Investigating the MACH-IV with item response theory and proposing the Trimmed MACH*

Article *in* Journal of Personality Assessment · November 2012

DOI: 10.1080/00223891.2012.742905 · Source: PubMed

1 author:

John F. Rauthmann
Bielefeld University
157 PUBLICATIONS **4,733** CITATIONS

SEE PROFILE

Journal of Personality Assessment
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hjpa20

Investigating the MACH-IV With Item Response Theory and Proposing the Trimmed MACH*

John F. Rauthmann [a]
[a] Department of Psychology, Humboldt-Universität zu Berlin, Germany
Version of record first published: 27 Nov 2012.

To cite this article: John F. Rauthmann (2012): Investigating the MACH-IV With Item Response Theory and Proposing the Trimmed MACH*, Journal of Personality Assessment, DOI:10.1080/00223891.2012.742905

To link to this article: http://dx.doi.org/10.1080/00223891.2012.742905

Journal of Personality Assessment, 1–10, 2013

ISSN: 0022-3891 print / 1532-7752 online
DOI: 10.1080/00223891.2012.742905

Routledge
Taylor & Francis Group

Investigating the MACH–IV With Item Response Theory and Proposing the Trimmed MACH*

JOHN F. RAUTHMANN

Department of Psychology, Humboldt-Universität zu Berlin, Germany

The MACH–IV was investigated ($N = 528$) with item response theory to elucidate its psychometric properties and suggest a trimmed version, the MACH*. The core content of the MACH–IV seemed to be cynicism/misanthropy and the MACH* was formed from the 5 most informative and precise MACH–IV items. The MACH* showed good internal consistency and construct and criterion validity comparable to the MACH–IV. The MACH–IV and MACH* measure most precisely at average to above average levels of Machiavellianism. Implications for theory and measurement of Machiavellianism are discussed.

Niccolò Machiavelli (1513/1981), a political advisor to the Medici family in Florence, introduced in his book *Il Principe* (*The Prince*) suggestions for rulers to obtain and maintain power: One should exploit the simple, fallible, and manipulable nature of (wo)men with long-term planned, pragmatic, cunning, and immoral strategies and tactics. His portrait of an effective ruler is incarnated in the personality dimension of *Machiavellianism,* introduced by Christie and Geis (1970) based on their literature analysis and extensive validation of their MACH scales. Since then, there has been considerable interest in Machiavellianism (Fehr, Samsom, & Paulhus, 1992; Jones & Paulhus, 2009; Rauthmann & Will, 2011; Wilson, Near, & Miller, 1996). The "golden standard" of measuring Machiavellianism has become Christie and Geis's (1970) MACH–IV, which is used in almost all studies (Fehr et al., 1992). However, it has yet to be investigated with sophisticated psychometric procedures and has not undergone reworking so far. Given that "Machiavellianism does represent a fundamental theme" (Wilson et al., 1996, p. 295) with ongoing popularity (e.g., Jonason, Webster, Schmitt, Li, & Crysel, 2012) and that developments of new Machiavellianism scales will have to rely on the MACH–IV as *the* benchmark criterion, the MACH–IV will likely continue to play a strong role in future Machiavellianism research. The work reported here elucidates the psychometric properties of the MACH–IV for the first time with item response theory (IRT), provides a trimmed version of the MACH–IV, and guides efforts for developing new Machiavellianism scales.

THE MACH–IV

More than 40 years ago, Christie and colleagues developed the MACH–IV scale with 20 items that are phrased as recommendations, quasi-facts, or statements (e.g., Anyone who completely trusts anyone is asking for trouble). People who endorse such items have been found to (a) think in a cold, strategic, and pragmatic way; (b) have cynical, misanthropic, and negativistic views; (c) be emotionally detached and callous; (d) be agentically motivated (e.g., for money, power, status) rather than communally (e.g., for love, family, harmony); and (e) use duplicity, exploitation, and manipulation tactics to push through their self-beneficial goals (Christie & Geis, 1970; Fehr et al., 1992; Jones & Paulhus, 2009, 2010; Rauthmann, 2011; Rauthmann & Will, 2011; Wilson et al., 1996, 1998). The MACH–IV can be deemed a reliable and valid scale (e.g., Fehr et al., 1992; Jones & Paulhus, 2009; Ramanaiah, Byravan, & Detwiler, 1994), and the body of evidence for its predictive abilities is impressive (for a review, see Fehr et al., 1992). Nonetheless, several problems have been noted, such as (a) response biases, (b) varying and low reliabilities, (c) varying factorial structures, (d) insufficient content and construct validity, and (e) some unclear relations to external criteria (for a review, see Fehr et al., 1992). It is thus not uncommon that the MACH–IV has been modified by either dropping (e.g., Ahmed & Stewart, 1981; Hunter, Gerbing, & Buster, 1982) or rewording items (e.g., Zook & Sipps, 1986)—which hints at problems with the scale.

A particular concern seems to be the dimensionality of the MACH–IV (and the Machiavellianism construct in general). Problems with its factorial structure have been repeatedly noted (e.g., Ahmed & Stewart, 1981; Corral & Calvete, 2000; Hunter et al., 1982; O'Hair & Cody, 1987; Vleeming, 1984; M. L. Williams, Hazelton, & Renshaw, 1975). Panitz (1989) found that it was difficult to assess a clear factorial structure for the MACH–IV, questioned what the scale actually measures, and concluded that the scale has no construct validity. More important, different factorial solutions, ranging from one to five factors, have been proposed (Ahmed & Stewart; Christie & Lehman, 1970; Corral & Calvete; Hunter et al., 1982; Kline & Cooper, 1983, 1984; Kuo & Marsella, 1977; O'Hair & Cody; Vleeming), and factor labels as well as the items in factors differ (see Hunter et al.; Rauthmann, 2011; Rauthmann & Will, 2011; Ray, 1983). This is aggravated by the fact that the dimensionality of the Machiavellianism construct is also conceptually unclear.

A work-around for the situation that (a) there is no single theory (yet) that prescribes a clear factorial structure of Machiavellianism and (b) the MACH–IV does not seem to a show a single, invariant factor structure has been to compute a single Machiavellianism score and treat it as a unitary construct. Indeed, Fehr et al. (1992) noted that "the total score may be more useful in predicting certain behaviors" (p. 109), which favors a

Received June 20, 2012; Revised August 25, 2012.

Address correspondence to John F. Rauthmann, Humboldt-Universität zu Berlin, Institut für Psychologie, Unter den Linden 6, 10099 Berlin, Germany; Email: jfrauthmann@gmail.com

unidimensional position. Whether this is theoretically sensible or not could be debated, but many research designs (especially in applied or organizational contexts) require short scales with less redundant "noise items." The varying factorial solutions and uncertainties regarding the validity of the MACH–IV might be an indication that the scale contains noise items. IRT analyses can provide a fresh look at the MACH–IV and suggest points of improvement, especially regarding specific items. A shorter scale can be constructed that functions as effectively as possible but more efficiently. This work thus examines with IRT what the MACH–IV predominantly captures and suggests a trimmed version, the MACH*, which is ridded of noise items.

Item Response Theory

IRT subsumes a set of mathematical models that probabilistically describe the relation between a participant's response to an item and his or her standing on the measured latent trait (θ) that accounts for that item response (Embretson & Reise, 2000; Morizot, Ainsworth, & Reise, 2007; van der Linden & Hambleton, 1997). For example, a person with a high level of Machiavellianism should more likely endorse the item "Anyone who completely trusts anyone is asking for trouble," whereas a person with a low level of Machiavellianism would be less likely to do so. IRT offers the advantage of computing item and test characteristics contingent on θ and thus regards item or test and person characteristics at once. The item characteristics curve (ICC) or category response curve (CRC) describes the probabilistic relation between a person's response to each category of an item and his or her level of θ (Reeve & Fayers, 2005). The *x*- or θ-axis represents the level of θ (usually from –3 to +3) given in standardized units ($M = 0, SD = 1$), whereas the *y*-axis represents the probability (from 0.00–1.00) of giving a certain category response. The θ-continuum is often divided into the categories of well below average ($\theta < -2$), below average ($-2 < \theta < -1$), average ($-1 < \theta < +1$), above average ($+1 < \theta < +2$), and well above average ($\theta > +2$).

