



LUNDS
UNIVERSITET

DEPARTMENT of PSYCHOLOGY

***Evaluative neutralisation as a means to reduce
social desirability in the SDO-7***

Alexander Dahlström

Master's Thesis (30 hp)

Spring 2021

Supervisors: Fredrik Björklund, Martin Bäckström

Abstract

An evaluative neutralisation process, whereby questionnaire items are phrased to be less evaluative, was used to create a new less socially desirable version of the social dominance orientation scale (SDO7). The main goal is to counteract the persistent floor effect which has affected the scale since its inception. The rationale behind the suggested correction is rooted with the problem of social desirability in self-rating inventories. A translated scale (Swedish) was used and it was constructed and tested over three different samples. The final sample consisted of 245 participants (61.5% female). It is argued and demonstrated that simple neutralisation can lessen the floor effect, increase variance and even result in an improved inventory, as measured by the scale's predictive power compared to 4 criterion variables. Paired samples *t*-tests indicated that the mean of the new scale was almost 1 point higher on a 7 point Likert-scale compared to the original scale, variance increased by 30%, normality of the SDO scores were better distributed and number of outliers dropped from 13 to 3. Correlations between the new scale and criterion variables were significantly stronger on half of the criterion-variables and stronger on the rest but without reaching significance. All of these improvements suggest that the new scale outperforms the original and that the evaluative neutralisation process was successful. The importance of an evaluative neutralisation for future SDO research is discussed in terms of construct validity and predictive power.

Keywords Evaluative Neutralisation, Social Dominance Orientation, Construct Validity, Social desirability

Acknowledgements

A special thanks to my two supervisors, Fredrik Björklund and Martin Bäckström, who very passionately offered guidance and expertise during this project. It is rare to have two, even one supervisor who offers such excellent feedback and advice.

And a professional thank you to one of the creators of the very field that has dominated my scientific inquiry as of late, Professor Jim Sidanius, who sadly passed away June 29th 2021, at the age of 75. With more than 330 published papers, he was a giant within social psychology and contributed heavily to prejudice research.

Intuitively, one could argue that equality is something that should be desirable by any given individual and that at least the vast majority of us would actively be striving for. This line of thought is erroneous however, at least according to Social Dominance Theory (SDT). The explanation put forward by Pratto et al. (1994) is that a society aims to minimise intergroup conflict by attempting to create a narrative that everyone can get behind. Hierarchy-legitimising myths (HLM) are produced by powerful groups in society in order to maintain and legitimise the current hierarchy. If everyone is under the assumption that a given social group is more adept at ruling, it would be hard to give credit to any potential criticism and people will tend to accept the status quo. HLM then enables society to function with less conflict, while anti-egalitarian policies can ensure that the hierarchy isn't challenged by keeping people in their place. An example of these HLMs can be sexist beliefs, such as the belief that women are more suited as house wives. If this belief is generally accepted, then women will pose little threat to the hierarchy since they sacrifice the chance of engaging in politics to change their situation. The odds of them voting for more progressive policies would also be low, given that they actually accept the HLMs as truths. Pratto et al. (1994) offers examples on a more ideological level, namely meritocracy. A person who holds a meritocratic ideology argues that the most merited person should get the job. Thus, if someone is successful, it ought to be because he or she earned it and consequently, if someone is struggling in life, it ought to be due to their lack of effort. This is purely the prescriptive nature of meritocracy. There is also a belief that society actually is meritocratic, which means that social issues such as discrimination, favouritism or nepotism isn't a problem. Such an ideology is the perfect example of a HLM, since it not only directly promotes the status quo, but it can also be used as an argument to counter egalitarian policies, such as affirmative action. The opposite of HLMs would be hierarchy-attenuating myths, HAMs (Pratto et al., 1994). Since HAMs are the direct opposite of HLMs, they would as such create friction in a highly hierarchic society and stimulate change. This type of socio-political and or prejudice research is becoming more and more important. The rise of populist right wing parties all over Europe demonstrates the importance of prejudice research (Forchtner, 2019). All of the mentioned socio-political attitudes are measurable via self-rating inventories. Such inventories are extremely useful, yet often flawed. The main objective of this paper is to improve the validity and reliability of the scale that is derived from SDT, namely Social Dominance Orientation (SDO).

Social Dominance Orientation

SDO measures a person's preference for a hierarchical societal design, resulting in some groups at the top and some groups at the bottom (Ho et al., 2015). There are two important aspects at play here, a dislike for egalitarianism and willingness to dominate outgroups. To anchor this in the previously mentioned SDT, SDO correlates positively with HLMs and negatively with HAMs (Pratto et al., 1994). An example item aiming to measure willingness to dominate from the SDO-7 is "Some groups of people must be kept in their place". Another item that focuses more on anti-egalitarianism reads "We should not push for group equality". A person who scores high on the SDO-7 would be characterised as one who sees the world as a competitive jungle, with finite resources (Ho et al., 2015). A person who attains a high score on the scale would as such be more prone to vote for anti-egalitarian policies and to use dominant and aggressive behaviour to ensure that the hierarchy is maintained. SDO is a very powerful predictor of racism, sexism, willingness to get ahead utilising dishonest methods, system justification, right wing politics and in more recent years, a strong correlation to climate denial has been established (Ho et al., 2015; Jylhä & Akrami, 2016; Pratto et al., 1994). During the ongoing pandemic, the scale has been used to indicate that a high SDO score correlates with depression (Shi et al., 2021). It can also be used to predict conspiracy mentality and scepticism regarding science and expert consensus (Dyrendal et al., 2021; Kerr & Wilson, 2021). On the other hand, SDO correlates negatively with egalitarianism, empathy and liberal politics (Ho et al., 2015; Pratto et al., 1994; Roy et al., 2021). Hudson et al (2019) found not only that higher levels of SDO correlates with a decrease in empathy, but also an increase in counter-empathy, which means taking pleasure in others suffering or misfortunes (often called *schadenfreude*). Consistent with previous SDO research, this *schadenfreude* is more pronounced when the misfortune of subjugated outgroups are targeted, compared to when ingroups are targeted. The two personality dimensions agreeableness and openness to experience from the big five also correlates negatively with SDO (Akrami & Ekehammar, 2006). It can even be used to predict job preference and education. Individuals who score highly on the SDO-scale tends to prefer jobs that are considered high in the hierarchy, such as within law enforcement (Ho et al., 2015).

The scale used to measure the construct has been revised many times. The latest edition is the SDO-7 (Ho et al., 2015). The SDO-7 features two sub-dimensions that were lacking in the previous versions which only measured SDO on a single dimension. These two new sub-dimensions are dominance (SDO-D) and anti-egalitarianism (SDO-E). The two sub-dimensions were first confirmed via confirmatory factor analysis, and later tested against a

range of different criterion variables. The different variables were sorted after dominance and anti-egalitarianism, and the two sub-dimensions were then used to predict the corresponding criterion. Examples of such criteria could be the previously mentioned meritocracy, which is a typical E-criterion and nationalism, which would be a prime example of a D-criterion. It is worth mentioning that SDO-E is actually a decent predictor for most D-criteria and vice versa. However, the correlations become more powerful when SDO-D predicts D-criteria and when SDO-E predicts E-criteria. This approach has opened up new venues in SDO research, enabling researchers to find the more driving factors for specific phenomena. One such example is that SDO-E seems to be the more driving factor of climate denial (Stanley et al., 2017). The SDO-scale can as such predict a range of negative phenomena through its positive correlations but also positive phenomena through its negative correlations, which definitely helps to illustrate the versatility, usefulness and importance of the scale.

The SDO-7 consists of 16 items, split equally between the two sub-dimensions of SDO-D and SDO-E. Half of the items are reversed, an equal number over the two sub-dimensions. This was another improvement and perhaps unintended critique towards the SDO-6, which wasn't originally designed with two distinct factors in mind. In the SDO-6, all items relating to dominance were pro-trait, while all the items relating to anti-egalitarianism were con-trait, items that needs to be reverse to calculate the score. This was a source of potential confound but has now been amended with the latest version of the scale (Ho et al., 2015).

Limitations and criticism of SDO

The SDO-7 and its predecessors have not gone unnoticed in the scientific community. Using Pubsearch, a search for social dominance orientation yields over 9000 hits and Pratto and Sidanius article from 1994 has 2589 citations on research gate. While the SDO-7 is a very powerful tool, it is not perfect. The construct has been criticised for being misleading. SDO was first described as a personality variable predicting social and political attitudes. This was criticised, since the items in the SDO don't seem to be measuring personality traits, but rather socio-political intergroup attitudes (Duckitt et al., 2002; Huang & Liu, 2005; Sidanius et al., 2006). The definition of SDO as a personality variable or not isn't a focus of this paper and the matter is more or less settled. It is however relevant to mention since a method used to affect the validity of personality inventories will be employed in this study.

