Getting shredded: Improving our understanding of muscle dysmorphia and related symptomology

Marita Cooper
July, 2018

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University.

Centre for Mental Health Research
Research School of Population Health
College of Medicine, Biology and the Environment
The Australian National University
Canberra, ACT, Australia

© Copyright by Marita Cooper 2018
All Rights Reserved
Declaration

I declare that this thesis is my own original work completed under the supervision of Emeritus Professor Kathleen Griffiths. Where additional authors were involved in this work, I declare that I have clearly stated my contribution. I affirm that this thesis has been completed in accordance with the Australian National University Research Award Guidelines for higher degree research.

Marita Cooper
July 2018
Author Contribution

All of the components of this thesis are based on the original work of the author, except where acknowledged as follows:

**Study 1: Muscle dysmorphia: Refining diagnostic criteria and improving nosological clarity.** The author designed the study in collaboration with JT, KE, DF, and KG. The author collected the articles which were then analysed collaboratively with BC. The author completed the data analysis and interpreted the results. The manuscript was prepared for submission by the author with revisions and final approval by all other co-authors.

**Study 2: Preliminary psychometric characteristics of a measure for appearance and performance-enhancing supplement and substance usage.** MC developed the APES items under the supervision of RH. The study design was developed by all three authors. MC executed the study and analyzed the data. MC wrote the first draft of the manuscript; RH and KG contributed through further commentary to the revision and approved the final manuscript.

**Study Three: Getting shredded: Development and validation of a measure of muscularity-oriented disordered eating.** The study design was developed by MC under the supervision of KG. MC developed the items, collected and analyzed the data. MC wrote the first draft of the manuscript; KG contributed to the revision of the paper and approved the final manuscript.

**Study Four: Bodybuilding to the max: Is our drive for muscularity based on body image dissatisfaction or distortion?** The study design was developed by MC under the supervision of KG. MC executed the study and analyzed the data. MC wrote the first draft of the manuscript; KG contributed through suggestions for analysis and critical editing of the manuscript and approved the final manuscript.

The journal articles presented in this thesis feature the author as the primary investigator in each instance. Other authors of these papers are listed below:


**Statement of Contribution**

This thesis is submitted as a Thesis by Compilation in accordance with [https://policies.anu.edu.au/ppl/document/ANUP_003405](https://policies.anu.edu.au/ppl/document/ANUP_003405)

I declare that the research presented in this Thesis represents original work that I carried out during my candidature at the Australian National University, except for contributions to multi-author papers incorporated in the Thesis where my contributions are specified in this Statement of Contribution.

**Title:** Muscle dysmorphia: Refining diagnostic criteria and improving nosological clarity

**Authors:** Marita Cooper, ANU; Dr Kamryn Eddy, Harvard University; Dr Jennifer Thomas; Harvard University; Dr Debra Franko, Northwestern University; Dr Kathleen Griffiths, ANU & Dr Bradley Carron-Arthur

**Publication outlet:** International Journal of Eating Disorders

**Current status of paper:** In preparation

**Contribution to paper:**

<table>
<thead>
<tr>
<th>Study concept and design</th>
<th>Analysis and interpretation</th>
<th>Manuscript preparation and revision</th>
<th>Total contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>75%</td>
<td>75%</td>
<td>80%</td>
<td>77%</td>
</tr>
</tbody>
</table>

Senior author or collaborating authors endorsement: _______________________________________________________

---

Marita Cooper

Candidate – Print Name

Signature

Date: 17/07/2018

---

Emeritus Professor Kathleen Griffiths

Primary Supervisor – Print Name

Signature

Date: 19/07/2018

---

Delegated Authority – Print Name

Signature

Date
Statement of Contribution

This thesis is submitted as a Thesis by Compilation in accordance with https://policies.anu.edu.au/ppl/document/ANUP_003405

I declare that the research presented in this Thesis represents original work that I carried out during my candidature at the Australian National University, except for contributions to multi-author papers incorporated in the Thesis where my contributions are specified in this Statement of Contribution.

Title: Preliminary psychometric characteristics of a measure for appearance and performance-enhancing supplement and substance usage.
Authors: Ms Marita Cooper, ANU; Dr Richard Hicks, Bond University; Dr Kathleen Griffiths, ANU
Publication outlet: Scandinavian Journal of Sports Medicine
Current status of paper: Under Review
Contribution to paper:

<table>
<thead>
<tr>
<th>Study concept and design</th>
<th>Analysis and interpretation</th>
<th>Manuscript preparation and revision</th>
<th>Total contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>75%</td>
<td>85%</td>
<td>83%</td>
</tr>
</tbody>
</table>

Senior author or collaborating authors endorsement: __________________________

Marita Cooper
Candidate – Print Name
Signature
Date

Emeritus Professor Kathleen Griffiths
Primary Supervisor – Print Name
Signature
Date

Delegated Authority – Print Name
Signature
Date
Statement of Contribution

This thesis is submitted as a Thesis by Compilation in accordance with https://policies.anu.edu.au/ppl/document/ANUP_003405.

