Mental Health Literacy of Depression:

A Preregistered Study Reconsidering Gendered Differences Using Filmed Disclosures

Viren Swami[1-2], Clemens Georg Grüneis[3], Martin Voracek[3], & Ulrich S. Tran[3]

[1]School of Psychology and Sports Sciences, Anglia Ruskin University, Cambridge, United Kingdom

[2]Centre for Psychological Medicine, Perdana University, Kuala Lumpur, Malaysia

[3]Department of Cognition, Emotion, and Methods in Psychology, School of Psychology, University of Vienna, Vienna, Austria

Address correspondence to: Prof. Viren Swami, School of Psychology and Sport Science, Anglia Ruskin University, East Road, Cambridge, Cambridgeshire CB1 1PT, United Kingdom. Email: viren.swami@aru.ac.uk.

**Abstract**

The study of mental health literacy is well-developed, but the basic methodology used in this research (i.e., the use of text-based vignettes) has not changed substantively in over two decades of work. Here, we developed novel filmed disclosures to re-assess mental health literacy of depression in a preregistered study utilising a randomised, fully within-subjects design. A total of 405 adults from Austria (57% women, age $M = 32.5$ years) viewed short (~3 min) filmed disclosures by a female or male target and were asked to report if they thought anything was wrong with the targets and, if so, to describe what they thought was wrong. Participants also rated the targets on a range of attitudinal dimensions and completed measures of conformity to masculine norms and expressivity. The majority of participants (93.8%) correctly identified that something was wrong with the targets and, of those that did, the majority (69.2%) correctly described cases of depression. Neither target nor participant gender significantly influenced symptom recognition. Gendered effects were also largely null in terms of perceived distress, treatment difficulty, sympathy, and likelihood of recommending help for the targets, and both conformity to masculine norms and expressivity had minimal impact on attitudinal dimensions. These results highlight the potential utility of filmed disclosures in the study of mental health literacy and suggest that gendered effects reported in previous studies may be an artefact of the use of text-based vignettes.

    **Keywords:** Mental health literacy; Filmed disclosures; Gender differences; Study preregistration; Within-subjects design

**Public Significance Statement**

Our results suggest that Austrian respondents have generally good mental health literacy of depression, which may have been facilitated through the use of filmed disclosures. Using filmed disclosures may help the public to better recognise symptoms of depression.

**Introduction**

Across the world, mental health disorders are highly prevalent: one meta-analysis of 174 surveys across 63 nations reported that 17.6% (95% *CI* = 16.3, 18.9) of respondents met criteria for common mental health disorders in the year preceding assessment, and 29.2% (95% *CI* = 25.9, 32.6) experienced a mental health disorder at some time during their lifetimes (Steel et al., 2014). Early help-seeking is vital for timely intervention and improved long-term outcomes (e.g., Boonstra et al., 2012; Dell'Osso et al., 2013), but most people with symptoms of mental health disorders do not seek or receive optimal care (Alonso et al., 2007). This has important implications, not only for individuals themselves (e.g., Vigo et al., 2016), but also their families and for societies more generally. The latter, for instance, includes not just direct costs, but also indirect costs to national economies (e.g., due to sick leave and decreased productivity; Sobocki et al., 2007).

Reasons for low rates of help-seeking for mental health disorders are multi-faceted and include structural barriers (e.g., lack of available services; Gulliver et al., 2010), cultural and religious beliefs (Hurley et al., 2020), and person-level factors, such as mistrust of psychiatric services (e.g., Soorkia et al., 2010) and public, perceived, and self-stigmatising attitudes toward mental health disorders (Clement et al., 2015). Another important person-level factor that has been identified in the literature is poor *mental health literacy*, which refers to a person's "knowledge and beliefs about mental disorders that aid the recognition, management, or prevention of these disorders" (Jorm et al., 1997, p. 143). Mental health literacy includes the ability to recognise and differentiate mental health disorders and symptoms, knowledge of how and where to seek information relating to mental health disorders, and attitudes that influence the ability to identify symptoms and seek appropriate help (Jorm, 2012, 2015, 2019; Jorm et al., 1997; Kelly et al., 2007; Spiker & Hammer, 2019),

as well as competencies that improve mental health and self-management capabilities (Kutcher et al., 2015, 2016).

A large body of evidence suggests that low levels of mental health literacy are reliably associated with poorer help-seeking attitudes and intentions, as well as lower service use for mental health disorders (for a review, see Furnham & Swami, 2018). Conversely, improving mental health literacy has the potential to empower communities to act for better mental health and thus reduce the social and economic burden of mental ill-health (Jorm, 2012, 2020; Kelly et al., 2007). Unfortunately, converging lines of evidence suggest that mental health literacy is generally poor (for a review, see Furnham & Swami, 2018), with lay respondents having difficulty recognising mental health disorders when presented in the form of text-based vignettes (e.g., Chen et al., 2017; Jorm et al., 2005; Swami, 2012), experiencing difficulty distinguishing real disorders from foils (Swami, Papanicolaou, et al., 2011; Swami, Persaud, et al., 2011), and often misunderstanding psychiatric labels (e.g., Furnham et al., 2009). When presented with text-based vignettes of depression, for example, only about a half of respondents in the United Kingdom (Swami, 2012) and about a third in China (Huang et al., 2019) correctly identified that a mental health disorder was depicted and that the disorder was depression.

However, the claim that lay respondents generally have low mental health literacy obscures marked gender differences (Lee et al., 2020). Thus, studies have consistently reported that women have significantly better mental health literacy than men (e.g., Cotton et al., 2016; Gorczynski et al., 2017; Hadjimina & Furnham, 2017; Swami, 2012). Among North American college students, for example, Kim and colleagues (2015) reported that women were more than twice as likely to correctly recognise cases of depression, as compared to men. Women have also generally been found to have significantly more positive attitudes toward mental health care, to be more sympathetic towards individuals with symptoms of mental health disorders, and to hold more positive views of help-seeking for mental ill-health

(e.g., Holzinger et al., 2012). However, target gender may be an additional factor to consider: in a representative sample of adults from the United Kingdom, Swami (2012) reported that women were more likely to correctly recognise cases of depression, but participants were also less likely to correctly identify the cases when the target was male.

These effects have been ascribed to the impact that gender role ideologies have on health attitudes and behaviours (Swami, 2012). More specifically, Courtenay's (2000) seminal idea that health practices are mediated and expressed through men's masculine performances is typically rallied to explain gendered differences in mental health literacy. In this view, the social construction and idealisation of hegemonic masculinity informs men's understanding of mental ill-health and inhibits both the development of empathetic responses to those with mental health disorders (Gilchrist & Sullivan, 2006) and positive help-seeking attitudes (Seidler et al., 2016). For instance, some research has suggested that, to the extent that depression is perceived as feminine and stigmatised (Gair & Camilleri, 2003; Galasiński, 2008), men (re)construct symptoms of depression as consistent with masculine ideals; that is, they have difficulty distinguishing psychological distress from expressions of traditional masculine attributes, such as risk-taking (Johnson et al., 2012; Oliffe, Robertson et al., 2010; Oliffe, Kelly et al., 2010). Moreover, to the extent that symptoms of mental health disorders are perceived as ostensibly 'inconsistent' with normative expectations and expressions of masculinity (Krumm et al., 2017), both women and men may be less likely to accurately recognise symptoms of depression in male compared with female targets (Gibbons et al., 2015; Swami, 2012; Swami et al., 2020).

While undoubtedly important, to date these theoretical positions have not been fully interrogated in relation to mental health literacy. One concern here is that, surprisingly so, the basic methodology for assessing symptom recognition – namely, the use of text-based vignettes – has not changed in any meaningful way since the seminal work of Jorm and colleagues (1997). Although text-based vignettes do have value (for a review, see Wei et al.,

2015), they have also been critiqued for having poor ecologic validity; that is, text-based vignettes are unlikely to translate adequately to real-life situations and the manner in which disclosures about mental health symptoms are communicated (Furnham & Swami, 2018). For example, text-based vignettes do not capture gesture, facial expressions, and tone of voice, all of which may have an influence of perceptions of others, particularly when issues of mental health are disclosed. In particular, it is possible that the use of text-based vignettes lacks sufficient complexity, as they present almost textbook-like information, which may affect ways in which responses to those cases are formed.

An alternative method for assessing mental health literacy is the use of filmed disclosures, in which actors vocalise case symptoms. To our knowledge, however, only one previous study has used this methodology: Marshall and Dunstan (2013) suggested that, in sample of Australian rural adolescents, filmed disclosures resulted in lower symptom recognition of depression compared to the use of text-based vignettes. However, the authors of that study did not fully report on the production of their filmed stimuli and did not control for severity of symptoms across gender, so it is difficult to generalise these findings. Of note, Furnham and Swami (2018) have suggested that the use of filmed disclosures would allow for greater engagement with content, which in turn may reduce gendered differences. In particular, when viewing films rather than reading text, viewers are more likely to construct mental models of an event based on information gleaned from the stimulus, as well as from their own personal experiences (Busselle & Bilandzic, 2008; Murphy et al., 2013), rather than rely on gendered scripts of behaviour. As a first contribution to the study of mental health literacy, then, the present study used filmed disclosures to assess mental health literacy; that is, rather than relying on text-based vignettes, we instead presented participants with filmed disclosures.