A more pressing limitation is a persisting floor effect that has been present in the scale since its inception. The SDO-7 is usually conducted with a 7 point Likert-scale. Despite the long range, averages fall rather low. Ho et al. (2015) reported averages in samples as low as 2.46. The global average is even lower using the SDO-6 scale. Fischer et al. (2012) conducted a metaanalysis using the SDO-6 and found the average of 27 countries to be as low as 1.82. The variance is also very limited, suggesting that the difference between a high SDO score and a low SDO score is usually very low. There are exceptions to the floor effect, but they are rare. One such exception would be Japan, who scored an average of 4.41 (Fisher et al., 2012). When variance is limited, potential correlation to any given criterion variable should be lower. Other problems also arise when a floor effect occurs. Skewness is heavily affected by this floor effect, which limits statistical options when analysing the data. Outliers will of course be more common with extreme skewness which offers more difficulties when conducting statistical analysis. If the floor effect is partially or completely eliminated, there is a possibility that the construct could become an even stronger predictor of different variables, while enabling more sensitive methods of analysis. One approach to reduce to the floor effect would be to flip the scale, and measure a propensity for equality. Unfortunately, this would only lead to a ceiling effect instead. The SDO-7 already contains con-trait items, which basically measures the opposite of SDO. As expected, these items tend to produce very high scores, which in turn contribute to the floor effect when the items are later reversed scored. This floor effect is a very apparent problem which will be the main focus of this paper.

Social desirability and evaluative neutralisation

One plausible explanation to the floor effect is that of social desirability. Social desirability is a response-bias that potentially affects respondents whenever the topic can be related to social norms (Gittelman et al., 2015). Stated in simpler terms, people have a tendency to over-report good behavior and under-report bad behavior. This also holds true for attitudes and political beliefs (Gittelman et al., 2015). This is something that the designers of the SDO-scales haven't focused on, historically and in the present. An item from the SDO-7 states that "Some groups of people are simply inferior to other groups". The framing of the items makes it clear that disagreement with the statement is the desirable option. The ideological jargon is highly reminiscent of Fascism. It is important to keep in mind that egalitarianism is today somewhat of a social norm, even though it doesn't always follow from people's voting behaviors (Hurwitz & Peffley, 1992). In other words, people do tend to voice

a support for egalitarianism, yet this doesn't guarantee that they would support policies that are designed to empower minorities.

There are a lot of different strategies to combat social desirability. Dwight and Donovan (2003) proposed using a specific type of instruction. By warning respondents that a faking behaviour could be identified and punished, a lowered response bias could be attained. The effect size was rather small but still significant. Attaching a measure of social desirability to a study can also give researchers clues regarding the sample and if some participants are displaying a higher degree of social desirability. This strategy is however a bit messy, since more items are added to often already sizable surveys and it doesn't seem to improve criterion validity. The ideal solution would be one where social desirability can be controlled and where criterion validity actually increased, which would mean an actual improvement of the scale.

A potential solution was proposed by Bäckström and Björklund (2013). They found that evaluative neutralisation could lessen the impact of social desirability when using personality inventories. The process of evaluative neutralisation is quite simple and intuitive. Find out what makes an item socially desirable to agree or disagree with and rephrase it so that the social desirable option becomes less apparent. A good example illustrating the process would be the highly socially undesirable introversion item "Avoid contact with others" which was rephrased to "Feel at ease even with being alone" (Bäckström, Björklund & Larsson, 2014). In this example, the substantive meaning remains very close to the original and further testing assured that the new item actually measured the same thing as the original. The items are assessing how one feels about being around other people. Yet the evaluative content, whether this is measuring a good or bad behaviour, is completely altered. Bäckström and Björklund (2013) identified items which were suspected to suffer from social desirability based on deviations from the midpoint of a Likert-scale. While more direct measures of social desirability were used, a PCA was conducted to indicate loadings from different, seemingly uncorrelated personality traits to the first PC. The rationale was that if the validity of the scale had been affected by social desirability, then a PCA based on all items would identify a social desirability loading. While the diagnosis of the problem was fairly complex and clever, the solution to the problem was almost simplistic. Undergraduate students were used to construct evaluative neutralised items based on the original scales. The result were new scales that kept the integrity of the five factor model intact, while significantly reducing the correlation between items and social desirability measures, as well as the loadings to the first PC. This is

a good indicator that construct validity improved. The inventory seems to have improved in regards to measuring what it was actually supposed to measure. Bäckström and Björklund (2013) concludes that this was by no means a miracle cure for social desirability, yet points out how a simple neutralisation process can significantly improve self-rating scales. The method has been replicated at a larger scale with similar results (Wood et al., 2021). More recent research has also indicated that construct and discriminant validity can be increased by utilising the evaluative neutralisation method (Bäckström & Björklund, 2020). There is no reason to believe that such a solution wouldn't be effective when it comes to the SDO-7. The scales are indeed inherently different. As we have concluded, the traits being measured by the big-five are more stable personality traits, while SDO seems to measure a more fluent trait that is heavily influenced by the social environment. However, they also have similarities. Both scales are self-rated measures with rather obvious socially desirable answers. This could arguably be the only needed requisite to make use of the proposed evaluative neutralisation. If this is true, Bäckström and Björklund's (2020) methods can be generalised to self-rated socio-political inventories.

Purpose

The aim of the present study is to attempt to improve on the existing SDO-7 scale by means of the evaluative neutralisation process suggested by Bäckström and Björklund (2013). The first hypothesis of this study is that the evaluative neutralisation of the items will result in a higher mean which is closer to the mid-point of the Likert-scale. In the present study, a 7-point Likert-scale is used, so the mid-point is 4. The second hypothesis is that variance will increase, as measured by the items' standard deviations. The rationale here is that people are more or less affected by social desirability. If the scales are indeed neutralised yet keep their validity, some will score the same, while some (enhancers) will score higher. If variance between the two scales remains unaffected, this will indicate a confounding factor. The increased variance should lessen the floor effect of the scale which in turn should produce more normal data, i.e less skewness, less kurtosis and fewer outliers. Obtaining normal data is important for statistical analyses. While a higher mean and increased variance offers an intrinsic value in form of a wider data-span, it is of course important that the scale still captures the characteristics of the SDO-7. This will be determined by investigating the correlation between the original SDO-7 and the neutralised scale, by testing the homogeneity of the two scales (Chronbach's alpha) and their respective predictive power on two SDO-D criterion variables and two SDO-E criterion variables. Factor analysis and Principal

components analysis will also be included for the same reason, to ensure that the new scale captures the characteristics of the SDO-7 and retains the four-factor solution. The third hypothesis is that the new scale will outperform the SDO-7 by producing stronger correlations to criterion variables.

Method

Participants

The participants ($N=321$) were recruited by convenience sampling over social media as well as via recruiting participants on the streets in Malmö and Lund. The response rates were low. For the final and largest survey, 1400 were registered as consenting on Qualtrics yet only 245 actually finished the survey. There were no requirements for participating in the study besides being at least 18 years old and understanding Swedish. See Table 1 for demographic information across the four samples.

Table 1. Number of participants, gender distribution and age distribution across the 4 samples

Sample	<i>N</i>	% Female	Age <i>M</i>	Age <i>SD</i>
Survey 1	30	73.30	36.60	12.10
Survey 2	16	56.30	41.44	13.38
Survey 3	30	66.70	40.43	14.56
Survey 4	245	61.50	49.21	14.68

Material

The neutralisation process was conducted by simply discussing and altering the items until a neutralised yet semantically relatable item was produced. The main goal of the neutralisation was to make the socially desirable answer less explicit, while keeping the actual meaning of the item intact. An example of a neutralised item would be “Some groups are simply not as capable as other groups” compared to the original “Some groups of people are simply inferior to other groups”. Here, the rationale is that it is more socially desirable to reject talk of inferiority, compared to acknowledging that some groups are not as capable. In order to arrive on the best solution, items were constructed, compared and amended in three consecutive surveys to arrive at the fourth and final scale. All the surveys included informed consent as well as demographic variables of age and gender. The first block would always be

the translated version of the SDO-7 scale (Grina et al., 2016) coupled with the neutralised items. The SDO-7 used in this paper can be found in appendix B. All of these items were randomised to prevent order effects as well as other potential biases. The second block for Survey 1 and 3 were demographic variables of age and gender, while Survey 2 and 4 contained an additional block with criterion variables.

The first survey consisted of 54 items, whereof 16 were the translated version of the SDO-7 (Grina et al., 2016). 16 of the more promising items were used to construct a comparable scale (SDO2.0). Items were selected based on mean, variance and corrected item-total correlations. An item was automatically discarded if it failed to correlate with the rest of the items in the scale.