Two parameters are commonly estimated: The threshold parameter *b* indicates the difficulty of an item response, and the discrimination parameter *a* the strength of the relationship between an item and the measured construct (Reeve & Fayers, 2005). Both parameters allow describing an item information curve (IIC). An IIC illustrates the θ range where an item best discriminates among individuals because higher information means more precision with less measurement error (i.e., more reliability). It can be used to identify those items with most information about the latent construct. The *y*-axis indicates the magnitude of information. Due to the additive nature of IICs, they can be summed up across all items to create the scale or test information function (TIF). The TIF indicates the total amount of information that the scale provides across different levels of θ and shows where it is most and least precise. More information, at different levels of θ, corresponds to less measurement error. Measurement precision is defined for different levels of θ in IRT, whereas classical test theory treats it as a constant across the entire scale (which is often unrealistic).

This work used a graded response model (GRM; Samejima, 1969, 1996) for the MACH–IV. It handles polytomous, ordered categories (e.g., attitudinal 5-point Likert-type scales with *disagree–agree* anchors) and lets discrimination parameters vary across items. For each item, one discrimination parameter *a* and $m - 1$ threshold parameters *b* (m = number of response categories) are estimated. In the GRM, item responses are considered as a series of $m - 1$ response dichotomies. A 5-point Likert-type response scale thus produces four threshold parameters $b_{1–4}$: Response Option 1 vs. Options 2, 3, 4, and 5 (b_1); Options 1 and 2 vs. 3, 4, and 5 (b_2); Options 1, 2, and 3 vs. 4 and 5 (b_3); and Options 1, 2, 3, and 4 vs. 5 (b_4).

This Study

This study investigates psychometric properties of the MACH–IV for the first time with IRT analyses. All 20 MACH–IV items are inspected regarding their discrimination parameter *a,* threshold parameters *b,* and IICs to examine which items perform well or poorly. This knowledge can be ideally used to design a streamlined version of the MACH–IV, the MACH*, that is reliable and content valid or, in IRT terminology, precise and informative.

First, discrimination parameters *a* indicate to what extent items are helpful for describing the measured latent construct. Thus, the relevance of each item to Machiavellianism can be evaluated. Second, item threshold parameters $b_{1–4}$ indicate the difficulty of each response option. Negative values are better for measuring low levels of Machiavellianism (i.e., they are easier to endorse for people), whereas positive values are better for measuring high levels of Machiavellianism (i.e., they are more difficult to endorse, unless one is more strongly Machiavellian). Third, IICs indicate how much information each item provides and can thus be used to identify poorly performing items. Items perform poorly if they have a flat curve (low information) or if two or more items share virtually similar information functions (redundancy). Such items can be discarded after carefully inspecting their content and relations to all other items. Fourth, we can evaluate what the MACH–IV actually measures by checking the item content of well-performing items. Fifth, the TIF indicates how precise the total MACH–IV score is at different levels of latent Machiavellianism. It can be used to examine the points of least and highest measurement precision (i.e., reliability). Lastly, IRT information can be used to generate the MACH*, a short (n of items < 10), unidimensional, valid, informative, and reliable version of the MACH–IV that measures similar content but is ridded of noise items that contribute to messy factorial structures. Addressing all these issues can help shed light on what exactly the MACH–IV measures, determine how well its items perform, and suggest recommendations for how to revise it.

Methods

Participants and Procedure

A sample of $N = 528$ (418 women, 101 men; 9 gave no indication) German native speakers with a mean age of about 33 years ($SD = 11$; range = 18–67) provided complete data for this study. The study was hosted as an online study on the platform PSYTESTS (www.psytests.de) for several months and assessed a variety of self-reported personality traits. Although participants were not asked about a current student status, it is safe to assume that the sample is largely based on a community population (evident by the mean age).

TABLE 1.—MACH–IV items (Christie & Geis, 1970).

No.	Item	Scoring	Specific Content[a]
1	The best way to handle people is to tell them what they want to hear.	+	Manipulation tactic (ingratiation)
2	When you ask someone to do something for you, it is best to give the real reasons for wanting it rather than giving reasons that might carry more weight.	–	Manipulation tactic (persuasion)
3*	Anyone who completely trusts anyone is asking for trouble.	+	Suspiciousness, misanthropy, cynicism
4	It is hard to get ahead without cutting corners here and there.	+	Rule-bending, immorality, agency-striving
5	Honesty is the best policy in all cases.	–	Duplicity/dishonesty
6*	It is safest to assume that all people have a vicious streak and it will come out when they are given a chance.	+	Suspiciousness, misanthropy, cynicism
7*	Never tell anyone the real reason you did something unless it is useful to do so.	+	Manipulation tactic (agenda concealment)
8	One should take action only when sure it is morally right.	–	Immorality
9	It is wise to flatter important people.	+	Manipulation tactic ingratiation
10	All in all, it is better to be humble and honest than important and dishonest.	–	Immorality, agency-striving
11	Barnum was very wrong when he said there's a sucker born every minute.	–	Misanthropy, cynicism
12	People suffering from incurable diseases should have the choice of being put painlessly to death.	+	Misanthropy,[b] cynicism, pragmatism
13	It is possible to be good in all respects.	–	Immorality, duplicity/dishonesty, misanthropy, cynicism
14*	Most people are basically good and kind.	–	Immorality, duplicity/dishonesty, misanthropy, cynicism
15	There is no excuse for lying to someone else.	–	Duplicity/dishonesty, immorality, misanthropy, cynicism
16	Most men forget more easily the death of their father than the loss of their property.	+	Agency-striving, misanthropy, cynicism
17	Most people who get ahead in the world lead clean, moral lives.	–	Agency-striving, misanthropy, immorality
18	Generally speaking, men won't work hard unless they're forced to do so.	+	Misanthropy
19*	The biggest difference between most criminals and other people is that criminals are stupid enough to get caught.	+	Misanthropy, cynicism
20	Most men are brave.	–	Misanthropy

Note. + = scored positively, − = scored negatively/reversely. Items marked with an asterisk (*) were used for the trimmed MACH*.
[a]Assigned according to the 46 Machiavellianism content areas from Rauthmann and Will (2011) and Rauthmann (2011).
[b]That endorsing (active) euthanasia for people suffering terribly from an incurable disease indicates a cynical, misanthropic standing must not reflect the author's opinion. Rather, this item was designed in 1970 to likely capture content associated with misanthropy and cynicism. No moral judgment of any form of euthanasia is intended here.