Beyond issues of methodology, extant research has only partially considered the impact that gender role ideologies have on mental health literacy. Thus, conformity to

masculine societal norms has been found to be negatively associated with help-seeking attitudes and behaviours for mental health disorders (Clark et al., 2020; Hammer et al., 2013; McDermott et al., 2018; Seidler et al., 2016; Vogel et al., 2011; Wong et al., 2017), as well as more negative attitudes toward psychiatric services (Berger et al., 2013). Conversely, an intervention that helped to counter masculine norms was found to improve help-seeking attitudes (Calear et al., 2017). Beyond the focus on help-seeking behaviours, however, studies have not fully examined the impact of masculine norms on other facets of mental health literacy, such as broader attitudes toward mental health disorders (e.g., perceived distress, treatment efficacy). Instead, the impact of masculine norms on these broader aspects of mental health literacy has generally been implicitly assumed, rather than empirically tested (Swami et al., 2020).

Beyond conformity to masculine norms, we are not aware of related research that has examined the impact of feminine gender role endorsement on mental health literacy. This is an important lacuna of research evidence, because the sociocultural feminisation of women constructs them as fragile and requiring of care (Oute et al., 2018). Moreover, the expression of distress and distress-related behaviours, such as crying, are often perceived as feminine (Gair & Camilleri, 2003; Gilchrist & Sullivan, 2006). One possibility, therefore, is that individuals who more strongly endorse feminine gender norms might be expected to develop and demonstrate more positive mental health literacy (cf. Furnham & Sjokvist, 2017). Of particular interest in this regard are attitudes toward individuals with mental health disorders, including expressed sympathy, but also attitudes toward help-seeking. Of course, while it should not be assumed that the adoption of feminine gender roles will necessarily operate in ways that are the polar opposite of conformity to masculine norms, we suggest this question represents an under-researched topic that could beneficially be investigated more fully.

**The Present Study**

The study was preregistered at https://aspredicted.org/md3xc.pdf and all study materials (including the utilised films) and data are available at https://osf.io/asujb/. The primary aim of the present study-was to conduct a re-assessment of gender differences in mental health literacy using filmed disclosures instead of text-based vignettes. Specifically, we examined the impact of participant and target gender on the recognition of cases of depression, which we selected given the high prevalence of depression worldwide and its role as one of the primary causes of disability (GBD 2017 Disease and Injury Incidence and Prevalence Collaborators, 2018). In addition, depression is perhaps the most widely studied mental health disorder in the mental health literacy canon (Swami & Furnham, 2018), which allowed us to draw targeted, meaningful comparisons with the extant research literature. Based on prior related findings (Gibbons et al., 2015; Swami, 2012), we expected that women would have better recognition levels than men (Hypothesis 1). Even though not preregistered, we also expected that recognition levels would be lower when the target was male. We further examined the impact of participant and target gender on attitudes toward cases of depression, which we operationalised in terms of perceived distress, treatment efficacy, sympathy, probability of recommending help-seeking behaviour, and the likelihood of recommending different forms of help. Again, we expected that women would hold more positive attitudes than men (Hypothesis 2) and that participants would have more positive attitudes when the target was a woman (not preregistered). A secondary aim of the present study was to examine the extent to which both conformity to masculine norms and feminine gender roles would moderate the expected gendered differences (Hypothesis 3). Even though not preregistered, we expected that conformity to masculine norms would be associated with more negative attitudes towards targets with depression, whereas femininity would be associated with more positive attitudes.

## Method

### Design and Study Materials

The study utilised a fully within-subjects design, with all participants receiving the same survey materials and viewing filmed disclosures by a female and male target in a randomised order.

### Participants

The sample was a non-probabilistic convenience sample recruited from Austria. A total of 501 individuals completed the survey, but we excluded participants who were under 18 years of age or who did not provide their age ($n = 12$), who failed an attention check item embedded halfway through the survey (see Materials; $n = 47$), and who had formal (university) education in psychology or related disciplines (e.g., psychotherapy, medicine, medical and health care; $n = 37$). The final sample, therefore, consisted of 405 participants, of whom the majority self-identified as Austrian citizens (80.5%; German = 10.9%; other nationalities = 8.6%). Participants ranged in age from 18 to 78 years ($M = 32.48$, $SD = 14.59$), and a slightly majority were women (57.0%). In terms of educational qualifications, 1.5% had completed minimum compulsory schooling, 6.7% had completed an apprenticeship, 48.4% had completed high school, and 43.5% had a tertiary qualification.

### Filmed Disclosures

To develop the stimuli used in the present study, two actors – a woman and a man – were instructed to role-play an individual with symptoms of depression, who would be describing their symptoms to a third party for the first time. Both actors were middle-aged Austrians recruited through a dedicated initiative (the *SimulationspatientInnen-Programm* or SPP), as implemented at the Medical University of Vienna, in the course of which professional actors are taught by medical experts to simulate a range of a physical and psychological symptoms for curriculum-based training purposes of medical undergraduates.

The script for both actors was identical, although the order of presentation of symptoms was re-arranged, and was based on the text-based vignettes used by Swami (2012). These vignettes describe an episode of moderate depression based on the *International Classification of Diseases-10* (*ICD-10*; World Health Organization, 1992) and of major depression according to the *Diagnostic and Statistical Manual of the American Psychiatric Association* (*DSM-V*; American Psychiatric Association, 2013). The vignettes were first translated into German using the parallel-blind technique (Behling & Law, 2000), before being converted into a first-person monologue. The professional actors were instructed to closely follow the script's wording, but with the freedom to add conjunctions, pauses, appropriate gestures, and facial expressions to increase credibility. Film and audio quality were ensured using professional lighting and recording equipment. Lighting, background, and camera placement were identical for both actors, who wore long-sleeved black clothing, no make-up, and no jewellery. After final mastering and editing, both films were about three minutes long.

The films were then pilot-tested by presenting them to 10 psychology Masters students. After watching each film, pilot participants were asked to rate the video quality, audio quality, and believability of the actors (1 = *very bad*, 7 = *very good*), as well as severity of symptoms (1 = *very low*, 7 = *very high*). These participants were also asked to the type of depressive disorder displayed (mild, moderate, severe without psychotic symptoms, severe with psychotic symptoms, and other), accompanied by an excerpt from the *ICD-10* diagnostic criteria. The quality of both films was good (video: $M = 5.91$, $SD = 1.22$; audio: $M = 5.27$, $SD = 1.56$). The male actor was somewhat more believable (male: $M = 5.36$, $SD = 1.75$) than the female actor (female: $M = 4.91$, $SD = 2.12$), although the effect size of the difference was small ($d = 0.23$). The attempt to mimic a moderate depressive episode was successful, with all participants identifying the symptoms as a depressive episode and 6 of them as "moderate"

for both cases. The estimated severity was also similar for both actors (male: $M = 4.73$, $SD =$ 1.35; female $M = 4.82$, $SD = 0.98$).

**Materials**

**Ratings of films**. In the main survey, the two films were shown in a randomised order and follow-up questions taken from Swami (2012) – and translated into German using the parallel-blind technique (Behling & Law, 2000) – were presented after each film. First, participants were asked via a *Yes*/*No* question whether believed "anything is wrong" with the targets. If participants selected a *Yes* response, they were presented with an open-ended question asking them what they thought was wrong. Responses to this open-ended question were coded into five categories (*No information*, *Depression*, *Other or non-specified disorder*, *Non-clinical distress*, and *Other*) by two independent judges (Cohen's κ = .96). For analytic purposes, these data were recoded as "*Correctly identified as depression*" and "*Not correctly identified as depression*". Following this, participants were asked to rate how distressing they believed the experience of the target to be (1 = *Not distressing at all*, 7 = *Extremely distressing*), the difficulty of treatment (1 = *Not difficult at all*, 7 = *Extremely difficult*), sympathy for the target (1 = *Not sympathetic at all*, 7 = *Extremely sympathetic*), and the likelihood that they would recommend the target seek help assuming they were friends (1 = *Not at all*, 7 = *Definitely*).