Using the 16 items identified in the first survey, a smaller survey was issued coupled with criterion variables to see if the new scale had predictive powers similar to that of the translated original scale. This survey consisted of the 16 translated original items from the SDO-7, the new 16-item SDO2.0 scale, as well as 6 items related to SDO-E and SDO-D criterion variables. These criterion variables were chosen from the Ho et al. (2015) study and the SDO-E criterion variable was political orientation on a left- to right-scale (adapted from the original liberal- to conservative-scale which isn't applicable in Sweden) and punitiveness, a measure of harsher law enforcement methods were used as a SDO-D criterion.

The third survey included 49 items, 16 of which were the translated original SDO-7, 16 were the SDO2.0 scale and the remaining 17 items were new neutralised items.

The fourth and final survey included the 16 translated original SDO-7 items, the 16 items from SDO2.0 and 5 more new neutralised items. Additionally, it contained 25 items related to 4 distinct criterion variables, whereof 2 were SDO-E and 2 were SDO-D. Many of the original SDO-7 criterion variables were deemed unfeasible for a Swedish population. For example, SDO-D criterion variables for the SDO-7 contained war legitimacy beliefs, support for the death penalty and nationalism. Sweden is a pacifist country which doesn't actively partake in wars, we have no death penalty, and while nationalism is probably a thing in any given country, the items in the SDO-7 criterion variable was highly adapted to the American population. For example, the second item in the scale read "This country must continue to lead the "Free World", which is hardly applicable to Sweden.

Criterion variables

Four different scales were used to produce two SDO-E and two SDO-D criterion variables in the final survey. For SDO-E, a 3-item scale measuring political orientation (left/right) (Dimdins et al., 2016) was used coupled with a 5-item scale measuring meritocracy (Zimmerman & Reyna, 2013). Note that the meritocracy scale actually has 10 items, 5 of which measures what society ought to be like and 5 which measures what society is really like. The prescriptive items were not included. For SDO-D, the translated version of dirty dozen dark triad was used (Garcia et al., 2018), as well as 4-item scale measuring Ethnic minority political exclusionism (MINEXCL) (Šram, 2020). The MINEXCL was chosen over the more traditional classical racism scale, since minority exclusionism is a more relevant measure in Sweden compared to the anti-black racism which has been used frequently in SDO-research in the USA. The scales measuring meritocracy and minority exclusionism were back translated to Swedish. All of these criterion variables can be found in appendix B.

Procedure

The estimated time to complete the survey varied greatly across the different surveys due to the different number of items in each. The shortest survey was estimated to take 8 minutes to complete, while the final survey was estimated to take 13 minutes to complete. The participants were first given a vague description of the purpose of the study, which was to improve an already existing scale measuring socio-political attitudes. The participants were also informed that they were free to exit the survey prematurely without consequences. The participants would then have to read and agree to an informed consent-form, stating that they are above the age of 18 and that they wish to proceed with the survey. After the survey was completed, the participants were given an email address and were encouraged to write any potential question or feedback they might have.

Ethical considerations

This study is in compliance with the ethical guidelines of the Swedish research council. All participants were given a description of the research they were contributing to, prior to participating. It was stated that a participant may at any point during the survey withdraw their consent by exiting the survey before submission. It was clearly stated that the participant would not face any adverse consequences for doing so. After the information was read, participants had to consent in order to continue to the actual survey. No data which can be used to trace or identify specific people other than age and gender was collected. Given the size of the sample and the nature of the recruitment, identifying a participant based on that

information was deemed impossible. The data will not be used for any other research purposes than the current study. All participants were given an email address to which they were advised to send inquiries regarding the study, should they want access to the finished article. This fulfils the ethical requirements of information, consent, confidentiality, utilisation and openness, put forward by the Swedish research council.

Results and discussion

Survey 1

The first survey contained 54 items, whereof 16 were the translated SDO-7. A 16-item prototype scale was constructed based on the mean, standard deviation and item-similarity to original scale. Some items had a score which was very close to the mid-point of the scale but were not included, either because A, their corrected item-total correlation was too low, indicating that they didn't actually measure what the rest of the scale measured, or B, the formulation of the item was problematic, meaning that it was semantically too different from the SDO-7. The semantic difference was of importance in order to keep the SDO-D and SDO-E sub dimensions intact. The descriptive statistics of all 54 items can be found in Table.1 in appendix A. The new scale selected from the neutralised items was named SDO2.0. To investigate if SDO2.0 produced a higher mean than the original SDO scale a paired sample *t*-test was conducted between the two scales and their subscales. The mean was significantly higher for SDO2.0-D ($M = 3.07, SD = 1.23$) compared to SDO-D ($M = 2.27, SD = 1.03$), $t(29) = 6.71, p < .001$. The difference in means was even greater for SDO2.0-E ($M = 3.70, SD = 1.30$) compared to SDO-E ($M = 2.05, SD = 0.90$), $t(29) = 8.88, p < .001$. These results indicate that the discrepancy was larger between the two means of the two sub-scales. A Pearson's correlation revealed a very strong significant correlation between the two scales, $r(28) = .88, p < .001$. This indicates that the items chosen for the second scale were closely related to the original scale.

Survey 2

In the second survey, two criterion variables were included and measured against the SDO-2.0 scale. This was a validity check to ensure that the new scale would actually correlate with criterion variables. Two previously tested criterion variables were used. Punitiveness, which measures a participants affinity for harsher treatment of criminals as a D-criterion (Ho et al., 2015) and political orientation (left/right-scale) on as an E-criterion (Dimdins et al.,

2016). Pearson’s correlations were estimated to compare the different dimensions with the criterion variables. The results can be seen in Table 2.

Table 2. Correlation between the different scales and criterion variables.

Dimension	Punitiveness	Politics
SDO	.53*	.73*
SDO2.0	.56*	.83*
SDO-D	.52*	.62*
SDO2.0-D	.39	.65*
SDOE	.45*	.72*
SDO2.0-E	.62*	.86*

*Note.** Correlation is significant at the 0.01 level.

The SDO2.0-D did not reach significance when correlated with a D-criterion.

Survey 3

The third survey was a last attempt to improve the items of the scale for the final survey. The results from the previous surveys indicated discrepancy between the subscales of the SDO2.0 scale. The anti-egalitarianism subscale’s mean improved more than the dominance subscale. Additionally, the SDO2.0-D failed to reach significance when correlated to a D-criterion. In light of this, 10 new dominance oriented items were included, as well as 7 new anti-egalitarian oriented items. The purpose of these new items was as such to make the discrepancy between the means of the two dimensions smaller by increasing the SDO2.0-D score, and to amend the potential problem regarding predicting d-criteria significantly. A principal component analysis was conducted to see if the original scale and the new scale could produce the 4 factors that were produced in the original study (Ho et al., 2015). The pattern matrix for the original scale can be seen in Table 3, and the pattern matrix for SDO2.0 can be seen in Table.4.

Table 3. Pattern matrix of SDO-7

Type	PC1	PC2	PC3	PC4
Pro-trait D	-.12	.46	.16	.60
Pro-trait D	.09	.68	.17	.00
Pro-trait D	.20	.84	-.01	-.09
Pro-trait D	-.48	.70	.33	.02
Con-trait D	.78	.00	-.13	.32
Con-trait D	.44	-.18	.38	.13
Con-trait D	.17	.91	-.67	.10
Con-trait D	.02	-.03	.06	.90
Pro-trait E	.06	-.02	.83	.01
Pro-trait E	.58	-.29	.07	.43
Pro-trait E	.17	.39	.31	.26
Pro-trait E	-.01	-.18	.90	.13
Con-trait E	.60	.16	.41	-.13
Con-trait E	.48	.20	.53	-.27
Con-trait E	.91	.14	-.00	-.17
Con-trait E	.11	.16	.53	.32

Note. Rotation method was Promax with Kaiser Normalisation.

Table 4. Pattern matrix of SDO2.0.

Type	PC1	PC2	PC3	PC4
Pro-trait D	-.15	.38	.79	-.08
Pro-trait D	.25	-.29	.88	-.04
Pro-trait D	.48	.31	.27	-.14
Pro-trait D	.74	-.19	.21	.26
Con-trait D	.67	.23	.01	.04
Con-trait D	.04	.01	-.05	.99
Con-trait D	-.12	.53	.54	.32
Con-trait D	-.14	.96	-.02	.01
Pro-trait E	1.00	-.39	-.01	.09
Pro-trait E	.95	-.09	.01	.01
Pro-trait E	.76	-.24	.23	-.14
Pro-trait E	.61	.23	.28	-.02

Con-trait E	.82	.30	-.20	.06
Con-trait E	.76	.37	-.11	-.10
Con-trait E	.78	.44	-.24	.02
Con-trait E	.59	.25	.24	-.10

These results are contradictory to previous research. Ho et al. (2015) managed to produce four distinct factors, based on substance (dominance/anti-egalitarianism) and direction (Pro-trait/Con-trait). Surprisingly, neither the original SDO-7 nor the SDO2.0 produced the expected 4-factor solution. Some vague patterns may be discerned, but the dimensions are not stable. No better solution was found using the alternative items introduced in this survey, as such, the SDO2.0 scale remained unchanged. A decision was made to include more items for the final survey in the weaker dimension of the scale (SDO2.0-D) to raise the mean further and to land on a more stable 4-factor solution. This led to a small alternation of the SDO2.0, one item was replaced in the con-trait d-dimension due to poor performance. The new and final scale was named SDO3.0 and can be found in appendix B.