Measures

Machiavellianism was measured with the MACH–IV (Christie & Geis, 1970; Shajek, 2007), using a 5-point Likert-type scale ranging from 1 (*I totally disagree*) to 5 (*I totally agree*). Items can be found in Table 1. Additionally, other scales were used to examine the construct and criterion validity of the MACH–IV and MACH*. These were the Dirty Dozen (narcissism, Machiavellianism, psychopathy: 12 items; Küfner, Dufner, & Back, 2012; Jonason & Webster, 2010), German Machiavellianism Scale (18 items; Henning & Six, 2008), Narcissistic Personality Disorder–17 (NPI–17; narcissism: 17 items; von Collani, 2008), Self-Report Psychopathy Scale–III (SRP–III; psychopathy: 30 items; K. M. Williams, Nathanson, & Paulhus, 2003), Big Five Inventory–Short SOEP Panel (15 items; Schupp & Gerlitz, 2010), and HEXACO Honesty/Humility (16 items; Lee & Ashton, 2004), all to be answered on a 5-point Likert-type scale ranging from 1 (*not like me at all*) to 5 (*totally like me*). Mean scores were computed in each instance. Narcissism and psychopathy were chosen because these form together with Machiavellianism the "Dark Triad of Personality" (Paulhus & Williams, 2002). However, Machiavellianism should show discriminant validity regarding the other two members of the Triad. The Big Five/Six are used because they have been linked to the Dark Triad (de Vries & van Kampen, 2010; Jakobwitz & Egan, 2006; Lee & Ashton, 2005) and Machiavellianism in particular (e.g., Ashton, Lee, & Son, 2000; Jones & Paulhus, 2009). The most consistent findings include negative relations with extraversion, agreeableness, conscientiousness, and honesty and humility (Fehr et al., 1992; Jones & Paulhus, 2009; Paulhus & Williams, 2002). Additionally, eight manipulation tactics frequently used by people high on Machiavellianism (Fehr et al., 1992; Jones & Paulhus, 2009; Wilson et al., 1996) were sampled with two self-generated items each on 7-point frequency-based Likert-type scales ranging from 1 (*never*) to 6 (*almost always*): emotional manipulation (I toy with others' emotions to influence them; I use others' feelings to influence them), ingratiation (I try to flatter important people; I make myself pleasant to important people), persuasion (I use my powers of persuasion to push through my views and interests; I persuade others for my own means), supplication (I let myself be helped by others to get what I want; I pretend to be weak and ask for help to have my wishes fulfilled), intimidation (I use intimidation to push through my interests; I intimidate others to push through my will), self-disclosure (I tell others secrets about me to win favor with them; I tell others intimate things about me to win their favor), betrayal (I betray others when I get the chance; I betray if it cannot be revealed that it was me), and revenge (I get back at others if that has no consequences for me; I get back at others if it cannot be traced back to me).

Data-Analytical Strategy

First, the 20 MACH–IV items were classified according to their content to guide the interpretation of IRT findings. Second, descriptive statistics (*M, SD,* item–total correlations, Cronbach's alpha, mean interitem correlation) as well as discrimination parameters a and threshold parameters $b_{1–4}$ were computed with IRTPRO (2012). Third, IICs of all 20 MACH–IV items and the TIF of the global MACH–IV score were inspected to examine

content coverage, amount of information, and levels of measurement precision. Fourth, a trimmed version of the MACH–IV, the MACH*, was generated based on IRT findings. The MACH* was then also examined with IRT analyses. Additionally, construct and criterion validity was examined by comparing (attenuated and fully disattenuated) correlation patterns of the MACH–IV and MACH*, respectively, across different criteria (narcissism, psychopathy, Big Six, and manipulation tactics). The following more detailed information can be obtained from the author on request: (a) tests for IRT assumptions (unidimensionality, local independence, and item fit), (b) ICCs/CRCs of all 20 MACH–IV items, and (c) the test characteristics curves of the MACH–IV and MACH*.

Results

The MACH–IV

Item content analysis. To investigate MACH–IV item content and inform later steps of item evaluation, each item was assigned to one or more of Rauthmann and Will's (2011) Machiavellianism aspects that they have extracted from a comprehensive literature review on Machiavellianism. This gives a clearer picture of which aspects the items represent (see, for a similar approach, Rauthmann, 2011). Table 1 summarizes the findings. The majority of items captured a negative and cynical view of people and the world. The specific content areas tapped were misanthropy (12 items), cynicism (9 items), immorality (6 items), duplicity/dishonesty (4 items), and agency-striving (4 items). Twelve items tapped more than one content area.

Descriptive statistics. Descriptive statistics are given in Table 2. Mean of the total MACH–IV score was 2.77 (SD = 0.51), with an internal consistency reliability of $\alpha = .82$ and a mean interitem correlation of .19 (r range = –.11–.53). The items with strongest item–total correlations were 3, 6, 7, 14, and 19, and those with the weakest were 4, 9, 12, 13, and 16.

TABLE 2.—Descriptive item statistics.

Items	*M*	*SD*	Item-Total *r*
1	2.24	0.97	.38
2	2.27	1.05	.37
3	**2.63**	**1.20**	**.54**
4	3.52	0.92	.31
5	2.46	1.02	.46
6	**2.74**	**1.18**	**.54**
7	**2.03**	**1.02**	**.60**
8	2.57	1.01	.38
9	2.37	0.98	.27
10	1.81	0.91	.42
11	3.30	1.29	.44
12	3.49	1.27	.29
13	3.14	1.24	.29
14	**2.81**	**1.00**	**.50**
15	3.17	1.06	.35
16	2.52	1.21	.27
17	3.88	0.84	.35
18	2.69	1.00	.35
19	**2.36**	**1.14**	**.49**
20	3.46	0.89	.36

Note. $N = 528$. Item-total r is the correlation of the item with the total scale score.
Bold items were selected for the MACH*. They also had the five highest item-total correlations.

TABLE 3.—Discrimination and threshold parameters of the MACH–IV.

Items	a	b_1	b_2	b_3	b_4
1	0.98	−1.38	0.70	2.34	5.10
2	0.91	−1.32	0.65	2.49	4.19
3	**1.53**	**−1.28**	**−0.00**	**0.93**	**2.25**
4	0.75	−4.92	−2.82	−0.37	3.01
5	1.01	−1.86	0.20	2.04	3.60
6	**1.61**	**−1.50**	**−0.08**	**0.87**	**1.98**
7	**1.98**	**−0.53**	**0.82**	**1.65**	**2.83**
8	0.83	−2.63	0.07	2.23	3.87
9	0.68	−2.09	0.26	3.12	6.68
10	0.96	−0.32	1.72	3.57	4.76
11	1.07	−2.20	−1.08	0.10	1.48
12	0.64	−3.49	−2.36	−0.20	1.71
13	0.61	−3.71	−1.09	0.36	3.09
14	**1.24**	**−2.51**	**−0.40**	**1.18**	**2.76**
15	0.74	−3.94	−1.55	0.60	3.20
16	0.67	−1.84	0.25	1.76	4.61
17	0.79	−6.57	−4.14	−1.17	1.65
18	0.84	−2.86	−0.25	1.68	4.38
19	**1.25**	**−1.07**	**0.40**	**1.49**	**2.92**
20	0.78	−6.10	−2.53	−0.09	3.17

Note. $N = 528$. Items shown in bold were selected for the MACH*. They also had the five highest discrimination parameters (i.e., above $a = 1.20$).

Discrimination and threshold parameters. IRT parameters a and $b_{1–4}$ can be found in Table 3. As judged by the discrimination parameter *a,* the most discriminative items were 3, 6, 7, 14, and 19. These items primarily captured cynical thinking about people and the world. The least discriminative items were 4, 9, 12, 13, 15, and 16.

For threshold b_1 (answering Option 1 vs. 2, 3, 4, and 5), the easiest items were 4, 17, and 20, and the most difficult were 7, 10, and 19. For threshold b_2 (answering Options 1 and 2 vs. 3, 4, and 5), the easiest items were 4, 17, and 20, and the most difficult were 3, 6, and 8. For threshold b_3 (answering Options 1, 2, and 3 vs. 4 and 5), the easiest items were 2, 9, and 10, and the most difficult items were 11, 12, and 20. For threshold b_4 (answering Options 1, 2, 3, and 4 vs. 5), the easiest items were 1, 9, and 10, and the most difficult items were 11, 12, and 17. Overall across all thresholds $b_{1–4}$, the most difficult items were 1, 9, and 10 (followed by 2, 6, and 17), and the easiest items were 4, 17, and 20 (followed by 11, 12, and 15).