Finally, participants were asked to rate the likelihood of recommending help for the target from nine different sources (None, Friends, Parents, Other family members, Counsellor, General practitioner, Psychologist/psychiatrist, Books, and Internet) on a 5-point scale (1 = *Not very likely*, 5 = *Very likely*). For the purposes of the present study, these data were subjected to a principal-axis factor analysis with oblique (direct oblimin) rotation. The same three latent factors with eigenvalues larger than unity were obtained for the films with the female (50.8% of variance explained) and male targets (48.7% of variance explained). The extracted factors were theoretically sound, as they represented Social Support (friends,

parents, other family members; Cronbach's α = .87, .87 for female and male targets),

Professional Help (counsellor, psychologist/psychiatrist, general practitioner, and reverse-

coded none; α = .69, .71), and Self-Help (books, Internet; α = .59, .53). We, therefore,

computed mean scores based on the results of the factor analysis.

**Conformity to masculine norms**. Participants were asked to complete the 46-item

short form (Parent & Moradi, 2009) of the Conformity to Masculine Norms Inventory

(CMNI; Mahalik et al., 2003), a German translation of which was developed for use in the

present study using the parallel-blind technique (Behling & Law, 2000). The CMNI-46 is a

measure of conformity to masculine norms along nine dimensions, namely Winning (6 items;

sample item: "In general, I will do anything to win"), Emotional Control (6 items; sample

item: "I never share my feelings"), Risk-Taking (5 items; sample item: "I enjoy taking risks"),

Violence (6 items; sample item: "Sometimes, violent action is necessary), Power Over

Women (4 items; sample item: "In general, I control the women in my life"), Playboy (4

items; sample item: "If I could, I would frequently change sexual partners"), Self-Reliance (5

items; sample item: "I hate asking for help"), Primacy of Work (4 items; sample item: "My

work is the most important part of my life"), and Heterosexual Self-Presentation (6 items;

sample item: "I would be furious if someone thought I was gay"). All items were rated on a 4-

point scale ranging from 1 (*Strongly disagree*) to 4 (*Strongly agree*). As an attention check, a

control question was added to the scale, which read: "Please select 'strongly agree' here".

Participants, who did not answer this question with *strongly agree*, were excluded from the study

(see Participants section, above).

Eighteen items were reverse-coded and subscale scores were computed as the mean of

items associated with each factor, so that higher scores reflected greater conformity to

masculine norms. Although scores on the CMNI-46 are multi-dimensional, we used total

scores to decrease complexity and for ease of interpretation. However, exploratory analyses

were also carried out with the subscale scores of Emotional Control, Risk-Taking, and Self-

Reliance, selected as being the most relevant constructs to mental health literacy. These exploratory analyses are presented in a respective subsection of the Results. Parent and Moradi (2009) reported that scores on the CMNI-46 have adequate internal consistencies, acceptable psychometric properties, and good construct and factorial validity. In addition, scores on the CMNI-46 have been shown to be invariant across gender (Parent & Smiler, 2013). In the present study, sample reliabilities (Cronbach's α) for subscale scores ranged from .73 to .92.

**Expressivity.** Participants were asked to complete the Gender Typicality Scale (GTS+; German original: Alstötter-Gleich, 2004). The GTS+ was developed based on the Bem Sex Role Inventory (BSRI; Bem, 1974), which contains items that are equally socially desirable for both genders, but widely seen as more typical for either women or men. The GTS+ thus assesses the degree of masculinity and femininity as separate subscales among both men and women, generally yielding somewhat larger sex differences on these dimensions than the original BSRI (Voracek et al., 2011). Participants were asked to rate 16 adjectives describing either typically feminine attributes (i.e., expressivity; sample item: "empathetic") or typically masculine attributes (i.e., instrumentality; sample item: "fearless"). Feminine and masculine items were presented in alternating order. Participants were asked to describe to what extent these characteristics accurately described them in daily life (1 = *Rarely*, 4 = *Almost always*). Altstötter-Gleich (2004) reported that the GTS+ measure has an underlying 2-dimensional factor structure (of 8 items each) and that these two scales show adequate internal consistency as well as adequate indices of construct validity insofar as scale scores are meaningfully associated with a range of theoretically-derived personality and psychological health constructs (see also Voracek et al., 2011). Cronbach's α in the present sample was .82 for Expressivity and .81 for Instrumentality.

For analytic purposes, we conducted a 2-step cluster analysis with the GTS+ subscales to identify stereotypically masculine and feminine subgroups. Theoretically, one should be

able to distinguish four subgroups within the 2-dimensional orthogonal space of expressivity and instrumentality: persons high both in expressivity and instrumentality ("androgynous" in Bem's, 1974, terminology), persons high in expressivity and low in instrumentality (stereotypically feminine), persons low in expressivity and high in instrumentality (stereotypically masculine), and persons low both in expressivity and instrumentality ("undifferentiated" according to the terminology of Bem, 1974). We were interested in finding out which subgroups could be empirically identified in our sample and proceeded to use this classification for further analysis. The cluster analysis produced a 2-cluster solution, with Expressivity being the most important predictor. We, therefore, divided the sample into two groups, expressing high or low stereotypically feminine traits. Average cluster quality was fair with a silhouette coefficient of 0.50. Although both clusters contained a considerable heterogeneity in participant gender, clusters were significantly related to gender, $\chi^2(1) = 18.43$ ($p < .001$). Based on odds ratios ($OR$), women were 2.47 (95% $CI = 1.63, 3.75$) times more likely to fall into the high Expressivity cluster than men.

**Procedures**

All procedures in this study adhered to the ethical standards of the 1964 Helsinki Declaration and its later amendments or comparable ethical standards, and with institutional guidelines of the [blinded for review]. Participation in the study did not affect the physical or psychological integrity, the right for privacy, or other personal rights or interests of participants and, as such, according to national law, this study was exempt from formal ethical approval. An online survey was created and hosted on SoSci Survey (https://www.soscisurvey.de) and participants were recruited via advertisements placed on social networking platforms between January 31 and March 30, 2020. All participants provided digital informed consent before being informed that they would need to use headphones/speakers and complete the survey in a quiet environment (repeated before the presentation of each film). Participants were then asked to provide their demographic

information (gender, age, nationality, highest educational qualification, educational field)

before being randomly presented with the film of the female ($n$ = 200) or male target ($n$ =

205). At the end of the film, the next page was automatically loaded and the film-based

questions described above were presented. Following this, participants were asked to

complete a 3-minute numeric distraction task (Verschuere et al., 2018) before viewing the

complementary film. This, in turn, was followed by a request to complete the GTS+ and

CMNI-46. Finally, participants received debriefing information and were asked to provide an

email address (stored independently of their survey responses) to be eligible for entry into two

prize draws worth €50 each.

## Results

### Confirmatory Analyses

**Symptom recognition.** We first examined whether participants thought "anything was

wrong" with the targets presented in the films. Participants indicated that something was

wrong in 91.9% of cases when the target was female and 95.6% of cases when the target was

male. To test for possible order effects (male vs. the female target presented first), we

performed two logistic regression analyses, using indication of "something was wrong" as

outcome and order of presentation as predictor. For the female target, indicating "something

was wrong" was more likely when the male target was presented first ($p$ = .002; $OR$ = 4.00,

95% $CI$ = 1.70, 9.46). This was indicative of a carry-over effect of the male target on the

female target. There was no such effect for the male target when the female target was

presented first ($p$ = .59). In order to control for this order effect, we utilised for further

analysis generalised linear mixed models (McCullagh & Nelder, 1989), which allow for a

binary outcome variable, using target gender as a within-subjects factor (female vs. male

target) and participant gender (women vs. men) and order of presentation (male vs. female

target presented first) as between-subjects factors. Specifically, we specified an unstructured

covariance matrix for the within-subjects factor. The model tested for main effects and the

interaction between target and participant gender.

Controlling for order effects, participants were overall more likely to indicate

"something was wrong" for the male than the female target, $OR = 2.60$ (95% $CI = 1.07, 6.33$),

$p =.035$. There were no effect of participant sex or indication of an interaction between target

and participant gender ($ps \geq .35$). Of those who believed that something was wrong with the

female target, the most common description of what was wrong was depression (69.8%),

followed by responses that were reflective of non-clinical distress (15.9%). Correctly

identifying the symptoms as reflective of depression did not depend on order of presentation

(logistic regression analyses; female target: $p = .62$; male target: $p = .12$). In the generalised

linear model (without the between-subjects factor of order), there were also no effects of

participant gender ($p = .09$), target gender ($p = .52$), or their interaction ($p = .21$).

**Distress, treatment difficulty, and sympathy.** Ratings of perceived distress,

treatment difficulty, and sympathy (see Table 1) were analysed individually with linear mixed

models (McLean et al., 1991), using target gender as a within-subjects factor (female vs. male

target) and participant gender as a between-subjects factor (women vs. men). We opted for

linear mixed models because of their flexibility in data modelling compared to conventional

mixed-design analyses of variance (e.g., concerning the assumption of variance

homogeneity). Again, we specified an unstructured covariance matrix for the within-subjects

factor. The baseline models tested for main effects and the interaction of the two factors.