Survey 4

Exploratory analysis of the final sample revealed outliers and normality issues with the original and the new scale. This was expected because of the previously mentioned floor effect. The normality statistics are summarised in Table 5.

Table 5. Normality data on SDO scores

Scale	Skewness	Kurtosis	No. Outliers
SDOtotal	1.37	1.12	13
SDO3.0total	0.90	0.09	3

Both kurtosis and skewness was closer to 0 in the new scale and the number of outliers dropped significantly. A decision was made to keep the outliers in the main result. This might sound counterintuitive and unconventional, but since the original scale produced more outliers than the new scale, it would be strictly beneficial for the study's purpose to simply remove the problematic data since this would make the difference between the two scales greater. Additionally, the problem formulated was a floor effect, with not enough participants yielding higher scores. Additional testing with outliers removed yielded similar results. All reported testing is as such with outliers included.

Internal reliability was assessed using Chronbach's alpha for the original scale, the new scale and their subscales. The alpha was slightly higher for the new scale. The results can be seen in Table 6.

Table 6. Internal reliability on the two scales and their subscales

Scale	No. Items	Chronbach's alpha
SDOtotal	16	0.92
SDO3.0total	16	0.93
SDO-D	8	0.85
SDO3.0-D	8	0.86
SDO-E	8	0.86
SDO3.0-E	8	0.89

The first hypothesis was a prediction that the new scale would produce a higher mean than the original scale. This mean would also be closer to the mid-point of the Likert-scale, which in this case is 4. This hypothesis was tested with paired samples *t*-test. Three tests were conducted, comparing the means of the two subscales and the total score with the scores of the new scale. The participants scored significantly higher on SDO3.0-D ($M = 2.60, SD = 1.32$) compared to SDO-D ($M = 1.86, SD = 1.02$), $t(244) = -16.76, p < 0.001$. The same pattern can be seen for SDO3.0-E ($M = 2.95, SD = 1.39$) compared to SDO-E ($M = 1.86, SD = 1.03$), $t(244) = -20.97, p < 0.001$. The final test revealed a significant total difference between the scales, SDO3.0total ($M = 2.78, SD = 1.28$) and SDOtotal ($M = 1.86, SD = 0.97$), $t(244) = -22.94, p < 0.001$. The results supports the first hypothesis but also the second. The second hypothesis predicted a greater variance due to the variance of social desirability. This can clearly be seen in the results above where the standard deviation has increased by around 30% for both the subscales and the total score. In order to statistically test the second hypothesis, a modified Morgan-Pitman test was conducted (Wilcox, 2015). The HC4 estimator technique was used. This is basically a special case of regression analysis which is tested against the assumption that the difference between two dependent variances equals 0 (homogeneity of variance). The technique is especially robust against handling outliers, which are kept in the present analysis. The result was highly significant, $t(244) = 9.80, p < 0.01$.

The difference between the two new sub-scales, SDO3.0-E and SDO3.0-D, also shrunk compared to the SDO2.0. As such, the discrepancy between the two subscales detected in the surveys 1-3 was amended.

The third hypothesis predicted that correlations between the subscales and criterion variables would be greater when compared to the original scale. The results of Pearson's correlation for the subscales and sum total can be seen in Table 7.

Table 7. Correlations between the different subscales and the four criterion variables

Scale	Politics r	Meritocracy r	Exclusionism r	Dark Triad r
SDO-D	.54**	.52**	.50**	.27**
SDO3.0-D	.63**	.60**	.57**	.35**
SDO-E	.65**	.63**	.47**	.27**
SDO3.0E	.63**	.64**	.52**	.22**
SDOtotal	.63**	.61**	.51**	.29**
SDO3.0total	.67**	.66**	.59**	.38**

Note. ** Correlation is significant at the 0.01 level. (Politics, Meritocracy = Criterion-E, Exclusionism, Dark Triad = Criterion-D)

The SDO3.0-D Subscale outperformed the SDO-D scale numerically on all criterion variables. However, the SDO-E scale outperformed the SDO3.0-E scale on politics and dark triad. Utilising the sum total for the scales yielded better results in some cases, especially for the SDO3.0 Scale. SDO3.0total outperformed the original scale in all regards, regardless of whether the subscale and criterion was a match or not. Utilising a modified Fisher's R-to-Z for dependent samples via an online calculator (Lee & Preacher, 2013), the two scales' correlations were examined to see if the new scale significantly outperformed the old scale. Comparisons were made between the sub-dimension and their corresponding criteria, as well as the total score for all criteria variables. The new D-dimension outperformed the old D-dimension, while no significant result were reached for the E-dimension. The new total score was significantly better across the all criterion variables. The results are displayed in Table 8.

Table 8. The results of the modified Fisher's R-to-Z

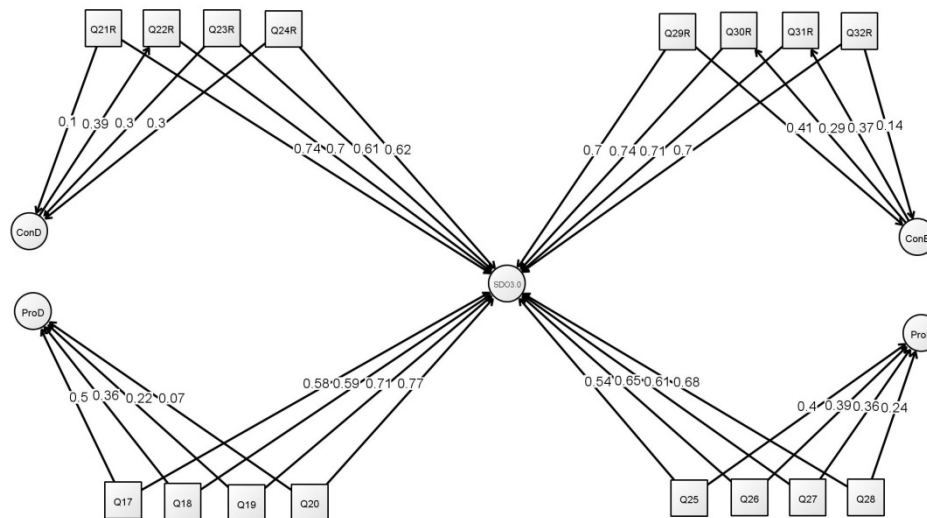
Criterion	z (Dimensions)	z (Total)	Best predictor
Politics	-.67	1.71*	SDO3.0Total
Meritocracy	-0.33	2.10*	SDO3.0Total
Exclusionism	2.43**	3.09**	SDO3.0Total
Dark Triad	2.43**	3.05**	SDO3.0Total

Note. The testing was conducted between the scales' corresponding dimensions (Dimension column) and scales' total score (Total column). ** Correlation is significant at the 0.01 level. * Correlation is significant at the 0.05 level.

The results from the Pearson's correlation indicated that the sub-dimensions of the new scale did not behave in line with previous research. In Table 3 above, it is clear that the original scale's strongest correlation matches with the appropriate criterion, yet this didn't happen for the SDO3.0-scale. Both the SDO-D and SDO3.0-D sub-dimensions should produce the two most powerful correlations when predicting Political Exclusionism. While this is true for the original scale, instead it was the combined score of SDO3.0-D and SDO3.0-E (SDO3.0total) that produced the superior correlation. An exploratory factor analysis was conducted using Onyx (Von Oertzen et al., 2015) to see if there was any distinction between the proposed factors. Analysis revealed that the four factors strongly correlated with each other and made no distinction between the two substantive dimensions. This explains why the SDO3.0total outperformed its own sub-dimensions. The factor ProD which would be the pro-trait dominance factor, should not correlate strongly with ConE, which is the Con-trait anti-egalitarianism factor. The correlation was strong and this indicates that these two factors that are supposed to be distinct could not be statistically distinguished. The results can be seen in Figure 1 in appendix A.

In the previous analysis, the four distinct sub-dimensions or factors could be observed. However, the general SDO-factor is not present in such a model. This general, overarching factor would correspond to the SDO3.0total factor. By including so called bi-factors, comparisons can be made between each individual item and the factor they belong to but also how the item relate to the overarching general factor. This gives us information regarding whether specific items load to their respective factor and also how much of the variance that each item contributes to the specific factor and the overarching general factor. Utilising a model with SDO3.0total set as a general factor, the four different factors were tested with a bi-factor analysis. The model can be seen in Figure 2, while Table 9 contains more details about the estimations of the model.