Item information curves. IICs of all 20 MACH–IV items are presented in Figure 1. As can be seen, most items had relatively flat information curves, with some exceptions. These exceptions showed stronger measurement precision at average to above average levels of θ, whereas well below average to below average levels of θ showed the least (but still good) measurement precision. Item 7 showed the strongest measurement precision in the above to well above average level of θ. Items 3, 6, 7, 14, and 19 provided the most information (i.e., over 0.4 at the highest point) across different levels of θ. All other items showed markedly flatter information curves. Items 4, 8, 9, 12, 13, 15, 16, 17, 18, and 20 provided only little or negligible information (i.e., beneath 0.2 at the highest point) across different levels of θ. Items 1, 2, 5, 10, and 11 had their maxima slightly above 0.2 at their highest points and thus performed on a medium level.

Looking at Table 1, the content of the most informative items was particularly cynicism/misanthropy. The content of the least

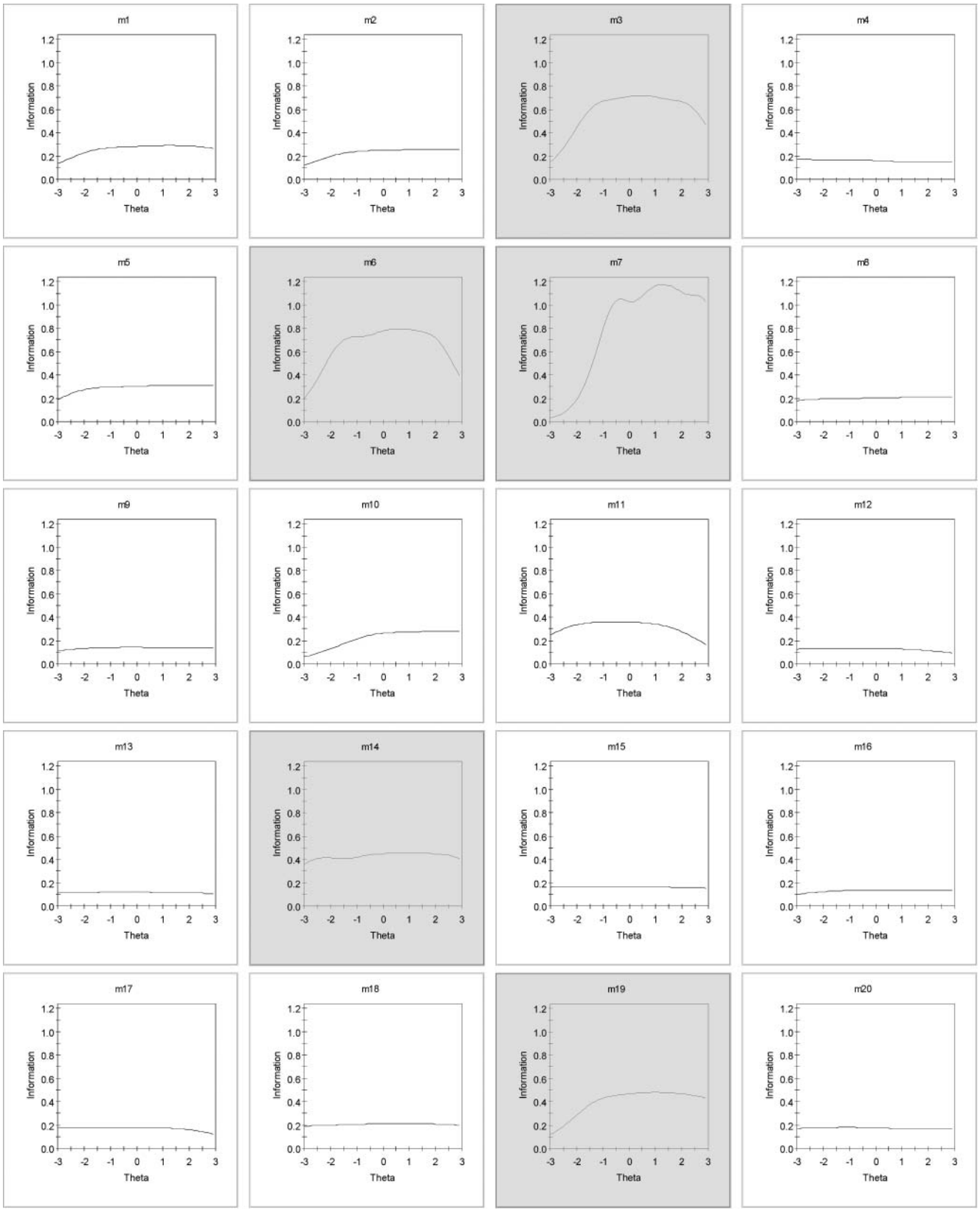

FIGURE 1.—Item information curves of the 20 MACH–IV items. *Note.* m1 to m20 denote the items found in Table 1 (Items 1–20). Gray shaded items provide most information (i.e., peaks > 0.40) to the overall MACH–IV score and were selected for the MACH*.

informative items was not tied together by a specific content they convey, but by how they were phrased. These items were (a) too radical or absolutistic (e.g., It is possible to be good in all respects), (b) (sad) pragmatic facts of life (e.g., It is wise to flatter important people), (c) realistically and statistically speaking wrong or highly unlikely (e.g., Most men are brave), or (d) in the meantime outdated or without such a controversial/negativistic touch that they used to have (e.g., People suffering from incurable diseases should have the choice of being put painlessly to death). Such items were not informative enough for what the MACH–IV measures at its core and could thus be dropped for a trimmed MACH* scale.

Test information function. The TIF of the MACH–IV total score can be found in Figure 2. Measurement precision was distributed from well below average to well above average levels of θ, with a minimum of approx. 4.0 (at $\theta \approx -3$) and a maximum of approximately 7.5 (at $\theta \approx 1$). The MACH–IV provided most precision in the average to above average and least precision in well below average levels of θ. This finding might correspond to the problem of identifying what exactly low levels of Machiavellianism are (e.g., the absence or inverse of Machiavellianism), whereas high levels are much more straightforward to interpret (and usually the focus of studies and interest).

*The Trimmed MACH**

Item selection. IRT findings point toward several points of improvement to maximize information, measurement precision, and item relevance. Items should be (a) neither too easy nor too difficult (i.e., mean over b_1 to b_4 between 0.00 and 1.50), (b) strongly correlated with the total MACH–IV score (i.e., item–total correlation > .45) as well as most informative and precise (i.e., discrimination $a > 1.20$), and (c) not redundant with each other. Items 3, 6, 7, 14, and 19 were chosen for the MACH* because these performed most favorably overall.

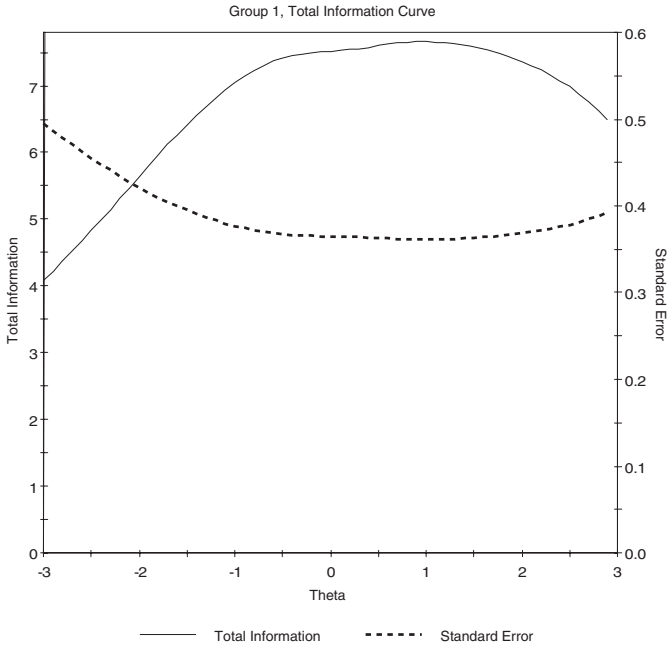

FIGURE 2.—Test information function of the MACH–IV total score.

Descriptive statistics. Mean of the total MACH* score was 2.51 ($SD = 0.81$), with an internal consistency reliability of $\alpha = .77$ (i.e., only slightly below the MACH–IV, but with only one fourth of items) and a mean interitem correlation of .41 (r range = .29–.53). Item–total correlations ranged from .47 to .65.