I declare that the research presented in this Thesis represents original work that I carried out during my candidature at the Australian National University, except for contributions to multi-author papers incorporated in the Thesis where my contributions are specified in this Statement of Contribution.

Title: Getting shredded: Development and validation of a measure of muscularity-oriented disordered eating
Authors: Marita Cooper, ANU & Dr Kathleen Griffiths, ANU

Publication outlet: International Journal of Eating Disorders

Current status of paper: Under Review

Contribution to paper:

<table>
<thead>
<tr>
<th>Study concept and design</th>
<th>Analysis and interpretation</th>
<th>Manuscript preparation and revision</th>
<th>Total contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>80%</td>
<td>75%</td>
<td>85%</td>
<td>80%</td>
</tr>
</tbody>
</table>

Senior author or collaborating authors endorsement: __________________________

Marita Cooper 17/07/2018
Candidate – Print Name Signature Date

Endorsed

Emeritus Professor Kathleen Griffiths 19/07/2018
Primary Supervisor – Print Name Signature Date

Delegated Authority – Print Name Signature Date
Statement of Contribution

This thesis is submitted as a Thesis by Compilation in accordance with https://policies.anu.edu.au/ppl/document/ANUP_003405

I declare that the research presented in this Thesis represents original work that I carried out during my candidature at the Australian National University, except for contributions to multi-author papers incorporated in the Thesis where my contributions are specified in this Statement of Contribution.

Title: Bodybuilding to the max: Is our drive for muscularity based on body image dissatisfaction or distortion?
Authors: Marita Cooper, ANU & Dr Kathleen Griffiths, ANU
Publication outlet: Body Image

Current status of paper: Submitted

Contribution to paper:

<table>
<thead>
<tr>
<th>Study concept and design</th>
<th>Analysis and interpretation</th>
<th>Manuscript preparation and revision</th>
<th>Total contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>75%</td>
<td>85%</td>
<td>83%</td>
</tr>
</tbody>
</table>

Senior author or collaborating authors endorsement: __________________________________________________________

Marita Cooper
Candidate – Print Name
Signature
Date

Emeritus Professor Kathleen Griffiths
Primary Supervisor – Print Name
Signature
Date

Delegated Authority – Print Name
Signature
Date
Acknowledgements

I would firstly like to acknowledge my parents, Bob and Jean Cooper who have consistently proven to be the most supportive parents in existence (trust me on that, I’m almost a doctor). They have consistently been there for me with not only financial support but, most importantly, with unconditional love. They have been sounding boards for my problems/concerns and always lent caring words and support. Just this once I will allow you to have your win (note that there is no oxford comma before the ‘and’ in the previous sentence) as the ultimate thanks you for all for your support and love. You are the best.

I would also like to thank my brother, Rohan, and his partner, Amy, as well as my two adorable nieces, Sophie Jean and Evie Beth. It has been hard being so far away and missing so much time with you, especially whilst the girls are so young and changing every day. Hopefully this will all be worth it and I will be able to be there for you all soon.

Next, Anthony, whilst you will get a joint acknowledgement with Nick in a minute for a bunch of things, I do need to single you out specifically. We have bonded since the first time we spoke during our Masters program when you said to me that you liked how I ended my sentences; from then onwards I knew our friendship was destined to be epic. I know that you hate research, but you chose to at least tolerate it for me which I truly appreciate as I know how much you hate making concessions on your dislikes for anyone. Over the past years you have read countless of my articles, heard about my research successes as well as my epic research failures and devastations, and through it all only complained a moderate amount. Outside of the research specific stuff, you are my own personal resilience coach who calls me out on schema related self-talk and remind me that I am not truly defective.

And Nick. I am so glad that you and Anth found each other. You are a truly thoughtful, caring, and amazing human being. Unlike Anth, you take genuine pleasure in learning and increasing your knowledge, so it has been great to have your enthusiasm around my research. As a couple, you and Anthony have both been there for me to listen, to feed me, and to support my PhD related poverty (or Louboutin related, but who knows).

To Simone. My strong girl Bong Soon. I know that especially towards the end of my research you have been struggling with a lot on your own plate. But, in spite of that, you have consistently made time to feed me (wow a lot of my acknowledgements come down to those who provide me food), watch Korean dramas and kpop videos with me (NCT choreo wow) and drink lots of wine and eat lots of cheese with me (and food yet again). Thank you for being there listening with an eager face and for caring even when this research stuff sounds like it is probably in a different language to that which you speak.