Figure 2. Confirmatory bi-factor analysis



Note. Standardised estimates between variables. Bi-factor set at SDO3.0.

Table 9. Results of bi-factor analysis

Factor	Item	Loadings std	<i>z</i>	<i>p</i>
ProD	Q17	.50	3.90	<.01
ProD	Q18	.36	3.48	<.01
ProD	Q19	.22	2.92	<.01
ProD	Q20	-.07	-1.08	.31
SDO3.0	Q17	.58	21.92	<.01
SDO3.0	Q18	.59	22.24	<.01
SDO3.0	Q19	.71	27.89	<.01
SDO3.0	Q20	.77	33.19	<.01
ConD	Q21R	.09	1.28	.20
ConD	Q22R	.39	3.80	<.01
ConD	Q23R	.30	3.17	<.01
ConD	Q24R	.30	3.37	<.01
SDO3.0	Q21R	.74	30.65	<.01
SDO3.0	Q22R	.70	27.89	<.01
SDO3.0	Q23R	.61	23.02	<.01
SDO3.0	Q24R	.62	23.39	<.01
ProE	Q25	.40	4.80	<.01
ProE	Q26	.39	5.01	<.01
ProE	Q27	.36	4.35	<.01
ProE	Q28	.24	3.21	<.01
SDO3.0	Q25	.54	20.82	<.01
SDO3.0	Q26	.65	24.37	<.01
SDO3.0	Q27	.61	22.71	<.01
SDO3.0	Q28	.68	25.70	<.01
ConE	Q29R	.41	5.03	<.01

ConE	Q30R	.29	4.24	<.01
ConE	Q31R	.37	4.91	<.01
ConE	Q32R	.14	1.97	.04
SDO3.0	Q29R	.69	26.51	<.01
SDO3.0	Q30R	.29	30.15	<.01
SDO3.0	Q31R	.71	27.49	<.01
SDO3.0	Q32R	.70	27.20	<.01

The loadings here illustrate the difference between the sub-factors and the total dimension of SDO3.0. Item Q20 and Q21R were not significant and had very low loadings on their respective sub-factor, yet they reached significance and correlated strongly with the total SDO-dimension. This helps explain why the new scale's sub-factors didn't produce results in line with previous research, while the SDO3.0total managed to outperform the original SDO7 in predicting criterion variables.

To test if each of the sub-factors actually contributed to the model fit, the model was tested by eliminating one of the four sub-factors at a time. By doing so, contribution of each sub-factor could be estimated. For example, if model fit improves or remains the same when a factor is removed then that specific factor is not supported and it can be questioned whether it should be part of the model. This was not the case however, the model fit actually decreased when any factor was eliminated. The result of the model fit (chi-squared test) is found in Table 10.

Table 10. Chi-squared test for model fit

Missing	Full model χ^2	-1 Factor χ^2	Difference χ^2	<i>p</i>
ProD	152.04	179.26	27.21	<.01
ConD	152.04	170.79	18.75	<.01
ProE	152.04	181.62	29.56	<.01
ConE	152.04	186.16	34.12	<.01

Unexpected findings

By dichotomising the sample after political orientation, a powerful skewness to the political left was found. Only 28 of the participants scored over 4 (which would be in the center on a left- to right-scale). Rerunning the dependent *t*-test on the two groups showed a much more powerful increase in means for the political right when comparing SDOtotal to SDO3.0total. Paired samples *t*-test revealed a significant higher score for participants who scored above 4 on politics, SDO3.0total ($M=4.90$, $SD =1.05$) compared to SDOtotal ($M =3.38$, $SD =1.09$), $t(27) =11.02$, $p < 0.01$. For the participants who scored below 4 on the

political scale the increase of the mean was smaller, SDO3.0total ($M=2.47$, $SD=1.01$) compared to SDOtotal ($M=1.65$, $SD=0.75$), $t(206)=21.19$, $p<0.01$. The increase of the mean for the political right was 1.52 while the political left increased with 0.82. For the full sample, the difference in means between the two groups was 0.92.

General discussion

In this study, an evaluative neutralisation was conducted on the SDO-7 scale to produce a new scale that would have greater variance, higher mean and as such a reduced floor effect. Additionally, the aim was to make the new scale superior to the original scale in regards to correlation to criterion variables. The underlying assumption was that a scale which suffered less from social desirability would display a better construct validity. This process included three different surveys to arrive at a third and final scale, named SDO3.0. The new scale did indeed produce a higher mean, greater variance, it suffered less from floor effects and because of this yielded better normality, as measured by skewness, kurtosis and number of outliers in the sample. Additionally, it produced somewhat stronger correlations to the four different criterion variables included, while only half of them (Criterion-D) managed to reach statistical significance. Measures of internal consistency were equal across the original scale and the SDO3.0.

Sub-dimensions of the new scale

The SDO3.0 scale failed to replicate the original four factors that Ho et al. (2015) produced. While the confirmatory factor analysis indicated that four factors were indeed present, these are clearly empirically driven and not theoretically driven, which is problematic since factor analyses are dependent on theory driven assumptions. If the factors were soundly based in SDT, which the SDO-7's factors appear to be, then the different sub-dimensions would correspond to their given criterion-variable. They did not, however, which would indicate that the current four-factor solution is inherently different from the SDO-7, which managed to produce sub-dimensions that actually corresponded to the correct criterion-variables (Ho et al., 2015). Sub-dimension Pro-D and Con-D both contained one item each (Q20 and Q21r) which failed to reach significance for their given factor. These items added nothing to their respective dimension, yet were more than adequate in providing variance for the total-dimension. Another possible explanation would be that the criterion-variables chosen for this project were faulty. Ethical considerations eliminated certain criteria, such as support for torture as a D-criterion. This type of research has been conducted before but arguably it is

more reasonable to use another fitting D-criterion which is less likely to cause distress to the participant. Relevance to the Swedish political climate was also considered, which is why political exclusionism was chosen over classic anti-black racism, which has been a common criteria in previous SDO research (Ho et al., 2015; Pratto et al., 1994). However, the single most important rationale behind choosing the different criteria was theory underlying SDT. Political exclusionism, which is a textbook example of a D-criterion, was chosen instead of the modern racism scale, which has certain items which are dangerously close semantically to an E-criterion. This paired with the fact that the original scale's sub-dimensions managed to differentiate between the criteria suggests that the current four-factor model is at fault, not the criterion-variables. The original study by Ho et al. (2015) was massive in comparison to this study. Over 2500 participants were drawn from Amazon M-turk and social science research platforms, and tested against 17 criterion-variables. As mentioned in the introduction, the two sub-dimensions are actually highly related and can be used to predict both corresponding criteria and the mismatched criteria. When using SDO-D to predict three different D-criteria, SDO-D was a stronger predictor in 28 cases, it was equally strong as SDO-E in 11 cases and in 1 case SDO-E actually outperformed SDO-D (Ho et al., 2015).

The most probable culprit for the four-factor solution is the evaluative neutralisation process. It quickly became apparent that the SDO-E subscale improved more dramatically than the SDO-D subscale, based on the *t*-test conducted in the first survey. Measures were taken to include more items to balance this discrepancy out but it was hard to improve on many of the first items produced for the SDO2.0-scale. There is a reason for this. The SDO-E dimension measures how unwilling a person is to spend resources on promoting equality, while the SDO-D dimension measures how willing a person is to dominate an outgroup (Ho et al., 2015). It is easier to formulate an item that is more neutral when one discusses egalitarianism compared to dominance. For example, it is easier to argue that whomever earns money ought to be able to keep it for themselves, compared to suggesting that some people have the right to forcefully dominate other groups. One possible explanation is that the neutralisation process took a wrong turn while processing the SDO-D dimension, resulting in an accidental semantic approach to the SDO-E dimension. Ideally, to neutralise the scale, one would like to avoid words such as dominance because of the negative valence which characterise such a word. However, the subscale dimension seems to suffer for it, since no fitting theme was used to replace it. In the original scale, words such as dominance and dominance related phrases such as “groups at the top” and “groups at the bottom” are

frequently used. In fact, out of the 8 items measuring dominance, 6 of them uses this kind of terminology (Ho et al., 2015) This kind of repetitiveness or thematic item construction is very helpful when distinct factors are to be constructed. It is also the very same theme that made the socially desirable answer obvious. This theme is not present in previous SDO-scales and makes its debut with the SDO-7, which was specifically designed to work in such a dichotomised manner (Ho et al., 2015; Pratto et al., 1994). Perhaps then, a more stable distinction between the new scale's sub-dimensions could be produced by using a more subtle dominance-theme. The beauty with SDO is that it can predict generalised prejudice extremely well without ever explicitly giving any hints as to what the difference comparison groups are. All you have to work with is vague formulations such as "groups at the top" and "groups at the bottom". If one would give an example group which is more socially acceptable to blame, for example convicted criminals compared to people who suffered disabilities through accidents, a more evaluative language could be kept. It is however unlikely that the scale would retain its validity in predicting generalised prejudice if such a solution was to be implemented.