Dimensionality. To examine dimensionality of the MACH*, an exploratory factor analysis (principal axis factoring, promax-rotation with $\kappa = 4$) was run over all items (Kaiser–Meyer–Olkin measure of sampling adequacy: .82; Bartlett's test of sphericity: approximately $\chi^2(10) = 640.33$, $p < .001$), which yielded one factor accounting for 41.76% of total variance. Factor loadings were .77 (Item 6), .68 (Item 3), .68 (Item 7), .54 (Item 19), and .53 (Item 14). A follow-up confirmatory factor analysis indicated superb fit for a unidimensional model, $\chi^2(5) = 7.09$, $p = .214$; $\chi^2/df = 1.42$; Tucker–Lewis Index (TLI) = .99; comparative fit index (CFI) = 1.00; root mean square error of approximation (RMSEA) = .02. The MACH* can thus be regarded a unidimensional measure.

IRT analyses. Discrimination and threshold parameters of MACH* items can be found in Table 4 and the IICs in Figure 3. As can be seen, Item 14 provided the least information at approximately 0.4 (quite evenly across all levels of θ) and Item 6 the most information at approximately 1.7 (the most around $-1.5 < \theta < 1.5$). Generally, though, all five items showed enough information and discrimination (see Table 4 and Figure 3). The TIF (Figure 4) indicated that the MACH* is most precise in the interval of $-1.5 < \theta < 1.5$, with maximum precision around $0.5 < \theta < 1.5$. The MACH* thus provided most information at average to above average levels of Machiavellianism, quite comparable to the MACH–IV. In relative terms to the MACH–IV, however, it provided less overall information (with a peak at slightly above 6.0 as compared to the MACH–IV with a peak at around 7.5) and less information at well to below average levels of θ. This might be expected by a scale that is four times shorter, though.

Construct validity. The MACH* was strongly correlated with a version of the MACH–IV that did not contain the MACH* items, $r = .63$ (disattenuated: .84), $p < .001$. To further examine and compare convergent and discriminant construct validity, the MACH–IV and MACH* were correlated with scales of Machiavellianism (German Machiavellianism Scale, Dirty Dozen Machiavellianism), narcissism (NPI–17, Dirty Dozen narcissism), and psychopathy (SRP–III global psychopathy, Dirty Dozen psychopathy). Findings are presented in Table 5 under Dark Triad. First, the MACH–IV and MACH* correlated similarly with the German Machiavellianism Scale.

TABLE 4.—Discrimination and threshold parameters of the MACH*.

Items	a	b_1	b_2	b_3	b_4
3	1.83	−1.19	0.01	0.87	2.04
6	2.48	−1.27	−0.05	0.74	1.66
7	1.87	−0.54	0.86	1.68	2.83
14	1.24	−2.52	−0.39	1.19	2.75
19	1.23	−1.10	0.41	1.53	2.93

Note. $N = 528$.

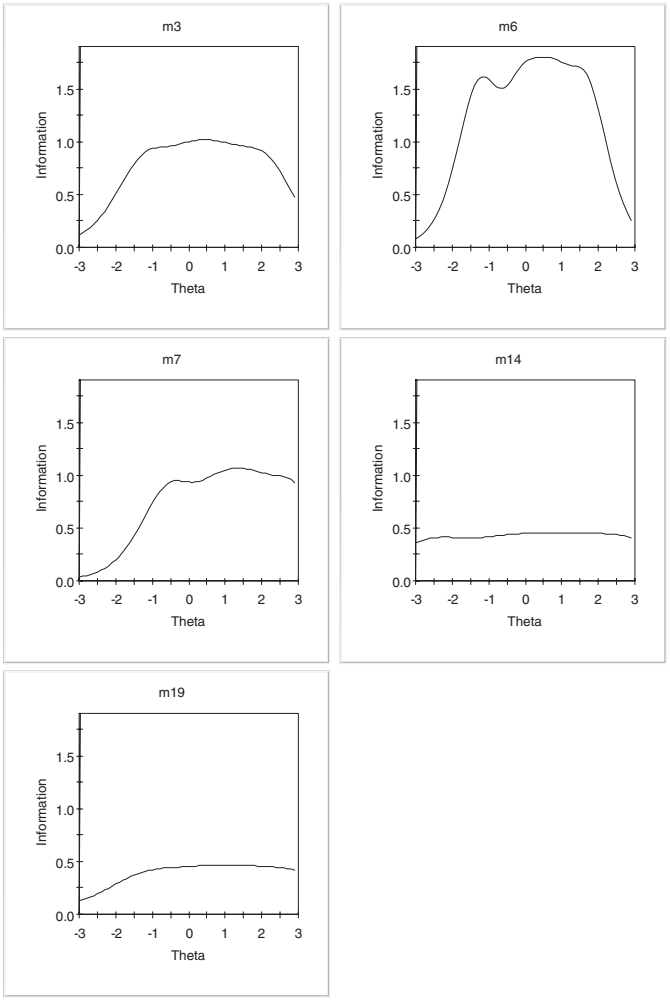

FIGURE 3.—Item information curves of the five MACH* items. *Note.* Items can be found in Table 1.

Second, the MACH* showed lower correlations with NPI–17 narcissism, Dirty Dozen narcissism, and SRP–III psychopathy, indicating better discriminant validity. Third, there was a peculiarity with the Dirty Dozen Machiavellianism and psychopathy scales: Both correlated strongly with the MACH–IV and the MACH*. Indeed, the correlation between Dirty Dozen psychopathy was larger for the MACH–IV and especially the MACH–IV. This pattern could hint at problems with the Dirty Dozen rather than the MACH scales (see similar findings in Jonason & Webster, 2010, Table 4, p. 424; Miller et al., in press). Together, these findings provide evidence that the MACH* shows good convergent and discriminant validity.

Criterion validity. Presence, magnitude, and direction of associations with external criteria should be largely similar between the MACH–IV and MACH*. To examine this, both scales were correlated with the Big Six (BFI–S Big Five + HEXACO Honesty/Humility) and eight manipulation tactics commonly used by Machiavellians (see Measures). As can be seen in Table 5, the correlation pattern of the MACH–IV and MACH* with 14 external criteria was highly similar, as indicated by a vector correlation of $r_v = .99$ ($p < .001$; computed from disattenuated correlations: $r_v = .87$, $p < .001$). However, the MACH* generally showed correlations lower in magnitude than the MACH–IV. Moreover, the correlation with neuroticism turned significant for the MACH* (not for the MACH–IV).

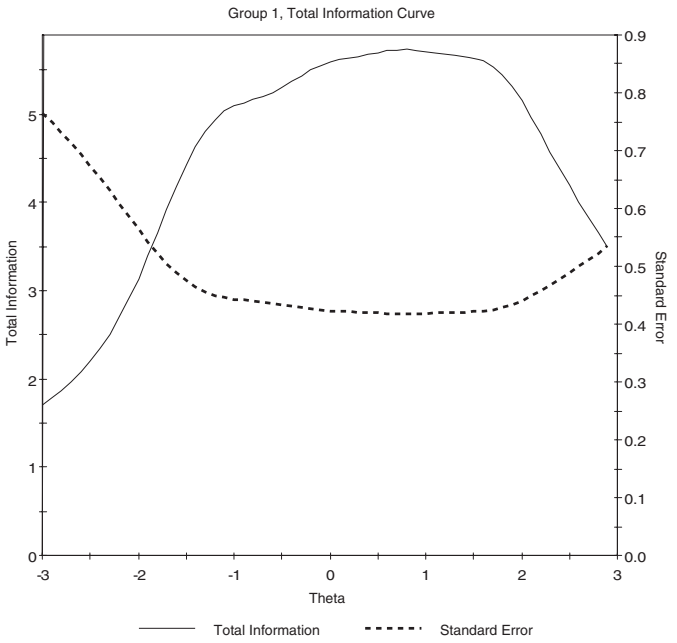

FIGURE 4.—Test information function of the MACH* total score.