Subsequent models either controlled for conformity to masculine norms (using centred

CMNI-46 total scores) or expressivity (using the two GTS+ clusters as a further between-

subjects factor, contrasting the clusters low vs. high in expressivity) to assess their impact on

the effects in the baseline models. For parsimony, the here presented models did not

incorporate interactions between participant gender and CMNI-46 total scores (as stated in the

preregistration). Respective tests did not reach significance and were therefore omitted from

the final models. Possible order effects were tested prior to the here presented analyses with linear mixed models as well. There were no order effects ($ps \geq .64$), except for treatment difficulty: ratings were overall higher when the male target was presented first ($d = 0.30$, 95% $CI = 0.10, 0.49$, $p = .003$). Again, this was indicative of a carry-over effect of the male target on the female target. Order of presentation was controlled for in all analyses of treatment difficulty.

The baseline models indicated significant main effects of target gender for perceived distress and treatment difficulty, and significant main effects of participant gender for perceived treatment difficulty and ratings of sympathy (see Table 2). However, effects were qualified by interactions for perceived distress and treatment difficulty. The female target's condition was rated as less distressing than the male target condition, but only among men (see Table 1; $d = -0.24$, 95% $CI = -0.44, -0.05$). The male target's condition was rated as more difficult to treat than the female target's condition, but again only among men ($d = 0.26$, 95% $CI = 0.06, 0.45$). Regardless of target's gender, women provided higher ratings of sympathy than men ($d = 0.37$, 95% $CI = 0.17, 0.56$).

Men reported significantly higher conformity to masculine norms (CMNI-46 total scores; see Table 1) than women, $t(398) = 8.44$, $p < .001$, $d = 0.85$, 95% $CI = 0.65, 1.05$, and also had higher subscale scores in Risk-Taking ($d = 0.46$, $p < .001$), but not in Emotional Control or Self-Reliance ($ds = 0.19$ and $-0.06$, $ps \geq .053$). However, conformity to masculine norms (CMNI-46 total scores) mostly did not impact the patterns observed in the baseline models (see Table 2). CMNI-46 total scores were positively associated with ratings of perceived distress and negatively with ratings of sympathy. Controlling for CMNI-46 total scores, the main effect of participant gender in ratings of sympathy was somewhat diminished in size and was nominally not significant anymore ($d = 0.19$, 95% $CI = -0.01, 0.39$; $p = .057$).

Participants in the high-expressivity cluster had higher ratings in sympathy than participants in the low-expressivity cluster (Table 2; $d = 0.38$, 95% $CI = 0.17, 0.58$), but controlling for expressivity did not change the patterns observed in the baseline models.

**Likelihood of recommending help and sources of help.** Rating of the likelihood of recommending help and the three sources of help (see Table 1) were similarly analysed with linear mixed models (see Table 3). There were no order effects of target condition ($ps \geq .47$). Women more likely recommended help than men ($d = 0.27$, 95% $CI = 0.08, 0.47$) and also more likely recommended professional help than men ($d = 0.35$, 95% $CI = 0.15, 0.55$). At the same time, both men and women were less likely to recommended professional help when the target was female as opposed to male ($d = -0.04$, 95% $CI = -0.07, -0.004$; controlling for the correlation between the two conditions).

Controlling for conformity to masculine norms (see Table 3), the main effect of participant gender for the likelihood of recommending help was diminished in size and lost its nominal significance ($d = 0.16$, 95% $CI = -0.04, 0.36$; $p = .108$). The main effect of participant gender for the likelihood of recommending professional help was similarly diminished in size ($d = 0.20$, 95% $CI = 0.003, 0.40$), but remained significant.

CMNI-46 total scores were negatively associated with ratings of recommending help and recommending professional help, and positively associated with ratings of recommending self-help. Controlling for CMNI-46 total scores increased the main effect of participant gender in these ratings, thus rendering them nominally significant. Accordingly, women appeared more likely to recommend self-help than men, after controlling for CMNI-46 total scores ($d = 0.24$, 95% $CI = 0.04, 0.44$).

**Exploratory Analyses**

Exploratory analyses of the ratings of perceived distress, treatment difficulty, sympathy, the likelihood of recommending help, and the sources of help controlled for the three CMNI-46 subscale scores of Emotional Control, Risk-Taking, and Self-Reliance

(instead for CMNI-46 total scores) in a common model. This did not change the previously obtained pattern of results (see Supplemental Materials, Tables S1 and S2). Emotional control was negatively associated with ratings of sympathy and the likelihood of recommending help, and positively associated with recommending social support. Self-Reliance was negatively associated with the latter rating. Risk-Taking was positively associated with recommending self-help.

## Discussion

In the present study, we used a novel method – namely, filmed disclosures – to examine mental health literacy of depression in a sample of adults from Austria in a randomised, fully within-subjects design. A first key finding from this study was that, irrespective of the target gender, almost all (93.8%) participants indicated that "something was wrong" and, of these participants, the majority (69.2%) correctly identified the targets as displaying symptoms of depression. Although direct comparisons with previous studies should be conducted cautiously because of sampling and methodological differences, it is notable that correct symptomatology identification in the present study was substantively higher than that reported recently among adults from China (36.9%; Huang et al., 2019), Vietnam (32.0%; Nguyen Thai & Nguyen, 2018), and the United Kingdom (53.4%; Swami, 2012), though admittedly not Australia (86.0%; Gibbons et al., 2015), where community education programmes are relatively well-developed (Jorm et al., 2007). Figures of correct symptomatology identification were also substantively higher than corresponding results of earlier studies with samples from central Europe (e.g., 39.8% in Lauber et al., 2003).

One possible explanation for the higher proportion of correct responses in the present study is our use of filmed disclosures. Although one previous study with Australian rural adolescents suggested that the use of filmed disclosures resulted in lower symptom recognition compared to the use of text-based vignettes (Marshall & Dunstan, 2013), it is also

possible that the use of filmed disclosures provides richer contextual information (e.g., gesture, facial expressions, tone of voice), which are more likely to activate relevant mental models that, in turn, facilitate symptom recognition or, at the very least, a perception that "something is wrong" (Busselle & Bilandzic, 2008; Murphy et al., 2013). Consistent with this suggestion, previous work has shown that exposure to social contact films (e.g., interviews with individuals with mental health disorders) resulted in significantly improved knowledge of mental health and reduced stigma (Koike et al., 2018; Yamaguchi et al., 2019). Of course, given the study design, we cannot rule out the possibility that the higher rates of symptom recognition in the present study compared to other recent and older studies reflect broader trends in improving mental health literacy (e.g., Schomerus et al., 2012).

In the present study, we also examined mental health literacy as a function of participant and target gender. However, in contrast to ~~both our~~ Hypothesis 1 and previous studies (e.g., Cotton et al., 2016; Gibbons et al., 2015; Gorczynski et al., 2017; Hadjimina & Furnham, 2017; Lee et al., 2020; Swami, 2012), we found little evidence of gender differences in symptom recognition as function of either participant or target gender. One exception was the finding that respondents were somewhat more likely to indicate that something was wrong when the target was male (95.6%) compared to female (91.9%), but even here the magnitude of the difference was negligible. It is possible that these null effects are again reflective of the use of filmed disclosures, rather than text-based vignettes. Given the likelihood that filmed disclosures enhance perceived realism in comparison to information delivered through text alone (Swami & Furnham, 2018), it might be suggested that filmed disclosures reduce or minimise any gendered discrepancies. To put it differently, it may be that the use of filmed disclosures improves the likelihood that participants correctly identify symptomatology in male targets and also that men will be more likely to correctly identify symptomatology in targets of all genders. Related to this explanation is the possibility of

ceiling effects: given the relatively high levels of symptom recognition in the present study, it is possible that any gendered effects become less easy to detect.

Beyond symptom recognition, there were also few gendered effects in terms of attitudes toward the cases of depression (Hypothesis 2). Men were more likely than women to indicate that the female target's condition was distressing and that the male target's condition was more difficult to treat, but effect sizes were small. In contrast, women provided higher ratings of sympathy compared to men irrespective of the target's gender, which likely reflect women's greater motivation and capacity for understanding others' thoughts, feelings, and experiences (e.g., Klein & Hodges, 2001; Macaskill et al., 2002). Women were also more likely than men to recommend that the targets should seek help, which is consistent with previous work (e.g., Swami, 2012), although the effect of target gender was not significant. In addition, women were also more likely than men to recommend that help be sought from professional sources (i.e., a counsellor, general practitioner, or psychologist/psychiatrist), while there were no gendered effects in terms of recommending help from social sources or self-help. Interestingly, both women and men were less likely to recommend professional help when the target was female as opposed to male, which also runs counter to our general expectations.