The four-factor solution in this study is inherently different from the one found in the original study (Ho et al., 2015). Keep in mind that all previous SDO-scales except for the SDO-7 never used the four-factor model. It is reasonable to assume that the new scale measures the same thing as the SDO-7, since the two scales correlate to a very high degree, while they can produce similar predictions with the different criterion variables chosen. Additionally, the entire reasoning behind dividing into two substantive dimensions was because by doing so, more accurate predictions could be made (Ho et al., 2015). When considering that the total score of new scale performed better when correlating to criterion variables, this limitation seems less significant.

Limitations - political skewness and representativeness

The sample was heavily skewed in regards to political orientation. The measure used here is a composite of three different measures, namely social political orientation, economical political orientation and an overall measure of political orientation. The score on this Likert-scale also ranged from 1 to 7. The mid-point is as such 4. Only 28 participants scored higher than a 4. This doesn't seem representative of the political climate in Sweden right now, considering we have a parliament with a majority belonging to the political right (Riksdagen, 2018). It would also seem as if the evaluative neutralisation had a much greater

effect on people who scored above 4 on political orientation. The difference between the means in the old scale and the new scale was nearly twice as large when compared to the difference between the means for the participants who scored below 4 on political orientation. There was never any initial hypothesis regarding such a phenomenon. The reason or reasons for this discrepancy is hard to theorise about considering the lack of data. Any explanation would be pure conjecture. To clarify, drawing conclusions on the underlying reasons for the discrepancy would be pure conjecture. There is however a very sizable difference between the political left and the political right in the sample. This discrepancy can still be discussed in terms of what it meant for the current study. This undoubtedly affects the results and therefore the generalizability but keep in mind that the results were weaker for the participants leaning towards the political left, which were in a clear majority. There is no reason to believe that all the improvements demonstrated in this study depended on the skewness of the sample. The results regarding the discrepancy between the political left and right indicates quite the opposite, the skewness seemed to have a deflating effect on the results rather than an inflating one. Arguably, the hypotheses would have found greater support in a more politically normal distribution. Another thing of note is the results for the SDO-7. The mean was very low (1.86) compared to the mean found in the original study (2.46) and is more comparable to the results found in the metaanalysis for the SDO-6 (1.82) (Fischer et al., 2012; Ho et al., 2015). There is no data in these articles regarding the distribution in regards to political orientation, which is unfortunate since it could've offered an explanation to the discrepancies seen in these results.

Data collection was severely hindered by the pandemic. Data was collected via postings in various groups on Facebook, as well as by recruitment via QR-code handed out to people (the interviewer was of course wearing a face mask and a visor while abiding to social distancing rules). The response frequency was low, which affects the representativeness of the sample. This became particularly problematic for the final survey, where over 1400 potential participants actually consented to the survey and as such got registered on Qualtrics, while only 245 actually finished the survey. It would seem that a lot of people saw the items, what they were about, and decided to withdraw their consent. The SDO-scale is inherently difficult to answer, which has been made clear through participants criticising the construct's vagueness in not explicitly stating which groups that are being discussed as "inferior". Ten emails with demands of clarification were received to the email that was put forward in the debrief. The Scale is also quite unpleasant for some participants, since it is measuring a trait that is highly socially undesirable. This fact offers yet another important reason to why

evaluative neutralisation potentially could benefit SDO research significantly. For the purpose of comparison, the original SDO-7 scale had to be included in this study. It is not hard to imagine however, that the response rate would go up if the original scale was excluded.

Future research

This study has indicated quite some potential with the evaluative neutralisation process within the SDO field. In the past, there has been a lack of focus of social desirability within the field, which becomes quite apparent if you simply read the items in any of the previous SDO-scales. The socially desirable answer is very apparent. Social desirability is difficult to deal with and it can heavily influence the results and thus the conclusions drawn from said results. With the floor effect partly eliminated, the number of outliers dropped from 13 to 3. That is another 10 participants worth of data that can be saved, with this very simple method. Since the new scale was able to produce stronger correlations than the original, an argument could be made for that the lessened burden of social desirability improved construct validity.

Future research should focus on stratification of the samples to make sure that more participants from the political right are included. Since the response rate was so low, a suggestion is made here to exclude the original SDO-7 scale and instead use past research as a reference point when comparing new and improved scales. The reason for this is the social desirability issue which seems to affect some participants. Some participants seem to be rather sensitive to the items in the SDO-7. As such, response bias might be reduced if such items are not presented to the participant. The evaluative neutralisation process should also focus on a more thematic approach, to make it more relatable to the SDO-7, given that a 4-factor solution is indeed a goal. One possible and very plausible explanation for the present study's failure to replicate a theory driven 4-factor model as discussed above was the semantic approach of SDO-D to SDO-E. By using a thematic approach, this problem could be avoided and it would help creating more distinct factors. Another phenomenon that should be investigated more thoroughly is the discrepancy by the participants who rated themselves as political left and political right. If stratification can lead to a more representative sample, it would be prudent to try to understand why the mean increases so much more for the political right. By including a social desirability scale, one could measure if the political right suffers more from social desirability. The political left is after all characterised by egalitarianism, which is a social norm (Hurwitz & Peffley, 1992). This might also be an explanation to why the political right

is more heavily affected.

Conclusion

The evaluative neutralisation process suggested by Bäckström and Björklund (2013) has indeed proven to be a very helpful tool for improving the SDO-scale. If you take into account that the simple process improves the scale's validity, normality, that it counteracts the floor effect and managed to increase predictive power, perhaps it is actually fitting to call it a “miracle cure”, even though the authors themselves would be reluctant to do so. This is not a claim that the evaluative neutralisation completely fixes the floor effect or skewness, but it does offer a sizable reduction to the problem of social desirability. It is a very simple process and with enough time and resources a much more stable scale with working sub-dimensions can be developed.

References

- Akrami, N., & Ekehammar, B. (2006). Right-wing authoritarianism and social dominance orientation: Their roots in Big-Five Personality Factors and facets. *Journal of Individual Differences, 27*(3), 117–126. <https://doi.org/10.1027/1614-0001.27.3.117>
- Bäckström, M., & Björklund, F. (2013). Social desirability in personality inventories: Symptoms, diagnosis and prescribed cure. *Scandinavian Journal of Psychology, 54*(2), 152–159. <https://doi.org/10.1111/sjop.12015>
- Bäckström, M., Björklund, F., & Larsson, M. R. (2014). Criterion validity is maintained when items are evaluatively neutralized: Evidence from a full-scale five-factor model inventory. *European Journal of Personality, 28*(6), 620–633. <https://doi.org/10.1002/per.1960>
- Bäckström, M., & Björklund, F. (2020). The properties and utility of less evaluative personality scales: Reduction of social desirability; increase of construct and discriminant validity. *Frontiers in Psychology, 11*, 1–15. <https://doi.org/10.3389/fpsyg.2020.560271>
- Dimdins, G., Sandgren, M., & Montgomery, H. (2016). Psychological variables underlying political orientations in an old and a new democracy: A comparative study between Sweden and Latvia. *Scandinavian Journal of Psychology, 57*(5), 437–445. <https://doi.org/10.1111/sjop.12314>
- Duckitt, J., Wagner, C., du Plessis, I., & Birum, I. (2002). The psychological bases of ideology and prejudice: testing a dual process model. *Journal of Personality and Social Psychology, 83*(1), 75–93.
- Dwight, S. A., & Donovan, J. J. (2003). Do warnings not to fake reduce faking? *Human Performance, 16*(1), 1–23. https://doi.org/10.1207/S15327043HUP1601_1