TABLE 5.—Construct and criterion validity of the MACH–IV and MACH* compared by patterns of correlations.

External Criteria	Descriptive Statistics			Correlations			
	M	*SD*	α	MACH–IV		MACH*	
Dark Triad							
GMS	2.59	0.75	.90	.68	(.79)	.60	(.72)
DD Machiavellianism	2.05	0.85	.85	.50	(.60)	.33	(.41)
DD Psychopathy	2.02	0.79	.71	.58	(.76)	.51	(.69)
DD Narcissism	2.63	0.89	.82	.21	(.25)	.15	(.18)
NPI–17 Narcissism	2.58	0.66	.88	.30	(.35)	.19	(.23)
SRP–III Psychopathy	2.49	0.43	.81	.49	(.60)	.36	(.46)
Big Six							
Neuroticism	3.17	0.85	.75	.02	(.03)	.09	(.12)
Extraversion	3.33	0.87	.82	−.23	(−.28)	−.21	(−.26)
Openness	3.78	0.79	.68	−.06	(.08)	−.03	(.04)
Agreeableness	3.68	0.65	.54	−.37	(−.56)	−.31	(−.48)
Conscientiousness	3.54	0.71	.66	−.30	(−.41)	−.20	(.28)
Honesty/Humility	3.47	0.55	.79	−.52	(−.65)	−.35	(−.49)
Manipulation tactics							
Emotional manipulation	2.06	1.18	.84	.45	(.54)	.33	(.41)
Ingratiation	2.48	1.29	.84	.32	(.39)	.21	(.26)
Persuasion	3.83	1.49	.76	.31	(.39)	.19	(.25)
Supplication	3.32	1.22	.51	.23	(.36)	.12	(.19)
Intimidation	1.59	1.00	.83	.33	(.40)	.23	(.29)
Self-disclosure	2.06	1.31	.86	.22	(.26)	.15	(.18)
Betrayal	1.95	1.20	.78	.52	(.65)	.40	(.52)
Revenge	2.31	1.53	.90	.49	(.57)	.39	(.47)

Note. $N = 528$. Correlations in brackets are fully disattenuated. GMS = German Machiavellianism Scale; DD = Dirty Dozen; NPI = Narcissistic Personality Inventory; SRP = Self-Report Psychopathy Scale. Correlations above |.09| are significant at least at $p < .05$.

Apart from these small differences, the MACH–IV and MACH* tapped the same correlates and outcomes. The MACH* can thus be seen as a shorter version of the MACH–IV with still comparable criterion validity to its longer version.

Discussion

*The MACH–IV and MACH**

IRT analyses yielded that (a) many MACH–IV items had rather low probabilities of endorsement (in different categories along the response scale), (b) approximately three quarters of items provided only little to negligible information, and (c) the strongest measurement precision was found in average to above average levels of Machiavellianism. The items with most measurement precision or information were those with cynicism and misanthropy content, suggesting that this theme lies at the core of what the MACH–IV is measuring. This provides an answer to researchers questioning what the instrument actually measures (e.g., Panitz, 1989). These items, also chosen for the MACH*, were "Anyone who completely trusts anyone is asking for trouble," "It is safest to assume that all people have a vicious streak and it will come out when they are given a chance," "Never tell anyone the real reason you did something unless it is useful to do so," "Most people are basically good and kind," and "The biggest difference between most criminals and other people is that criminals are stupid enough to get caught." These items are all indicative of a cynical view of people in general (e.g., that they should not the trusted, they are basically evil, they will betray you, they are stupid, etc.). The other items not performing so well were mixed in their specific content and classifications, but they had in common that they were (a) poorly phrased, such as too radical (e.g., Item 13: It is possible to be good in all respects) or reflecting (sad but) common facts (e.g., Item 9: It is wise to flatter important people); (b) too specific or puzzling to respondents (e.g., Item 11: Barnum was very wrong when he said there's a sucker born every minute); or outdated (e.g., Item 12: People suffering from incurable diseases should have the choice of being put painlessly to death). These items might introduce noise to the MACH–IV that can contribute to the factorial invariance often found (see earlier).

The five MACH* items sampling cynical and misanthropic content were chosen plainly on empirical grounds of IRT analyses guided by this question: Which items are the most informative and precise? However, it also makes sense conceptually to postulate cynical and misanthropic beliefs at the core of the construct Machiavellianism, not just the measure. First, the focus on misanthropy—a cynical view on people in general—resembles strongly one of the key topics identified across the range of the diverse literatures roamed by Christie and colleagues (which they dubbed "Views" originally), namely that human kind was basically weak, fallible, and susceptible so that one can and should take advantage of others to one's benefit. Second, endorsing cynical and misanthropic beliefs might covary with other typically Machiavellian tendencies (e.g., callousness and immorality) and lead to Machiavellian phenotypical behaviors such as exploitation and manipulation. Thus, misanthropy could be at the heart of Machiavellianism: Cynical, misanthropic belief systems could be the underlying core that generates, facilitates, intensifies, maintains, and also justifies devious, self-beneficial interactions with other people. Third, Machiavellians have been found to harbor negative views of others and themselves, which can be taken as an indicator of broad negativistic person views (Rauthmann, 2012). Together, the narrower content of the new MACH* scale is empirically supported by IRT analyses, makes theoretical sense, and posits a plausible core factor of Machiavellianism that gives rise to many other Machiavellian phenomena.

In psychometric terms, the MACH* is quite formidable for a five-item scale and indeed comparable to the MACH–IV in terms of construct and criterion validity. The five MACH* items show good internal consistency, information, and precision. Moreover, the MACH* measures more efficiently and economically what the MACH–IV captures; it is not fuzzy such as the broader MACH–IV for which factorial validity is hard to establish (e.g., Panitz, 1989). The MACH* also showed reduced correlations with the other Dark Triad members narcissism and psychopathy and thus better discriminant abilities.

Limitations and Further Points of Improvement

One limitation of this work is the sole empirical, inductive approach to the MACH* items. However, in the absence of a clear theory on Machiavellianism (Fehr et al., 1992; Wilson et al., 1996) and the need for a streamlined MACH–IV scale, the application of IRT seems a relevant and important undertaking. The MACH* can be regarded as the most precise and informative of what the MACH–IV has to offer. Another limitation might be the narrower scope of the MACH* as its content is closely tailored to cynical and misanthropic beliefs. Construct and criterion validity, however, inform us that the MACH* performs similarly to its longer counterpart. Moreover, cynical and misanthropic views on people as the core of Machiavellianism makes conceptually more sense than other content areas, such as dishonesty, deceit, agentic striving, or manipulation and exploitation. The latter are better seen as behavioral consequences that ensue from Machiavellian cynicism and misanthropy, not facets of Machiavellianism per se. If they were facets, then the prediction of behavioral criteria (e.g., cheating, lying, exploiting, etc.) would be a matter of circularity because the Machiavellianism scales would assess exactly what they are trying to predict.

How could the MACH* be even further improved? First, items can be framed from an "I"-perspective to generate more personal relevance for respondents so that they actually describe themselves. The current items use statements to measure the level of endorsement that should be indicative of underlying levels of Machiavellianism: Essentially attitudes and beliefs are sampled (Kraut & Price, 1976). A more personal and direct approach would be phrasing items such as "I think low of people," "I am suspicious of others," or "I use people as means to an end." Such items might increase predictive validity, as self-descriptions can be more indicative of actual behavior and social outcomes than abstract, impersonal attitudes that need not manifest in one's everyday behavior. Second, the generic Likert-type scale with (dis)agreement anchors could be replaced by a frequency-based Likert-type scale (from *almost never* to *almost always*) when using "I"-perspective items to enhance psychometric properties, such as predictive validity (see Rauthmann & Denissen, 2011). Third, different forms of misanthropy might be distinguished, such as being suspicious of others (i.e., suspicion), thinking that other people are worthless (i.e., active devaluation or derogation), and seeing people as means toward some other self-serving end (i.e., instrumentality). Fourth, invariance across different "groups" (e.g., nations, culture, ethnicity, sex, age classes, or experimental designs) could be established for

the MACH*. IRT can help examine whether groups differ consistently in their responses to items, but with same θ levels, by differential item functioning (Embretson & Reise, 2000).