And Paul. You are so good. Not just in the obvious way (as in that way you are amazing) but in the way that makes someone a genuinely good human. You are
thoughtful and sweet, and I am so lucky to have found you (honestly the amount of crazy coincidences that led to us meeting is insane). You are so utterly unselfish and know exactly how to support me without me saying a word. You calm me like no one else can. I miss you terribly and can’t wait for what the future holds. Saranghae.

Thank you to the absolutely incredible Thesis Boot Camp team and, in particular, Victoria Firth-Smith, Associate Professor Inger Mewburn, and Dr Thuy Do. I could not have completed this project without you. I didn’t really know what I would get out of the weekend (or if I could write without wine) but my writing has grown immeasurably under each of your guidance. Victoria your warmth is infectious and you are inspirationally caring and kind. Inger your tips for writing have transformed my writing, themes and rhemes have changed my life. And Thuy you have taught me to write beautifully without it sounding as obnoxiously verbose. Thank you all and to anyone thinking of attending please do!

To Liz and Lou, I know it has been a last minute whirlwind having you onboard, but I am so incredibly grateful for both of you joining the team and your great feedback. I don’t think anyone can underestimate the importance of having good supervision and I really appreciate the support and feedback you have provided. Especially with Lou you researching outside your area and Liz you meant to be enjoying time off, I could not have asked for a better panel.

Thank you to Jenny, Kamryn, and Debby, my first real mentors. I have been able to believe in myself so much more because each of you believed in me. Debby, I loved that you got my odd Australian terminology and emboldened me to continue to travel for my research. Kamryn, you are so incredibly kind and encouraged me to be confident in my abilities and knowledge. Jenny, your advice has been invaluable and your humility is inspiring. You all have such immense research skill and knowledge, however, I hope that if I gain one thing from each of you it is to become more curious, supportive, and humble.

Kathy. I don’t know how to thank you for every way in which you have supported me throughout the past four years. There is no way I would have gotten this far without you. You have agreed to my seemingly insane changes and last minute ideas. You have supported me through the lowest and highest stages of this process and been there when I had no one on my side. You have made me a better researcher, writer, and human. Thank you.
Abstract
Muscle dysmorphia (MD) is a syndrome characterised by a preoccupation with one’s perception of their inadequate muscularity and engagement in risky body change behaviours to enhance muscularity. Recent systematic reviews examining MD have reported insufficient empirical support for the diagnosis, a lack of reliability in current diagnostic criterion, and inadequate delineation of MD from existing diagnoses. Considering the importance of accurate conceptualization of psychiatric diagnoses for early screening and identification of symptomology as well as effective treatment and prevention, there is an urgent need to formulate an accurate clinical picture of MD. Consequently, this project sought to develop a clear clinical picture of MD to improve the assessment of MD and to extend our understanding of the classification of MD to inform future research and treatment protocols. This thesis comprised four original studies (one meta-analytic review and three empirical investigations) formulated to achieve these aims. The initial study was a systematic and meta-analytic review of research papers on MD utilising Robins and Guze’s (1970) criteria for validating psychiatric diagnoses. Our findings revealed that individuals with MD demonstrated more significant symptomology than controls indicating they likely fit a clinical diagnosis. Despite this, the symptomatic profiles, treatment response, and familial links of MD could not be meaningfully differentiated from traditional eating disorders (anorexia or bulimia nervosa) or body dysmorphic disorder. The subsequent original investigations included the development and preliminary psychometric evaluation of two novel measures, the Appearance and Performance Enhancing Supplement/Substance Scale (APES) and the Eating for Muscularity Scale (EMS) as well as a community-based
study of muscle enhancing behaviors. These papers addressed ergogenic substance use, eating pathology, and body image disturbance, which had been identified from the initial review and Study 1 as potentially delineating MD from existing diagnoses. Studies 3 and 4 described the development and validation of the APES, an 18-item measure of supplement and substance use, as well as the EMS, a 30-item measure of muscularity-oriented eating pathology. The APES demonstrated strong internal consistency, medium to strong concurrent validity, and good to excellent test-retest reliability. Similarly, the EMS demonstrated high internal consistency, moderate to strong concurrent validity with associated measures, and strong test-retest reliability. Study 4 comprised an examination of the role of body image disturbance in predicting MD behaviours. This study, undertaken in a community sample of 265 participants, found that body image dissatisfaction, and not body image distortion, was associated with increased engagement with MD-related behaviours. Overall, our findings supported MD as a clinically relevant condition; however, we were unable to demarcate MD from existing diagnoses. Future research should utilise the APES and EMS as well as examine body image distortion to differentiate MD and similar psychiatric conditions. It is hoped that exploration of these areas, particularly within clinical populations, will continue to be a focus in research leading to increased knowledge regarding the aetiology, course, prognosis, heritability, and prevalence of MD to build more effective screening and treatment guidelines.
Table of Contents