- Dyrendal, A., Kennair, L. E. O., & Bendixen, M. (2021). Predictors of belief in conspiracy theory: The role of individual differences in schizotypal traits, paranormal beliefs, social dominance orientation, right wing authoritarianism and conspiracy mentality. *Personality and Individual Differences, 173*. <https://doi.org/10.1016/j.paid.2021.110645>
- Fischer, R., Hanke K., & Sibley C. G., (2012). Cultural and institutional determinants of social dominance orientation: A cross-cultural meta-analysis of 27 societies. *Political Psychology, 33*(4), 437–467.
- Forchtner, B. (2019). Climate change and the far right. *WIREs: Climate Change, 10*(5), N.PAG.
- Garcia, D., Persson, B. N., Nima, A. A., Brulin, J. G., Rapp-Ricciardi, M., & Kajonius, P. (2018). IRT analyses of the Swedish dark triad dirty dozen. *Heliyon, 4*(3). <https://doi.org/10.1016/j.heliyon.2018.e00569>
- Gittelman, S., Lange, V., Cook, W. A., Frede, S. M., Lavrakas, P. J., Pierce, C., & Thomas, R. K. (2015). Accounting for social-desirability bias in survey sampling: A model for predicting and calibrating the direction and magnitude of social-desirability bias. *Journal of Advertising Research, 55*(3), 242–254. <https://doi.org/10.2501/JAR-2015-006>
- Grina, J., Bergh, R., Akrami, N., & Sidanius, J. (2016). Political orientation and dominance: Are people on the political right more dominant? *Personality and Individual Differences, 94*, 113–117. <https://doi.org/10.1016/j.paid.2016.01.015>
- Ho, A. K., Sidanius, J., Kteily, N., Sheehy-Skeffington, J., Pratto, F., Henkel, K. E., Foels, R., & Stewart, A. L. (2015). The nature of social dominance orientation: Theorizing and measuring preferences for intergroup inequality using the new SDO₇ scale. *Journal of Personality and Social Psychology, 109*(6), 1003–1028. <https://doi-org.ludwig.lub.lu.se/10.1037/pspi0000033.supp> (Supplemental)
- Huang, L.-L., & Liu, J. H. (2005). Personality and social structural implications of the situational priming of social dominance orientation. *Personality and Individual Differences, 38*(2), 267–276. <https://doi-org.ludwig.lub.lu.se/10.1016/j.paid.2004.04.006>
- Hudson, S. T. J., Cikara, M., & Sidanius, J. (2019). Preference for hierarchy is associated with reduced empathy and increased counter-empathy towards others, especially out-group targets. *Journal of Experimental Social Psychology, 85*. <https://doi.org/10.1016/j.jesp.2019.103871>
- Hurwitz, J., & Peffley, M. (1992). Traditional versus social values as antecedents of racial stereotyping and policy conservatism. *Political Behavior, 14*(4), 395–421. <https://doi.org/10.1007/BF00992042>
- Jylhä, K., Akrami, N. (2016) Social dominance orientation and climate change denial: The role of dominance and system justification. *Personality and Individual Differences, 2015*:108. doi:10.1016/j.paid.2015.05.041.

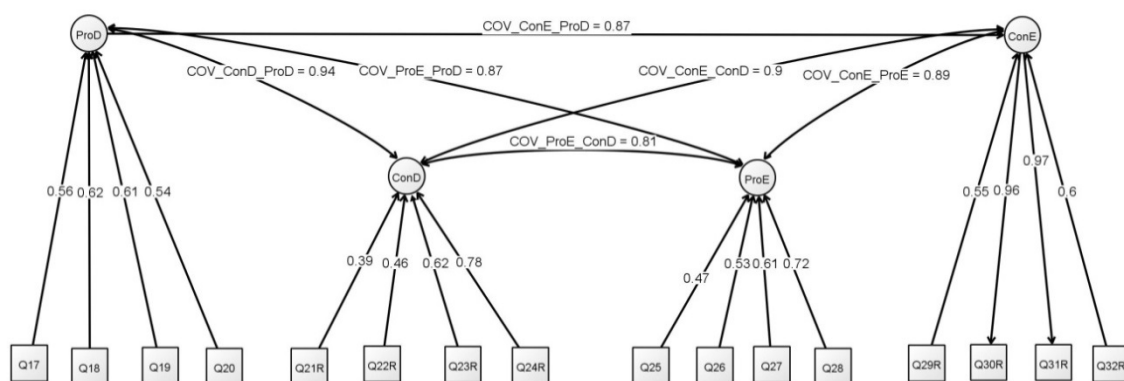
- Kerr, J. R., & Wilson, M. S. (2021). Right-wing authoritarianism and social dominance orientation predict rejection of science and scientists. *Group Processes & Intergroup Relations*, 24(4), 550–567. <https://doi.org/10.1177/1368430221992126>
- Lee, I. A., & Preacher, K. J. (2013). Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Available from <http://quantpsy.org>.
- Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, 67(4), 741–763. <https://doi-org.ludwig.lub.lu.se/10.1037/0022-3514.67.4.741>
- Riksdagen. (2018, September 17) *Den nya riksdagen efter valet*. <https://www.riksdagen.se/sv/valet-2018/den-nya-riksdagen-efter-valet/>
- Roy, S., Neumann, C. S., Jones, D. N., Gari, A., & Šram, Z. (2021). Psychopathic propensities contribute to social dominance orientation and right-wing authoritarianism in predicting prejudicial attitudes in a large European sample. *Personality and Individual Differences*, 168. <https://doi.org/10.1016/j.paid.2020.110355>
- Shi, J., Chen, Z., Wang, X., Teng, F., Yang, Y., & Chen, H. (2021). Dominate others, hurt self: Social dominance orientation predicts depression during the COVID-19 pandemic. *Personality and Individual Differences*, 175. <https://doi-org.ludwig.lub.lu.se/10.1016/j.paid.2021.110710>
- Sidanius, J., Sinclair, S., Pratto, F. (2006). Social dominance orientation, gender, and increasing educational exposure. *Journal of Applied Social Psychology*. 36. 1640 - 1653. <https://doi-org.ludwig.lub.lu.se/10.1111/j.0021-9029.2006.00074.x>
- Šram, Z. (2020). The development of a 9-item scale to measure anti-immigrant attitude toward the middle east refugees. *Romanian Journal of Applied Psychology*, 22(2), 26–32. <https://doi.org/10.24913/rjap.22.2.01>
- Stanley, S. K., Wilson, M. S., Sibley, C. G., & Milfont, T. L. (2017). Dimensions of social dominance and their associations with environmentalism. *Personality and Individual Differences*, 107, 228–236. <https://doi-org.ludwig.lub.lu.se/10.1016/j.paid.2016.11.051>
- Von Oertzen, T., Brandmaier, A. M., & Tsang, S. (2015). Structural equation modeling with Ω nyx. *Structural Equation Modeling*, 22(1), 148–161. <https://doi.org/10.1080/10705511.2014.935842>
- Zimmerman, J. L., & Reyna, C. (2013). The meaning and role of ideology in system justification and resistance for high- and low-status people. *Journal of Personality and Social Psychology*, 105(1), 1–23. <https://doi-org.ludwig.lub.lu.se/10.1037/a0032967>

Wilcox, R. R. (2015). Comparing the variances of two dependent variables. *Journal of Statistical Distributions and Applications* 2, Article number:7.

Wood J. K., Anglim., J. & Horwood, S. (2021). A less evaluative measure of big five personality: comparison of structure and criterion validity. *European Journal of Social Psychology*. <https://doi.org/10.1177/08902070211012920>

Appendix A

Figure 1 Exploratory factor analysis of the SDO3.0



Note. Standardised estimates between variables.

Descriptive Statistics

	N	Min	Max	Mean	Std. Deviation
Ibland måste vissa grupper hållas under kontroll för att få ordning i samhället	30	1	7	3.57	1.906
Det vore bäst om vissa grupper kunde fås att acceptera sin plats i samhället	30	1	5	2.10	1.269
Ett välfungerande samhälle kräver att vissa grupper klättrar högre medan andra stannar på sin plats	30	1	5	2.53	1.525
Ett samhälle flyter på bättre när vissa grupper har mer inflytande än andra	30	1	6	2.70	1.685
En uttalad hierarki ser till att rätt sorts grupp hamnar i beslutsställning	30	1	7	2.67	1.626
Vissa grupper klarar helt enkelt av mindre än vad andra grupper gör	30	1	7	3.33	1.936
Det är OK att starkare grupper tenderar att hamna i en bättre position i samhället än andra	30	1	6	2.80	1.627
Det kan inte undvikas att grupper som har lite sämre förutsättningar får mindre inflytande	30	1	7	3.60	1.793
Det är bara naturligt att vissa grupper klättrar högre än andra	30	1	6	2.97	1.564
Vissa grupper står antagligen över andra grupper för att de är bättre lämpade att lyckas	30	1	5	2.67	1.493

Att tillåta vissa grupper att klara sig bättre än andra medför positiva konsekvenser	30	1	5	2.73	1.337
Lägre samhällsgrupper förtjänar lika stor del av kakan som högre	30	1	6	2.80	1.584
Grupper som presterar mindre bra är lika förtjänta av respekt som de som presterar bättre	30	1	5	1.97	1.351
Ingen enskild grupp borde få ställa och styra	30	1	7	2.77	1.832
Det finns inte någon samhällsgrupp som är mer lämpad att styra	30	1	7	3.30	1.985
Grupper som har hamnat högt i samhället måste lämna plats åt grupper som hamnat lägre	30	1	7	3.80	1.769
Det är samhällets uppgift att se till att grupper som hamnat lågt inte tvingas kvar där	30	1	6	2.33	1.184
Samhället bör sträva bort från alla typer av hierarkiska strukturer	30	1	7	3.90	1.788
Alla samhällsgrupper bör vara helt jämnt representerade i beslutsfattande positioner	30	1	7	3.83	1.877
Inga grupper ska ha mer att säga till om än andra	30	1	7	3.10	1.918
Gruppers lika inflytande behöver inte prioriteras särskilt högt	30	1	7	2.27	1.363