Implications for Theory and Conceptualization: The Problem of Dimensionality

Machiavellianism is in theory probably best seen as a multidimensional construct, but is usually treated in measurement as a unidimensional one (Rauthmann, 2011; Rauthmann & Will, 2011). Interestingly though, Panitz (1989) stated that "the issue is not whether Machiavellianism is a single or multi-dimensional construct, but whether the construct is measurable using the Mach IV scale" (p. 963). Regardless, given (a) the theoretical and empirical controversy around the factorial structure of Machiavellianism and the MACH–IV, (b) the value of homogenous constructs (Smith, McCarthy, & Zapolski, 2009), (c) the widespread customs of computing only one total score of the MACH–IV, and (d) the general need for short and efficient but effective scales, the unidimensional MACH* can be very helpful for researchers and practitioners. There is no doubt that we could distinguish different facets of Machiavellianism (for examples, see Rauthmann, 2011; Rauthmann & Will, 2011), possibly with different correlates and consequences (see Hunter et al., 1982). However, theory should suggest which and how many facets there are, how they are interrelated, which correlates and trajectories they entail, and how an according measure should look like. The focus around cynical, misanthropic beliefs that spur associated structures and processes such as insensitivity and callousness, which lead to overt manipulation behavior, might be a fruitful avenue for future Machiavellianism theorization.

Implications for Scale Development: Short and New Scales

The MACH* could be used as an alternative, streamlined version of the MACH–IV. If researchers want to tap the constructs measured by longer, firmly established scales with less noise items but as closely as possible, short scales can be devised using IRT analyses. Similar to the approach taken here, an NPI* from the NPI (see Ackerman, Donnellan, & Robins, 2012 for an IRT analysis of the NPI) and an SRP* from the SRP–III could be constructed, also with about five items each. Together, the NPI*, MACH*, and SRP* could tap the full Dark Triad as the "Filthy Fifteen" with only 15 instead of about 90 items. Dark Triad research is blossoming (e.g., Jones & Paulhus, 2010; O'Boyle, Forsyth, Banks, & McDaniel, 2012; Rauthmann, 2012), but most research designs cannot handle full, multidimensional measures of each member. Thus, short scales, diligently constructed and maximally convergent to their longer versions, will be needed to study the Dark Triad in concert (for existing scales, see Jonason & Webster, 2010).

This work did not plead for a completely different measurement approach to Machiavellianism, but only proposed an "updated" MACH* version. However, the time seems ripe 40 years after Christie and Geis's (1970) seminal work on Machiavellianism to propose new measures based on sound theory and more strongly on what existing literature has to offer (see, e.g., Fehr et al., 1992; Jones & Paulhus, 2009; Rauthmann, 2011; Rauthmann & Will, 2011; Wilson et al., 1996). Some alternative measures already exist (e.g., Machiavellian Behavior Scale Mach–B: Aziz, May, & Crotts, 2002; Aziz & Meeks, 1990; Dirty Dozen Machiavellianism: Jonason & Webster, 2010; MACH–VI: Jones & Paulhus, 2009; Machiavellian Personality Scale MPS: Dahling, Whitaker, & Levy, 2009; Machiavellianism scale for the workplace: Kessler et al., 2010), but they remain virtually unused. The MACH–IV has been and still is the golden standard of measuring individual differences in Machiavellianism, but this is potentially dangerous because "personality measures assume a life of their own and implicitly become the construct; that is, a construct is operationally defined by the method that is used to measure it" and once the measure "becomes popular it is equated with the construct" (Schimmack, 2010, p. 243). This might make it difficult for new measures to amend, extend, or even replace the MACH–IV from its pole position. In any case, future Machiavellianism scales would ideally (a) be grounded in and guided by theory, (b) measure Machiavellianism multidimensionally, (c) use an economic shorter style scale, (d) be examined with both classical test theory and IRT analyses in comprehensive validation programs, and (e) capture correlates and outcomes similar to the MACH–IV but with incremental precision.

Conclusion

This work investigated the MACH–IV with IRT analyses and provided a psychometrically sound, streamlined version of it, the five-item MACH*, which can be considered the "best of" what the MACH–IV has to offer. Analyses point toward cynical and misanthropic beliefs as core content of what the MACH–IV measures and that these can be best assessed at average to above average levels of Machiavellianism. Future research on Machiavellianism can build on these insights and tackle conceptualization and measurement of Machiavellianism.

References

Ackerman, R. A., Donnellan, M. B., & Robins, R. W. (2012). An item response theory analysis of the Narcissistic Personality Inventory. *Journal of Personality Assessment*, *94*, 141–155.

Ahmed, S. M. S., & Stewart, R. A. C. (1981). Factor analysis of the Machiavellian scale. *Social Behavior and Personality*, *9*, 113–115.

Ashton, M. C., Lee, K., & Son, C. (2000). Honesty as the sixth factor of personality: Correlations with Machiavellianism, primary psychopathy, and social adroitness. *European Journal of Personality*, *14*, 359–368.

Aziz, A., May, K., & Crotts, J. C. (2002). Relations of Machiavellian behaviors with sales performance of stockbrokers. *Psychological Reports*, *90*, 451–460.

Aziz, A., & Meeks, J. (1990). *A new scale for measuring Machiavellianism.* Unpublished paper. School of Business and Economics, College of Charleston, Charleston, SC.

Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. New York, NY: Academic.

Christie, R., & Lehman, S. (1970). The structure of Machiavellian orientations. In R. Christie & F. Geis (Eds.), *Studies in Machiavellianism* (pp. 359–387). New York, NY: Academic.

Corral, S., & Calvete, E. (2000). Machiavellianism: Dimensionality of the Mach IV and its relation to self-monitoring in a Spanish sample. *Spanish Journal of Psychology*, *3*, 3–13.

Dahling, J. J., Whitaker, B. G., & Levy, P. E. (2009). The development and validation of a new Machiavellianism scale. *Journal of Management*, *35*, 219–257.

de Vries, R. E., & van Kampen, D. (2010). The HEXACO and 5DPT models of personality: A comparison and their relationships with psychopathy, egoism, pretentiousness, immorality, and Machiavellianism. *Journal of Personality Disorders*, *24*, 244–257.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

Fehr, B., Samsom, D., & Paulhus, D. L. (1992). The construct of Machiavellianism: Twenty years later. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in personality assessment* (Vol. 9, pp. 77–116). Hillsdale, NJ: Erlbaum.

Henning, H., & Six, B. (2008). Machiavellismus [Machiavellianism]. In A. Glöckner-Rist (Ed.), *Zusammenstellung sozialwissenschaftlicher Items und Skalen* [Compilation of items and scales for the social sciences]. ZIS Version 12.00. Bonn, Germany: GESIS.

Hunter, J. E., Gerbing, D. W., & Buster, F. J. (1982). Machiavellian beliefs and personality: Construct invalidity of the Machiavellian dimension. *Journal of Personality and Social Psychology*, *43*, 1293–1305.

IRTPRO. (2012). *User's guide to IRTPRO 2.1 for Windows*. Skokie, IL: SSI Systems Inc.

Jakobwitz, S., & Egan, V. (2006). The Dark Triad and normal personality traits. *Personality and Individual Differences*, *40*, 331–339.

Jonason, P. K., & Webster, G. D. (2010). The Dirty Dozen: A concise measure of the Dark Triad. *Psychological Assessment*, *22*, 420–432.