In addition, we examined the impact of conformity to masculine norms on attitudes toward the cases of depression (Hypothesis 3), but all effects were generally null or weak. Greater conformity to masculine norms was, however, associated with significantly greater ratings of perceived distress and lower ratings of sympathy, and controlling for total CMNI-46 rendered the gender difference in ratings of sympathy non-significant. Greater conformity to masculine norms was also significantly associated with lower recommendations to seek help and to seek help from professional sources, but were significantly associated with stronger recommendations to self-help. In broad outline, these findings are consistent with previous findings suggesting that conformity to masculine norms are negatively associated

with help-seeking attitudes for mental health (Clark et al., 2020; Hammer et al., 2013; McDermott et al., 2018; Seidler et al., 2016; Vogel et al., 2011; Wong et al., 2017) and more negative attitudes toward psychiatric services (Berger et al., 2013). It should be noted, however, that effect sizes in the present study were relatively weak.

Of course, it should be noted that the above analyses were run with total CMNI-46 scores, which may obscure lower-order associations with facets of conformity to masculinity. However, running the exploratory analyses with selected CMNI-46 facets that were selected as being the most relevant to mental health literacy did not alter our basic pattern of findings. In addition, the present study also examined the impact of expressivity – a core component of feminine gender role endorsement – on attitudes toward the cases of depression, but effects were non-significant beyond the finding that participants in the high expressivity cluster had higher ratings in sympathy than participants in the low expressivity cluster. That is, beyond the general finding that those higher in expressivity were more likely to be sympathetic to targets with depression, endorsement of expressivity did not appear to impact attitudes toward the cases. It should be noted, however, that expressivity may be better conceptualised as a personality trait rather than a gender role (Auster & Ohm, 2000; Woodhill & Samuels, 2003), and future research could extend our findings by including measuring feminine gender role endorsement more directly, such as through the use of the Conformity to Feminine Norms Inventory (Mahalik et al., 2005).

**Strengths and Limitations**

Major strengths of the present study include: (1) study preregistration; (2) the fully within-subjects design; (3) the randomised stimulus sequence; (4) the utilization of filmed disclosures instead of text vignettes; (5) open study materials and data. While the present study thereby extends the literature on mental health literacy in a number of important ways, several limitations should be acknowledged. First, although efforts were made to standardise the filmed disclosures across target gender, it is possible that there were subtle differences

across conditions that influenced results. For instance, although the quality of the two films and perceived severity of the cases were rated as similar across target gender in our pilot study, the male actor was rated as significantly more believable (although the effect size of the difference was small). We, therefore, cannot rule out the possibility that the lack of gendered effects in the present study was, at least partly, a function of the believability of the actors. However, this criticism needs to be balanced by the likelihood that the use of filmed disclosures increases ecologic validity compared to text-based vignettes, which are straightforward and do not require the same levels of active interpretation (cf. Wijeratne & Harris, 2009).

In a related vein, it may be that responses to filmed disclosures by singular actors are idiosyncratic to some extent, shaped by difficult-to-control factors such as likeability, physical attractiveness, or perceived similarity (see Swami, 2021). Increasing the number of utilised actors would afford greater possibilities for controlling for factors such as these, but would also bring new practical challenges (e.g., the cost of producing additional films). Relatedly, while we maintain that filmed disclosures are more ecologically valid than the use of text-based vignettes, films are unlikely to fully mirror the way disclosures about mental health occur in everyday life. The latter are likely to be much more discursive, with room for interaction, question-asking, and reciprocal self-disclosure – all of which are absent in our films. This issue may be improved upon in the future through the use of actors in conversation with participants, which may be useful in generating in-depth qualitative narratives by respondents.

Further, although efforts were made to reduce the likelihood that participants would have noticed similarities across the two films (e.g., re-arranging the order of presentation of symptoms, the use of a distraction task), we observed order effects for two of the investigated outcomes ("something was wrong" and treatment difficulty), which may be a consequence of differences in the believability of the two actors in the filmed disclosures. The use of even

more similar videos or of a between-subjects design may be desirable in future research, though larger subsamples would be needed in trade-off. Relatedly, although the recruitment of a community sample avoids the limitations of a college sample, our method of recruitment means that our sample is unlikely to be representative of the wider population in Austria and/or Central Europe. The recruitment of a representative sample of the population would, therefore, be welcome in future research. Finally, it should be noted that the present study was not specifically set up to examine the utility of filmed disclosures in relation to text-based vignettes. By necessity, we have cautiously drawn comparisons across methodologies in this discussion, but recommend that future work specifically compare outcomes as a function of methodology (e.g., Marshall & Dunstan, 2013).

**Conclusion**

The present study provides new data on mental health literacy of depression in an adult sample from Austria using a methodology that has been infrequently applied in extant research. Using filmed disclosures, we found that participants' ability to correctly identify cases of depression was substantively higher than in a number of recent studies that have used text-based vignettes. One tentative conclusion, which will require replication, is that text-based vignettes – which offer limited naturalism compared to filmed disclosures – may artefactually deflate or underestimate symptom recognition abilities. A second important finding from the present study was the general lack of gendered effects, as well as the lack of effects of conformity to masculine norms and expressivity. It is difficult to conclusively point to reasons for this, although one possibility is that gendered effects are less likely to emerge when the naturalism of stimuli is improved. Given that the basic methodology for assessing mental health literacy has not changed substantively in over two decades of research (Wei et al., 2015), the use of filmed disclosures may offer a much-needed and improved tool in the arsenal of researchers.

# References

Alonso, J., Codony, M., Kovess, V., Angermeyer, M. C., Katz, S. J., Haro, J. M., De
 Girolamo, G., De Graaf, R., Demyttenaere, K., Vilagut, G., Almansa, J., Lépine, J. P., &
 Brugha, T. S. (2007). Population level of unmet need for mental healthcare in Europe.
 *British Journal of Psychiatry*, *190*(4), 299-306.
 https://doi.org/10.1192/bjp.bp.106.022004

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental
 Disorders* (5th ed.). American Psychiatric Association.

Altstötter-Gleich, C. (2004). Expressivität, Instrumentalität und psychische Gesundheit: Ein
 Beitrag zur Validierung einer Skala zur Erfassung des geschlechtsrollenbezogenen
 Selbstkonzepts [Expressivity, instrumentality, and mental health: Validation of a scale
 assessing gender role self-concept]. *Zeitschrift für Differentielle und Diagnostische
 Psychologie*, *25*(3), 123-139. https://doi.org/10.1024/0170-1789.25.3.123

Auster, C. J., & Ohm, S. C. (2000). Masculinity and femininity in contemporary American
 society: A reevaluation using the Bem Sex-Role Inventory. *Sex Roles*, *43*(7/8), 499-528.
 https://doi.org/10.1023/A:1007119516728

Behling, O., & Law, K. (2000). *Translating questionnaires and other research instruments*.
 Sage. https://doi.org/10.4135/9781412986373

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and
 Clinical Psychology*, *42*(2), 155-162. https://doi.org/10.1037/h0036215

Berger, J. L., Addis, M. E., Green, J. D., Mackowiak, C., & Goldberg, V. (2013). Men's
 reactions to mental health labels, forms of help-seeking, and sources of help-seeking
 advice. *Psychology of Men and Masculinity*, *14*(4), 433-443.
 https://doi.org/10.1037/a0030175

Boonstra, N., Klaasen, R., Sytema, S., Marshall, M., de Haan, L., Wunderink, L., & Wiersma, D. (2012). Duration of untreated psychosis and negative symptoms: A systematic review and meta-analysis of individual patient data. *Schizophrenia Research*, *142*(1-3), 12-19. https://doi.org/10.1016/j.schres.2012.08.017

Busselle, R., & Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: A model of narrative comprehension and engagement. *Communication Theory*, *18*(2), 255-280. https://doi.org/10.1111/j.1468-2885.2008.00322.x

Calear, A. L., Banfield, M., Batterham, P. J., Morse, A. R., Forbes, O., Carron-Arthur, B., & Fisk, M. (2017). Silence is deadly: A cluster-randomised controlled trial of a mental health help-seeking intervention for young men. *BMC Public Health*, *17*, 834. https://doi.org/10.1186/s12889-017-4845-z

Chen, S., Wu, Q., Qi C., Deng, H., Wang, X., He, H., Long, J., Xiong, Y., & Liu, T. (2017). Mental health literacy about schizophrenia and depression: A survey among Chinese caregivers of patients with mental disorder. *BMC Psychiatry*, *17*, 89. https://doi.org/10.1186/s12888-017-1245-y

Clark, L. H., Hudson, J. L., Rapee, R. M., & Grasby, K. L. (2020). Investigating the impact of masculinity on the relationship between anxiety specific mental health literacy and mental health help-seeking in adolescent males. *Journal of Anxiety Disorders*. Advance online publication. https://doi.org/10.1016/janxdis.2020.102292

Clement, S., Schauman, O., Graham, T., & Maggioni, F. (2015). What is the impact of mental health-related stigma on help-seeking? A review of quantitative and qualitative studies. *Psychological Medicine*, *45*(1), 11-27. https://doi.org/10.1017/S0033291714000129