Jämställdhetspolitik skulle behöva tonas ner på vissa områden	30	1	5	2.13	1.196
Det läggs redan ner för mycket tid och resurser på jämställdhetspolitik	30	1	7	2.23	1.524
Det är omöjligt att garantera att alla grupper får likvärdiga utfall i livet	30	3	7	5.33	1.539
Det är inte okej att tvinga fram jämställdhet om grupper riskerar att bli orättvist behandlade	30	1	7	3.80	1.730
Det kan bli orättvist om man påtvingar jämlikhet mellan alla grupper	30	1	7	3.60	1.831
Tyvärr finns det ibland mer akuta saker att ägna sig åt i samhället än just jämställdhet	30	2	7	4.10	1.470
Fokuserar vi enbart på jämställdhet så riskerar samhället att inte utvecklas	30	1	7	3.57	1.813
Vi måste lägga all vår energi på att se till att alla grupper har lika stor chans att lyckas	30	1	7	3.03	1.629
Jämställdhet måste ha företräde före all annan samhällsutveckling	30	1	7	4.27	1.818
Samhället bör lägga mycket resurser på att skapa identiska förhållanden för olika grupper	30	1	7	4.17	1.763
Samhället ska se till att grupper som hamnat högt delar med sig mer än vad de gör idag	30	1	7	3.10	1.709
Samhället behöver garantera att alla grupper får samma ekonomiska förutsättningar	30	1	7	3.23	1.924

Vi måste göra allt i vår makt för att se till att alla grupper har samma chans att lyckas här i livet, oavsett kostnaden	30	1	7	2.97	1.542
Samhället bör erbjuda grupper med svårigheter ett obegränsat antal tillfällen att lyckas	30	1	7	4.50	1.697
Jämställdhet mellan grupper är det viktigaste som finns	30	1	7	3.47	2.224
Andra samhällsfrågor får komma i andra hand tills vi kan garantera lika behandling mellan olika grupper	30	1	7	4.63	1.629
Vissa grupper av människor måste hållas på sin plats	30	1	5	2.40	1.589
Det är nog bra att vissa grupper är överordnade och andra är underordnade	30	1	6	2.37	1.497
Det ideala samhället kräver att vissa grupper är överordnade och att andra är underordnade	30	1	6	2.63	1.712
Vissa grupper är helt enkelt underlägsna andra grupper	30	1	7	2.33	1.688
Underordnade grupper förtjänar att lyckas lika mycket som överordnade grupper	30	1	5	1.67	.994
Ingen grupp borde få dominera samhället	30	1	7	2.67	1.583
Underordnade grupper borde inte hindras från att avancera	30	1	6	1.87	1.306
Att vissa grupper tillåts dominera andra är en dålig idé	30	1	5	2.23	1.194
Vi borde inte eftersträva jämlikhet mellan olika grupper	30	1	4	1.57	1.006

Att garantera alla grupper samma livskvalitet är inte något att sträva mot	30	1	7	1.50	1.167
Det är orättvist att försöka göra alla grupper jämlika	30	1	6	2.57	1.406
Jämlikhet mellan grupper borde inte vara vårt främsta mål	30	1	7	2.80	1.808
Vi borde arbeta för att ge alla grupper samma möjligheter att lyckas	30	1	5	1.73	1.081
Vi bör göra vad vi kan för att alla grupper ska få lika villkor	30	1	5	2.00	1.390
Oavsett vad som krävs så borde vi sträva efter att ge alla grupper samma möjligheter i livet	30	1	6	2.20	1.270
Jämlikhet mellan grupper borde vara ett samhällsideal	30	1	5	2.07	1.285

Appendix B

SDO3.0

Ibland måste vissa grupper hållas under kontroll för att få ordning i samhället

Vissa grupper klarar helt enkelt av mindre än vad andra grupper gör.

Det är bara naturligt att vissa grupper klättrar högre än andra.

Att tillåta vissa grupper att klara sig bättre än andra medför positiva konsekvenser.

Lägre samhällsgrupper förtjänar lika stor del av kakan som högre.

Att vissa grupper dominerar andra är något som samhället till varje pris måste komma tillrätta med.

Det finns inte någon samhällsgrupp som är mer lämpad att styra.

Inga grupper ska ha mer att säga till om än andra.

Det är inte okej att tvinga fram jämställdhet om grupper riskerar att bli orättvist behandlade.

Det kan bli orättvist om man påtvingar jämlikhet mellan alla grupper.

Tyvärr finns det ibland mer akuta saker att ägna sig åt i samhället än just jämställdhet.

Fokuserar vi enbart på jämställdhet så riskerar samhället att inte utvecklas.

Jämställdhet måste ha företräde före all annan samhällsutveckling.

Vi måste göra allt i vår makt för att se till att alla grupper har samma chans att lyckas här i livet, oavsett kostnaden.

Jämställdhet mellan grupper är det viktigaste som finns.

Grupper som har hamnat högt i samhället måste lämna plats åt grupper som hamnat lägre.

SDO-7

Vissa grupper av människor måste hållas på sin plats.

Det är nog bra att vissa grupper är överordnade och andra är underordnade.

Det ideala samhället kräver att vissa grupper är överordnade och att andra är underordnade.

Vissa grupper av människor är helt enkelt underlägsna andra grupper.

Underordnade grupper förtjänar att lyckas lika mycket som överordnade grupper.

Ingen grupp borde få dominera i samhället.

Underordnade grupper borde inte hindras från att avancera.

Att vissa grupper tillåts dominera andra är en dålig idé.

Vi borde inte eftersträva jämlikhet mellan olika grupper.

Att garantera alla grupper samma livskvalité är inte något att sträva mot.

Det är orättvist att försöka göra alla grupper jämlika.

Jämlikhet mellan grupper borde inte vara vårt främsta mål.

Vi borde arbeta för att ge alla grupper samma möjligheter att lyckas.

Vi bör göra vad vi kan för att alla grupper ska få lika villkor.

Oavsett vad som krävs så borde vi sträva efter att ge alla grupper samma möjligheter i livet.

Jämlikhet mellan grupper borde vara ett samhällsideal.

Minexcl

Medlemmarna av etniska minoritetsgrupper som inte är lojala mot vår stat borde bli berövade på medborgarskapet.

Etniska minoritetspartier borde inte tillåtas att vara medlemmar i vår riksdag.

Det finns vissa etniska minoritetsgrupper som borde uppmuntras att lämna landet.

Det finns inget behov för etniska minoritetsgrupper att organisera sig i politiska partier.

Dirty dozen dark triad

Jag tenderar att manipulera andra för att få det jag vill.

Jag har bedragit eller ljugit för att få det jag vill.

Jag har använt mig av smicker för att få min vilja igenom.

Jag tenderar att utnyttja andra för att nå mina egna mål.

Jag tenderar att sakna ånger.

Jag tenderar att inte bry mig om moralen i mina handlingar.

Jag tenderar att vara kall eller okänslig.

Jag tenderar att vara cynisk.

Jag tenderar att vilja att andra beundrar mig.

Jag tenderar att vilja att andra uppmärksammar mig.

Jag tenderar att söka prestige och status.

Jag tenderar att förvänta mig speciella tjänster av andra.

Political orientation (Left-to-right)

- 1) I vilken utsträckning betraktar du dig själv som ekonomiskt vänster (med betoning på minskning av ekonomiska klyftor i samhället, även om det leder till att ekonomiska resurser omfördelas) eller ekonomiskt höger (med betoning på individers rätt att fullt ut avnjuta resultaten av deras ekonomiska framgång, även om det leder till ekonomiska klyftor)?
- 2) I vilken utsträckning betraktar du dig själv som socialt liberal (med betoning på uttryck av individuella fri- och rättigheter över värmandet av traditioner och sociala normer)

eller socialt konservativ (med betoning på värnandet av traditioner och sociala normer över individuella fri- och rättigheter)?

3) Hur skulle du skatta din egen politiska orientering?

Meritocracy

Du ser här ett antal påståenden som mäter hur du tycker att samhället faktiskt ÄR. Indikera på skalan hur väl påståendet stämmer överens med din uppfattning.

- 1) Människor som jobbar hårt blir framgångsrika.
- 2) Om människor jobbar hårt så uppnår de det de vill.
- 3) Om man jobbar hårt så kan man bli framgångsrik oavsett vilket kön, etnicitet eller klass man tillhör.
- 4) Diskriminering hindrar ingen från att nå framgång, om personen arbetar hårt.
- 5) Sverige är ett öppet samhälle där alla individer kan nå hög status genom hårt arbete.
- 6) Framgång i det svenska samhället är lika möjligt för alla individer.