Jonason, P. K., Webster, G. D., Schmitt, D. P., Li, N. P., & Crysel, L. (2012). The antihero in popular culture: A life history theory of the Dark Triad. *Review of General Psychology*, *16*, 192–199.

Jones, D. N., & Paulhus, D. L. (2009). Machiavellianism. In M. R. Leary & R. H. Doyle (Eds.), *Handbook of individual differences in social behavior* (pp. 93–108). New York, NY: Guilford.

Jones, D. N., & Paulhus, D. L. (2010). Differentiating the Dark Triad within the interpersonal circumplex. In L. M. Horowitz & S. N. Strack (Eds.), *Handbook of interpersonal theory and research* (pp. 249–267). New York, NY: Guilford.

Kessler, S., Bandelli, A., Spector, P., Borman, W., Nelson, C., & Penny, L. (2010). Re-examining Machiavelli: A three-dimensional model of Machiavellianism in the workplace. *Journal of Applied Social Psychology*, *40*, 1868–1896.

Kline, P., & Cooper, C. (1983). A factor-analytic study of measures of Machiavellianism. *Personality and Individual Differences*, *4*, 569–571.

Kline, P., & Cooper, C. (1984). An investigation of the items in the Machiavellian scales. *Journal of Social Psychology*, *124*, 251–252.

Kraut, R. E., & Price, J. D. (1976). Machiavellianism in patients and their children. *Journal of Personality and Social Psychology*, *33*, 782–786.

Küfner, A. C. P., Dufner, M., & Back, M. (2012). *Das Dreckige Dutzend und die Niederträchtigen Neun—Zwei Kurzskalen zur Erfassung von Narzissmus, Machiavellismus und Psychopathie* [The Dirty Dozen and the Naughty Nine: Two short scales for the assessment of narcissism, Machiavellianism, and psychopathy]. Manuscript submitted for publication.

Kuo, H. K., & Marsella, A. J. (1977). The meaning and measurement of Machiavellianism in Chinese and American college students. *Journal of Social Psychology*, *101*, 165–173.

Lee, K., & Ashton, M. C. (2004). The HEXACO Personality Inventory: A new measure of the major dimensions of personality. *Multivariate Behavioral Research*, *39*, 329–358.

Lee, K., & Ashton, M. C. (2005). Psychopathy, Machiavellianism, and narcissism in the Five-factor model and the HEXACO model of personality structure. *Personality and Individual Differences*, *38*, 1571–1582.

Miller, J. D., Few, L. R., Seibert, L. A., Watts, A., Zeichner, A., & Lynam, D. R. (in press). An examination of the Dirty Dozen measure of psychopathy: A cautionary tale about the costs of brief measures. *Psychological Assessment*.

Morizot, J., Ainsworth, A. T., & Reise, S. T. (2007). Toward modern psychometrics: Application of item response theory models in personality research. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 407–423). New York, NY: Guilford.

Machiavelli, N. (1981). *The prince*. New York, NY: Bantam Classics. (Original work published in 1513)

O'Boyle, E. H., Forsyth, D. R., Banks, G. C., & McDaniel, M. A. (2012). A meta-analysis of the Dark Triad and work behavior: A social exchange perspective. *Journal of Applied Psychology*, *97*, 557–579.

O'Hair, D., & Cody, M. J. (1987). Machiavellian beliefs and social influence. *Western Journal of Speech Communication*, *51*, 279–303.

Panitz, E. (1989). Psychometric investigation of the Mach IV scale measuring Machiavellianism. *Psychological Reports*, *64*, 963–968.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality*, *36*, 556–563.

Ramanaiah, N. V., Byravan, A., & Detwiler, F. R. J. (1994). Revised NEO Personality Inventory profiles of Machiavellianism and non-Machiavellian people. *Psychological Reports*, *75*, 937–938.

Rauthmann, J. F. (2011). Towards multifaceted Machiavellianism: Content, factorial, and construct validity of a German Machiavellianism Scale. *Personality and Individual Differences*, *52*, 345–351.

Rauthmann, J. F. (2012). The Dark Triad and interpersonal perception: Similarities and differences in the social consequences of narcissism, Machiavellianism, and psychopathy. *Social Psychological and Personality Science*, *3*, 487–496.

Rauthmann, J. F., & Denissen, J. J. A. (2011). I often do it vs. I like doing it: Comparing a frequency- and valency-approach to extraversion. *Personality and Individual Differences*, *8*, 1283–1288.

Rauthmann, J. F., & Will, T. (2011). Proposing a multidimensional Machiavellianism conceptualization. *Social Behavior and Personality*, *39*, 391–404.

Ray, J. J. (1983). Defective validity of the Machiavellianism scale. *Journal of Social Psychology*, *119*, 291–292.

Reeve, B. B., & Fayers, P. (2005). *Applying item response theory modeling for evaluating questionnaire item and scale properties.* In P. Fayers & R. D. Hays (Eds.), Assessing quality of life in clinical trials: Methods of practice (2nd ed., pp. 55–73). New York, NY: Oxford University Press.

Samejima, F. (1969). Estimation of latent trait ability using a response pattern of graded scores. *Psychometric Monograph*, *18*.

Samejima, F. (1996). Graded response model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 85–100). New York, NY: Springer.

Schimmack, U. (2010). What multi-method data tell us about construct validity. *European Journal of Personality*, *24*, 241–257.

Schupp, J., & Gerlitz, J.-Y. (2008). BFI-S: Big Five Inventory—SOEP. In A. Glöckner-Rist (Ed.), *Zusammenstellung sozialwissenschaftlicher Items und Skalen* [Compilation of items and scales for the social sciences]. Version 12.00. Bonn, Germany: GESIS.

Shajek, A. (2007). *Entwicklung zweier Kurzskalen zur Messung von Machiavellismus und Egoismus. Dokumentation eines Instrumententests auf Basis des SOEP-Pretests 2006* [Development of two short scales for the measurement of Machiavellianism and egoism. Documentation of an instrument test based on the SOEP-pretest 2006.]. Berlin, Germany: DIW Berlin.

Smith, G. T., McCarthy, D., & Zapolski, T. (2009). On the value of homogeneous constructs for construct validation, theory testing, and the description of psychopathology. *Psychological Assessment*, *21*, 272–284.

van der Linden, W. J., & Hambleton, R. K. (1997). *Handbook of modern item response theory*. New York, NY: Springer.

Vleeming, R. G. (1984). The nomothetical network of a Machiavellianism scale. *Psychological Reports*, *54*, 617–618.

von Collani, G. (2008). *Modifizierte deutsche Versionen des*Narcissistic Personality Inventory *(NPI-d)* [Modified German versions of the Narcissistic Personality Inventory (NPI-d)]. In A. Glöckner-Rist (Ed.), *Zusammenstellung sozialwissenschaftlicher Items und Skalen* [Compilation of items and scales for the social sciences]. Version 12.00. Bonn, Germany: GESIS.

Williams, K. M., Nathanson, C., & Paulhus, D. L. (2003, August). *Structure and validity of the Self-Report Psychopathy Scale–III in normal populations*. Presentation at the 11th annual convention of the American Psychological Association, Toronto, Canada.

Williams, M. L., Hazelton, V., & Renshaw, S. (1975). The measurement of Machiavellianism: A factor analytic and correlational study of Mach IV and Mach V. *Speech Monographs*, *42*, 151–159.

Wilson, D. S., Near, D., & Miller, R. R. (1996). Machiavellianism: A synthesis of the evolutionary and psychological literatures. *Psychological Bulletin*, *119*, 285–299.

Wilson, D. S., Near, D., & Miller, R. R. (1998). Individual differences in Machiavellians as a mix of cooperative and exploitative strategies. *Evolution and Human Behavior*, *19*, 203–212.

Zook, A., & Sipps, G. J. (1986). Reliability and sex-differences with a gender-free Mach IV. *Journal of Social Psychology*, *126*, 131–132.