Cotton, S. M., Wright, A., Harris, M. G., Jorm, A. F., & Mcgorry, P. D. (2006). Influence of gender on mental health literacy in young Australians. *Australian and New Zealand Journal of Psychiatry*, *40*(9), 790-796. https://doi.org/10.1080/j.1440-1614.2006.01885.x

Courtenay, W. H. (2000). Constructions of masculinity and their influence on men's well-

being: a theory of gender and health. *Social Science and Medicine*, *50*(10), 1385-1401.

https://doi.org/https://doi.org/10.1016/S0277-9536(99)00390-1

Dell'Osso, B., Glick, I. D., Baldwin, D. S., & Altamura, A. C. (2013). Can long-term

outcomes be improved by shortening the duration of untreated illness in psychiatric

disorders? A conceptual framework. *Psychopathology*, *46*(1), 14-21.

https://doi.org/10.1159/000338608

Furnham, A., Daoud, Y., & Swami, V. (2009). "How to spot a psychopath": Lay theories of

psychopathy. *Social Psychiatry and Psychiatric Epidemiology*, *44*, 464-472.

https://doi.org/10.1007/S00127-008-0459-1

Furnham, A., & Sjokvist, P. (2017). Empathy and mental health literacy. *Health Literacy and

Research and Practice*, *1*(2), e31-e40. https://doi.org/103928/24748307-20170328-01

Furnham, A., & Swami, V. (2018). Mental health literacy: A review of what it is and why it

matters. *International Perspectives in Psychology*, *7*(4), 240-257.

https://doi.org/10.1037/ipp0000094

Gair, S., & Camilleri, P. (2003). Attempting suicide and help-seeking behaviours: Using

stories from young people to inform social work practice. *Australian Social Work*,

*56*(2), 83-93. https://doi.org/10.1046/j.0312-407X.2003.00064.x

Galasiński, D. (2008). *Men's discourses of depression.* Palgrave Macmillan.

GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. (2018). Global,

regional, and national incidence, prevalence, and years lived with disability for 354

diseases and injuries for 195 countries and territories, 1990-2017: A systematic analysis

for the Global Burden of Disease Study 2017. *The Lancet*, *392*(10159), 1789-1858.

https://doi.org/10.1016/S1040-6736(18)32279-7

Gibbons, R. J., Thorsteinsson, E. B., & Loi, N. M. (2015). Beliefs and attitudes towards

mental illness: An examination of the sex differences in mental health literacy in a

community sample. *PeerJ*, *3*, e1104. https://doi.org/10.7717/peerj.1004

Gilchrist, H., & Sullivan, G. (2006). Barriers to help-seeking in young people: Community

beliefs about youth suicide. *Australian Social Work*, *59*(1), 73-85.

https://doi.org/10.1080/0312407050044979

Gorczynski, P., Sims-Schouten, W., Hill, D., & Wilson, J. C. (2017). Examining mental

health literacy, help seeking behaviours, and mental health outcomes in UK university

students. *Journal of Mental Health Training, Education and Practice*, *12*(2), 111-120.

https://doi.org/10.1108/JMHTEP-05-2016-0027

Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to

mental health help-seeking in young people: A systematic review. *BMC Psychiatry*, *10*,

113-128. https://doi.org/10.1186/1471-244X-10-113

Hadjimina, E., & Furnham, A. (2017). Influence of age and gender on mental health literacy

of anxiety disorders. *Psychiatry Research*, *251*, 8-13.

https://doi.org/https://doi.org/10.1016/j.psychres.2017.01.089

Hammer, J. H., Vogel, D. L., & Heimerdinger-Edwards, S. R. (2013). Men's help seeking:

Examination of differences across community size, education, and income. *Psychology

of Men and Masculinity*, *14*(1), 65-75. https://doi.org/10.1037/a0026813

Holzinger, A., Floris, F., Schomerus, G., Carta, M. G., & Angermeyer, M. C. (2012). Gender

differences in public beliefs and attitudes about mental disorder in western countries: A

systematic review of population studies. *Epidemiology and Psychiatric Sciences*, *21*(1),

73–85. https://doi.org/10.1017/S2045796011000552

Huang, D., Yang, L. H., & Pescosolido, B. A. (2019). Understanding the public's profile of

mental health literacy in China: A nationwide study. *BMC Psychiatry*, *19*, 20.

https://doi.org/10.1186/s12888-018-1980-8

Hurley, D., Swann, C., Allen, M. S., Ferguson, H. L., & Vella, S. A. (2020). A systematic review of parent and caregiver mental health literacy. *Community Mental Health Journal*, *56*, 2-21. https://doi.org/10.1007/s10597-019-00454-0

Johnson, J. L., Oliffe, J. L., Kelly, M. T., Galdas, P., & Ogrodniczuk, J. S. (2012). Men's discourses of help-seeking in the context of depression. *Sociology of Health and Illness*, *34*(3), 345-361. https://doi.org/10.1111/j.1467-9566.2011.01372.x

Jorm, A. F. (2012). Mental health literacy: Empowering the community to take action for better mental health. *American Psychologist, 67*(3), 231-243. https://doi.org/10.1037/a0025957

Jorm, A. F. (2015). Why we need the concept of "mental health literacy". *Health Communication*, *30*(12), 1166-1168. https://doi.org/10.1080/10410236.2015.1037423

Jorm, A. F. (2019). The concept of mental health literacy. In O. Okan, U. Bauer, Levin-Zamir, D., Pinheiro, P., & Sørensen, K. (Eds.), *International handbook of health literacy: Research, practice and policy across the lifespan* (pp. 53-65). Policy Press.

Jorm, A. F., Kitchener, B. A., Kanowski, L. G., & Kelly, C. M. (2007). Mental health first aid training for members of the public. *International Journal of Clinical and Health Psychology*, *7*(1), 141-151.

Jorm, A. F., Korten, A. E., Jacomb, P. A., Christensen, H., Rodgers, B., & Pollitt, P. (1997). Public beliefs about causes and risk factors for depression and schizophrenia. *Social Psychiatry and Psychiatric Epidemiology*, *32*(3), 143-148. https://doi.org/10.1007/BF00794613

Jorm, A. F., Nakane, Y., Christensen, H., Yoshioka, K., Griffiths, K. M., & Wata, Y. (2005). Public beliefs about treatment and outcome of mental disorders: A comparison of Australia and Japan. *BMC Medicine*, *3*, 12-28. https://doi.org/10.1186/1741-7015-3-12

Kelly, C. M., Jorm, A. F., & Wright, A. (2007). Improving mental health literacy as a strategy to facilitate early intervention for mental disorders. *Medical Journal of Australia*, *187* (7 Suppl), S26-S30. https://doi.org/10.5694/j.1326-5377.2007.tb01332.x

Kim, J. E., Saw, A., & Zane, N. (2015). The influence of psychological symptoms on mental health literacy of college students. *American Journal of Orthopsychiatry*, *85*(6), 620-630. https://doi.org/10.1037/ort0000074

Klein, K. J. K., & Hodges, S. D. (2001). Gender differences, motivation, and empathic accuracy: When it pays to understand. *Personality and Social Psychology Bulletin*, *27*(6), 720-730. https://doi.org/10.1177/0146167201276007

Koike, S., Yamaguchi, S., Ojio, Y., Ohta, K., Shimada, T., Watanabe, K., Thornicroft, G., & Ando, S. (2018). A randomised controlled trial of repeated filmed social contact on reducing mental illness-related stigma in young adults. *Epidemiology and Psychiatric Sciences*, *27*(2), 199-208. https://doi.org/10.1017/S2045796016001050

Krumm, S., Checchia, C., Koesters, M., Killiam, R., & Becker, T. (2017). Men's views on depression: A systematic review and metasynthesis of qualitative research. *Psychopathology*, *50*(2), 107-124. https://doi.org/10.1159/000455256

Kutcher, S., Bagnell, A., & Wei, Y. (2015). Mental health literacy in secondary schools: A Canadian approach. *Child and Adolescent Psychiatric Clinics of North America*, *24*(2), 233-244. https://doi.org/10.1016/j.chc.2014.11.007

Kutcher, S., Wei, Y., & Coniglio, C. (2016). Mental health literacy: Past, present, and future. *Canadian Journal of Psychiatry*, *61*(3), 154-158. https://doi.org/10.1177/0706743715616609

Lauber, C., Nordt, C., Falcato, L., & Rössler, W. (2003). Do people recognise mental illness? *European Archives of Psychiatry and Clinical Neurosciences*, *253*(5), 248-251. https://doi.org/10.1007/s00406-003-0439-0

Lee, H. Y., Hwang, J., Ball, J. G., Lee, J., Yu, Y., & Albright, D. L. (2020). Mental health

    literacy affects mental health attitude: Is there a gender difference? *American Journal of*

    *Health Behavior*, *44*(3), 282-291. https://doi.org/10.5993/AJHB.44.3.1

Macaskill, A., Maltby, J., & Day, L. (2002). Forgiveness of self and others and emotional

    empathy. *The Journal of Social Psychology*, *142*(5), 663-665.

    https://doi.org/10.1080/00224540209603925

Mahalik, J. R., Locke, B. D., Ludlow, L. H., Diemer, M. A., Scott, R. P., Gottfriend, M., &

    Freitas, G. (2003). Development of the Conformity to Masculine Norms Inventory.

