of Strategic Security*, *6*(4), 44-66. Available at: https://www.jstor.org/stable/26466749?seq=1#metadata_info_tab_contents

Lösel, F., King, S., Bender, D., & Jugl, I. (2018). Protective factors against extremism and violent radicalization: A systematic review of research. *International Journal of Developmental Science*, *12*(1-2), 89-102. DOI: 10.3233/DEV-170241

McCauley, C., & Moskalenko, S. (2017). Understanding political radicalization: The two-pyramids model. *American Psychologist*, *72*(3), 205-216. DOI: 10.1037/amp0000062

Mednick, S. A. (1968). The remote associates test. *The Journal of Creative Behavior*. 2, 213–214. DOI: 10.1002/j.2162-6057.1968.tb00104.x

Meeus, J., Duriez, B., Vanbeselaere, N., & Boen, F. (2010). The role of national identity representation in the relation between in-group identification and out-group derogation: Ethnic versus civic representation. *British Journal of Social Psychology*, *49*(2), 305-320. DOI: 10.1348/014466609X451455

Moghaddam, F. M. (2005). The staircase to terrorism: A psychological exploration. *American Psychologist*, *60*(2), 161-169. DOI: 10.1037/0003-066X.60.2.161

Mun, Y., Oprins, E., Van Den Bosch, K., Van Der Hulst, A., & Schraagen, J. M. (2017, September). Serious gaming for adaptive decision making of military personnel. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 1168-1172). Sage CA: Los Angeles, CA: SAGE Publications.

Obaidi, M., Bergh, R., Akrami, N., & Anjum, G. (2019). Group-based relative deprivation explains endorsement of extremism among western-born muslims. *Psychological Science*, *30*(4), 596-605. DOI: 10.1177/0956797619834879

Perry, G., Wikström, P. O. H., & Roman, G. D. (2018). Differentiating right-wing extremism from potential for violent extremism: the role of criminogenic exposure. *International Journal of Developmental Science*, *12*(1-2), 103-113. DOI: 10.3233/DEV-170240

Rokeach, M. (1948). Generalized mental rigidity as a factor in ethnocentrism. *The Journal of Abnormal and Social Psychology*, *43*(3), 259-278. DOI: 10.1037/h0056134

Sageman, M. (2004). *Understanding terror networks*. Philadelphia, USA: University of Pennsylvania Press.

Savage, S., Khan, A., & Liht, J. (2014). Preventing violent extremism in Kenya through value complexity: Assessment of Being Kenyan Being Muslim. *Journal of Strategic Security*, *7*(3), 1-26. Available at: https://www.jstor.org/stable/26465191?seq=1#metadata_info_tab_contents

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2020). Monte Carlo Power Analysis for Indirect Effects. Available at: https://schoemanna.shinyapps.io/mc_power_med/?fbclid=IwAR0lzCyFXSOfi2Eaco76j49VdhjmQ1chjEEpcR4bDv002OyuNCIH8_lyZ_U

Schuurman, B. (2018). Research on terrorism, 2007–2016: A review of data, methods, and authorship. *Terrorism and Political Violence*, 1-16. DOI: 10.1080/09546553.2018.1439023

Sidanius, J. (1985). Cognitive functioning and sociopolitical ideology revisited. *Political Psychology*, 637-661. DOI: 10.2307/3791021

Sidanius, J., & Ekehammar, B. (1976). Cognitive functioning and socio-political ideology A multidimensional and individualized analysis. *Scandinavian Journal of Psychology*, *17*(1), 205-216. DOI: 10.1111/j.1467-9450.1976.tb00232.x

Stephens, W., Sieckelinck, S., & Boutellier, H. (2019). Preventing violent extremism: A review of the literature. *Studies in Conflict & Terrorism*, 1-16. DOI: 10.1080/1057610X.2018.1543144

Tausch, N., & Becker, J. C. (2013). Emotional reactions to success and failure of collective action as predictors of future action intentions: A longitudinal investigation in the context of student protests in Germany. *British Journal of Social Psychology*, *52*(3), 525-542. DOI: 10.1111/j.2044-8309.2012.02109.x

Tausch, N., Becker, J. C., Spears, R., Christ, O., Saab, R., Singh, P., & Siddiqui, R. N. (2011). Explaining radical group behavior: Developing emotion and efficacy routes to normative and nonnormative collective action. *Journal of personality and social psychology*, *101*(1), 129-148. DOI: 10.1037/a0022728

Van Hiel, A., & Mervielde, I. (2003). The measurement of cognitive complexity and its relationship with political extremism. *Political Psychology*, *24*(4), 781-801. DOI: 10.1046/j.1467-9221.2003.00354.x

Van Hiel, A., Onraet, E., Crowson, H. M., & Roets, A. (2016). The relationship between right-wing attitudes and cognitive style: A comparison of self-report and behavioural measures of rigidity and intolerance of ambiguity. *European Journal of Personality*, *30*(6), 523-531. DOI: 10.1002/per.2082

Van Hiel, A., Onraet, E., & De Pauw, S. (2010). The relationship between social-cultural attitudes and behavioral measures of cognitive style: A meta-analytic integration of studies. *Journal of Personality*, *78*(6), 1765-1800. DOI: 10.1111/j.1467-6494.2010.00669.x

Van Zomeren, M., Leach, C. W., & Spears, R. (2012). Protesters as "passionate economists" a dynamic dual pathway model of approach coping with collective disadvantage. *Personality and Social Psychology Review*, *16*(2), 180-199. DOI: 10.1177/1088868311430835

Vergani, M., Iqbal, M., Ilbahar, E., & Barton, G. (2018). The three Ps of radicalization: Push, pull and personal. A systematic scoping review of the scientific evidence about radicalization into violent extremism. *Studies in Conflict & Terrorism*, 1-32. DOI: 10.1080/1057610X.2018.1505686

Webster, D. M., & Kruglanski, A. W. (1997). Cognitive and social consequences of the need for cognitive closure. *European Review of Social Psychology*, *8*(1), 133-173. DOI: 10.1080/14792779643000100

Zmigrod, L. (2020). The role of cognitive rigidity in political ideologies: theory, evidence, and future directions. *Current Opinion in Behavioral Sciences*, *34*, 34-39. DOI: 10.1016/j.cobeha.2019.10.016

Zmigrod, L., Rentfrow, P. J., & Robbins, T. W. (2018). Cognitive underpinnings of nationalistic ideology in the context of Brexit. *Proceedings of the National Academy of Sciences*, *115*(19), E4532-E4540. DOI: 10.1073/pnas.1708960115

Zmigrod, L., Rentfrow, P. J., & Robbins, T. W. (2019a). Cognitive inflexibility predicts extremist attitudes. *Frontiers in Psychology*, *10*, 989. Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00989/full

Zmigrod, L., Rentfrow, P. J., & Robbins, T. W. (2019b). The partisan mind: Is extreme political partisanship related to cognitive inflexibility?. *Journal of Experimental Psychology: General*.