    *Psychology of Men & Masculinity*, *4*(3), 3-25. https://doi.org/10.1037/1524-9220.4.1.3

Mahalik, J. R., Morray, E. B., Coonerty-Femiano, A., Ludlow, L. H., Slattery, S. M., &

    Smiler, A. (2005). Development of the Conformity to Feminine Norms Inventory. *Sex*

    *Roles*, *52*(7/8), 417-435. https://doi.org/10.1007/S11199-0053709-7

Marshall, J. M., & Dunstan, D. A. (2013). Mental health literacy of Australian rural

    adolescents: An analysis using vignettes and short films. *Australian Psychologist*, *48*(2),

    119-127. https://doi.org/10.1111/j.1742-9544.2011.00048.x

McDermott, R. C., Smith, P. N., Borgogna, N., Booth, N., Granato, S., & Sevig, T. D. (2018).

    College students' conformity to masculine role norms and help-seeking intentions for

    suicidal thoughts. *Psychology of Men and Masculinity*, *19*(3), 340-

    351. https://doi.org/10.1037/men0000107

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). Chapman and

    Hall.

McLean, R. A., Sanders, W. L., & Stroup, W. W. (1991). A unified approach to mixed linear

    models. *The American Statistician*, *45*(1), 54-64.

Murphy, S. T., Frank, L. B., Chatterjee, J. S., & Baezconde-Garbanti, L. (2013). Narrative

    versus non-narrative: The role of identification, transportation and emotion in reducing

health disparities. *Journal of Communication*, *63*(1), 116-137.

https://doi.org/10.1111/jcom.12007

Nguyen Thai, Q. C., & Nguyen, T. H. (2018). Mental health literacy: Knowledge of

depression among undergraduate students in Hanoi, Vietnam. *International Journal of*

*Mental Health Systems*, *12*, 19. https://doi.org/10.1186/s13033-018-0195-1

Oliffe, J. L., Robertson, S., Kelly, M.T., Roy, P., & Ogrodniczuk, J. (2010). Connecting

masculinity and depression among international male university students. *Qualitative*

*Health Research*, *20*(7), 987-998. https://doi.org/10.1177/1049732310365700

Oliffe, J. L., Kelly, M. T., Johnson, J. L., Bottorff, J. L., Gray, R. E., Ogrodniczuk, J. S., &

Galdas, P. M. (2010). Masculinities and college men's depression: Recursive

relationships. *Health Sociology Review*, *19*(4), 465-477.

https://doi.org/10.5172/hesr.2010.19.4.465

Oute, J., Tondora, J., & Glasdam, S. (2018). "Men just drink more than women. Women have

friends to talk to": Gendered understandings of depression among healthcare

professionals and their implications. *Nursing Inquiry*, *25*(3), e12241.

https://doi.org/10.1111/nin.12241

Parent, M. C., & Moradi, B. (2011). An abbreviated tool for assessing conformity to

masculine norms: Psychometric properties of the Conformity to Masculine Norms

Inventory-46. *Psychology of Men and Masculinity*, *12*(4), 339-353.

https://doi.org/10.1037/a0021904

Parent, M. C., & Smiler, A. P. (2013). Metric invariance of the Conformity to Masculine

Norms Inventory-46 among women and men. *Psychology of Men and Masculinity*,

*14*(3), 324-328. https://doi.org/10.1037/a0027642

Schomerus, G., Schwahn, C., Holzinger, A., Corrigan, P. W., Grabe, H. J., Carta, M. G., &

Angermeyer, M. C. (2012). Evolution of public attitudes about mental illness: A

systematic review and meta-analysis. *Acta Psychiatrica Scandinavica*, *125*(6), 440-452.

https://doi.org/10.1111/j.1600-0447.2012.01826.x

Seidler, Z. E., Dawes, A. J., Rice, S. M., Oliffe, J. L., & Dhillon, H. M. (2016). The role of

masculinity in men's help-seeking for depression: A systematic review. *Clinical*

*Psychology Review*, *49*, 106-118. https://doi.org/10.1016/j.cpr.2016.09.002

Sobocki, P., Lekander, I., Borgström, F., Ström, O., & Runeson, B. (2007). The economic

burden of depression in Sweden from 1997 to 2005. *European Psychiatry*, *22*(3), 146–

152. https://doi.org/10.1016/j.eurpsy.2006.10.006

Soorkia, R., Snelgar, R., & Swami, V. (2011). Factors influencing attitudes towards seeking

professional psychological help among South Asian students in Britain. *Mental Health,*

*Religion, and Culture*, *14*(6), 619-623. https://doi.org/10.1080/13674676.2010.494176

Spiker, D. A., & Hammer, J. H. (2019). Mental health literacy as theory: Current challenges

and future directions. *Journal of Mental Health*, *28*(3), 238-242.

https://doi.org/10.1080/09638237.2018.1437613

Steel, Z., Mamane, C., Iranpour, C., Chey, T., Jackson, J. W., Patel, V., & Silove, D. (2014).

The global prevalence of common mental disorders: A systematic review and meta-

analysis, 1980-2013. *International Journal of Epidemiology*, *43*(2), 476-493.

https://doi.org/10.1093/ije/dyu038

Swami, V. (2012). Mental health literacy of depression: Gender differences and attitudinal

antecedents in a representative British sample. *PLoS ONE*, *7*(11), e49779.

https://doi.org/10.1371/journal.pone.0049779

Swami, V. (2021). *Attraction explained: The science of how we form relationships* (2nd ed.).

Routledge.

Swami, V., Barron, D., Smith, L., & Furnham, A. (2020). Mental health literacy of maternal

and paternal postnatal (postpartum) depression in British adults. *Journal of Mental*

*Health*, *29*(2), 217-224. https://doi.org/10.1080/09638237.2019.1608932

Swami, V., Papanicolaou, A., & Furnham, A. (2011). Examining mental health literacy and
its correlates using the overclaiming technique. *British Journal of Psychology*, *102*(3),
662-675. https://doi.org/10.1111/j.2044-8295.2011.02306.x

Swami, V., Persaud, R., & Furnham, A. (2011). The recognition of mental health disorders
and its association with psychiatric scepticism, knowledge of psychiatry, and the Big
Five personality factors: An investigation using the overclaiming technique. *Social
Psychiatry and Psychiatric Epidemiology*, *46*(3), 181-189.
https://doi.org/10.1007/s00127-010-0193-3

Verschuere, B., Meijer, E. H., Jim, A., Hoogesteyn, K., Orthey, R., McCarthy, R. J.,
Skowronski, J. J., Acar, O. A., Aczel, B., Bakos, B. E., Barbosa, F., Baskin, E., Bègue,
L., Ben-Shakhar, G., Birt, A. R., Blatz, L., Charman, S. D., Claesen, A., Clay, S. L., …
Yıldız, E. (2018). Registered replication report on Mazar, Amir, and Ariely (2008).
*Advances in Methods and Practices in Psychological Science*, *1*(3), 299-317.
https://doi.org/10.1177/2515245918781032

Vigo, D., Thornicroft, G., & Atun, R. (2016). Estimating the true global burden of mental
illness. *Lancet Psychiatry*, *3*(2), 171-178. https://doi.org/10.1016/S2215-
0366(15)00505-2

Vogel, D. L., Heimerdinger-Edwards, S. R., Hammer, J. H., & Hubbard, A. (2011). "Boys
don't cry": Examination of the links between endorsement of masculine norms, self-
stigma, and help-seeking attitudes for men from diverse backgrounds. *Journal of
Counseling Psychology*, *58*(3), 368–382. https://doi.org/10.1037/a0023688

Voracek, M., Pietschnig, J., Nader, I. W., & Stieger, S. (2011). Digit ratio (2D:4D) and sex-
role orientation: Further evidence and meta-analysis. *Personality and Individual
Differences*, *51*(4), 417-422. https://doi.org/10.1016/j.paid.2010.06.009

Wei, Y., McGrath, P. J., Hayden, J., & Kutcher, S. (2015). Mental health literacy measures evaluating knowledge, attitudes, and help-seeking: A scoping review. *BMC Psychiatry*, *15*, 291. https://doi.org/10.1186/s12888-015-0681-9

Wijeratne, C., & Harris, P. (2009). Late life depression and dementia: A mental health literacy survey of Australian general practitioners. *International Psychogeriatrics*, *21*(2), 330-337. https://doi.org/10.1017/S1041610208008235

Wong, Y. J., Ho, M. H. R., Wang, S. Y., & Miller, I. S. (2017). Meta-analyses of the relationship between conformity to masculine norms and mental health-related outcomes. *Journal of Counseling Psychology*, *64*(1), 80-93. https://doi.org/10.1037/cou0000176

Woodhill, B. M., & Samuels, C. A. (2003). Positive and negative androgyny and their relationship with psychological health and well-being. *Sex Roles*, *48*(11/12), 555-565. https://doi.org/10.1023/A:1023531530272

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. World Health Organization.

Yamaguchi, S., Ojio, Y., Ando, S., Bernick, P., Ohta, K., Watanabe, K., Thornicroft, G., Shiozawa, T., & Koike, S. (2019). Long-term effects of filmed social contact or internet-based self-study on mental health-related stigma: A 2-year follow-up of a randomised controlled trial. *Social Psychiatry and Psychiatric Epidemiology*, *54*(1), 33-42. https://doi.org/10.1007/s00127-018-1609-8

Yousaf, O., Popat, A., & Hunter, M. S. (2015). An investigation of masculinity attitudes, gender, and attitudes toward psychological help-seeking. *Psychology of Men and Masculinity*, *16*(2), 234-237. https://doi.org/10.1037/a0036241

**Table 1**

*Means and Standard Deviations of the Ratings and the Study Variables, Differentiated for*

*Participant Gender*

| Rating or variable | Women | Men |
|---|---|---|
| Distress | 6.29 (0.99) / 6.29 (0.96) | 5.99 (120) / 6.27 (0.98) |
| Treatment difficulty | 4.41 (1.33) / 4.38 (1.37) | 4.43 (1.36) / 4.70 (1.37) |
| Empathy | 5.77 (1.23) / 5.96 (1.03) | 5.49 (1.28) / 5.55 (1.18) |
| Recommending help | 6.62 (0.85) / 6.73 (0.70) | 6.50 (0.97) / 6.50 (1.02) |
| Professional help | 4.30 (0.62) / 4.32 (0.61) | 4.00 (0.78) / 4.09 (0.71) |
| Social support | 2.96 (1.09) / 2.97 (1.05) | 2.88 (1.14) / 2.89 (1.14) |
| Self-help | 2.02 (0.97) / 2.03 (0.98) | 1.89 (0.86) / 1.89 (0.86) |
| CMNI-46 total score | 84.16 (12.54) | 95.50 (14.26) |
| Emotional Control | 2.11 (0.80) | 2.26 (0.72) |
| Risk-Taking | 1.94 (0.60) | 2.23 (0.68) |
| Self-Reliance | 2.17 (0.78) | 2.12 (0.68) |
| GTS+ Expressivity | 2.98 (0.56) | 2.74 (0.50) |
| GTS+ Instrumentality | 2.45 (0.53) | 2.71 (0.54) |

*Note*. CMNI-46 = Conformity to Masculine Norms Inventory-46; GTS+ = Gender Typicality

Scale. Ratings are provided for female (left) and male (right) target conditions.

**Table 2**

*Results for Perceived Distress, Perceived Treatment Difficulty, and Ratings of Empathy*

| Effect | Estimate (*SE*) | | |
| --- | --- | --- | --- |
| | Baseline model | Controlling for masculinity | Controlling for expressivity |
| Distress | | | |
| Intercept | 6.27 (0.07)*** | 6.32 (0.07)*** | 6.21 (0.09)*** |
| Participant gender | 0.02 (0.10) | -0.06 (0.10) | -0.01 (0.10) |
| Target gender | -0.27 (0.08)** | -0.27 (0.09)** | -0.27 (0.09)** |
| Participant gender × target gender | 0.27 (0.11)* | 0.26 (0.11)* | 0.28 (0.11)* |
| Masculinity | | -0.008 (0.003)* | |
| Expressivity | | | 0.11 (0.09) |
| Treatment difficulty | | | |
| Intercept | 4.56 (0.12)*** | 4.51 (0.13)*** | 4.62 (0.14)*** |
| Order | 0.28 (0.12)* | 0.28 (0.12)* | 0.29 (0.12)* |
| Participant gender | -0.32 (0.14)* | -0.26 (0.15) | -0.30 (0.14)* |
| Target gender | -0.27 (0.09)** | -0.27 (0.09)** | -0.27 (0.09)** |
| Participant gender × target gender | 0.30 (0.12)* | 0.30 (0.12)* | 0.31 (0.12)* |
| Masculinity | | 0.005 (0.005) | |
| Expressivity | | | -0.14 (0.13) |
| Empathy | | | |
| Intercept | 5.55 (0.08)*** | 5.67 (0.09)*** | 5.35 (0.10)*** |

| Effect | Estimate (SE) | | |
| --- | --- | --- | --- |
| | Baseline model | Controlling for masculinity | Controlling for expressivity |
| Participant gender | 0.40 (0.11)*** | 0.22 (0.12) | 0.31 (0.11)** |
| Target gender | -0.07 (0.08) | -0.06 (0.08) | -0.07 (0.08) |
| Participant gender × target gender | -0.11 (0.10) | -0.11 (0.10) | -0.12 (0.10) |
| Masculinity | | -0.02 (0.004)*** | |
| Expressivity | | | 0.40 (0.11)*** |

*Note.* Intercepts captured the ratings of men in the male target condition. Participant gender contrasted men (baseline) vs. women, target gender the male target (baseline) vs. the female target. Order of presentation of the target conditions (male vs. female target first) was only controlled for in the analysis of treatment difficulty ratings (see main text). The measure of masculinity (CMNI-46 total scores) was centered before it was entered in the models. Expressivity contrasted clusters low (baseline) vs. high in expressivity. *$p < .05$, **$p < .01$, ***$p < .001$.

**Table 3**

*Results for Likelihood of Recommending Help and Three Sources of Help*

| Effect | Estimate (*SE*) | | |
|---|---|---|---|
| | Baseline model | Controlling for masculinity | Controlling for expressivity |
| Recommending help | | | |
| Intercept | 6.50 (0.06)*** | 6.55 (0.07)*** | 6.48 (0.08)*** |
| Participant gender | 0.23 (0.09)** | 0.15 (0.09) | 0.22 (0.09)* |
| Target gender | 0.00 (0.08) | 0.01 (0.08) | 0.00 (0.08) |
| Participant gender × target gender | -0.11 (0.10) | -0.12 (0.10) | -0.11 (0.10) |
| Masculinity | | -0.007 (0.003)** | |
| Expressivity | | | 0.05 (0.08) |
| Professional help | | | |
| Intercept | 4.09 (0.05)*** | 4.14 (0.05)*** | 4.06 (0.06)*** |
| Participant gender | 0.23 (0.07)*** | 0.14 (0.07)* | 0.21 (0.07)** |
| Target gender | -0.09 (0.04)* | -0.09 (0.04)* | -0.09 (0.04)* |
| Participant gender × target gender | 0.07 (0.05) | 0.08 (0.05) | 0.07 (0.05) |
| Masculinity | | -0.007 (0.002)** | |
| Expressivity | | | 0.07 (0.06) |
| Social support | | | |
| Intercept | 2.89 (0.08)*** | 2.85 (0.09)*** | 2.84 (0.10)*** |
| Participant gender | 0.08 (0.11) | 0.15 (0.12) | 0.07 (0.11) |

| Effect | Estimate (*SE*) | | |
|---|---|---|---|
| | Baseline model | Controlling for masculinity | Controlling for expressivity |
| Target gender | -0.01 (0.04) | -0.01 (0.04) | -0.01 (0.04) |
| Participant gender × target gender | -0.003 (0.05) | 0.003 (0.06) | -0.01 (0.05) |
| Masculinity | | 0.005 (0.004) | |
| Expressivity | | | 0.10 (0.11) |
| Self-help | | | |
| Intercept | 1.89 (0.07)*** | 1.84 (0.07)*** | 1.90 (0.09)*** |
| Participant gender | 0.14 (0.09) | 0.24 (0.10)* | 0.13 (0.10) |
| Target gender | -0.003 (0.04) | -0.003 (0.04) | -0.003 (0.04) |
| Participant gender × target gender | -0.003 (0.05) | 0.01 (0.05) | -0.004 (0.05) |
| Masculinity | | 0.010 (0.003)** | |
| Expressivity | | | -0.01 (0.09) |

*Note*. Intercepts captured the ratings of men in the male target condition. Participant gender contrasted men (baseline) vs. women, target gender the male target (baseline) vs. the female target. The measure of masculinity (CMNI-46 total scores) was centered before it was entered in the models. Expressivity contrasted clusters low (baseline) vs. high in expressivity. *$p<.05$, **$p<.01$, ***$p<.001$.