Declaration ........................................................................................................................................... 2
Author Contribution ........................................................................................................................... 3
Acknowledgements .......................................................................................................................... 8
Abstract .................................................................................................................................................. 10
CHAPTER 1 ........................................................................................................................................... 13
  1. Overview of this chapter ............................................................................................................... 13
  2. Muscle dysmorphia ....................................................................................................................... 14
     2.1 History of MD ............................................................................................................................ 14
     2.2 Prevalence, risk factors and the burden of muscle dysmorphia ........................................... 15
     2.3 Symptomatology and classification of muscle dysmorphia .................................................. 18
     2.5 Assessment and treatment of MD .............................................................................................. 30
     2.6 Statement of the problem ......................................................................................................... 32
  3. The current research ....................................................................................................................... 33
     3.1 The aims of this project ............................................................................................................. 33
     3.2 Method ..................................................................................................................................... 34
     3.3 Overview of the studies ............................................................................................................ 35
     3.4 Anticipated contributions and limitations of the research ..................................................... 37

CHAPTER 2: Study 1 ............................................................................................................................. 52
  Foreword .......................................................................................................................................... 52
  Muscle dysmorphia: Refining diagnostic criteria and improving nosological clarity ........................................... 55

CHAPTER 3: Study 2 ............................................................................................................................. 104
  Foreword .......................................................................................................................................... 104
  Preliminary psychometric characteristics of a measure for appearance and performance-enhancing supplement and substance usage ................................................................. 106

CHAPTER 4: Study 3 ............................................................................................................................. 129
  Foreword .......................................................................................................................................... 129
  Getting shredded: Development and validation of a measure of muscularity-oriented disordered eating .......................................................... 131

CHAPTER 5: Study 4 ............................................................................................................................. 162
  Foreword .......................................................................................................................................... 162
  Bodybuilding to the max: Is our pursuit of muscularity driven by body image dissatisfaction or distortion? .................................................................................................................. 164

CHAPTER 6 ........................................................................................................................................... 183
  Overview .......................................................................................................................................... 183
  6.1 Summary of overall project findings ......................................................................................... 183
  6.2 Clinical and research implications of findings ........................................................................... 186
     6.2.1 MD as a clinically relevant population ................................................................................. 186
     6.2.2 Delineation of MD from existing diagnoses ....................................................................... 190
     6.2.3 Assessment of MD ............................................................................................................... 192
  6.2 Methodological strengths and limitations of this project ......................................................... 194
     6.2.1 Strengths ............................................................................................................................... 194
     6.2.2 Limitations ............................................................................................................................ 197
  6.3 Future research directions ......................................................................................................... 199
  6.4 Conclusion .................................................................................................................................. 202
CHAPTER 1
Muscle dysmorphia (MD) is a recently described cluster of symptoms characterised by a preoccupation with perceived insufficient muscularity (American Psychiatric Association, 2013). Although the exact prevalence of MD is at present unknown, it has been suggested that MD is as common as thinness-oriented body image disorders such as anorexia and bulimia nervosa (Grieve, Truba, & Bowersox, 2009). This is particularly concerning given that MD is associated with a heightened risk of substance dependence, suicide, and comorbid mental health diagnoses (Pope, Pope, Menrad, Fay, Olivardia, & Phillips, 2005). Despite the public health significance of MD, the scientific community remains in disagreement over its classification and diagnosis (Chung, 2001; Foster, Shorter, & Griffiths, 2013; Murray & Baghurst, 2013).

This thesis seeks to improve the current evidence base for MD with the aim of improving our understanding of its aetiology and classification and, ultimately, its treatment. The work, which is a thesis by compilation, comprises three papers under review (Chapters 3 to 5) one paper in preparation (Chapter 2), preceded by an introductory context chapter (Chapter 1) and concludes with an overview of the findings (Chapter 6).

1. Overview of this chapter
This chapter provides an overall context for the current PhD project and its constituent studies. It includes an overview of MD including a critical review of the literature that has informed the project and its design and an outline of the rationale for and the papers included in this thesis. Further background specific to each study, including its link to the overarching goals of this thesis, is included in the foreword to each publication.
2. Muscle dysmorphia

2.1 History of MD

MD was first identified in a study examining anabolic steroid use in male bodybuilding populations (Pope, Katz, & Hudson, 1993). In the course of this study, Pope and colleagues (1993) identified a subset of their sample who described themselves as “small and puny,” despite their objectively large musculature. These individuals reported engaging in maladaptive behaviours to increase their muscularity and hide perceived inadequacies in their physique (Pope et al., 1993). One participant in this study described his symptomology as a “reverse anorexia,” similar to anorexia nervosa but with reversed body image disturbance and behaviours; a term which the authors subsequently adopted in their description of the phenomenology. Reverse anorexia was described as a preoccupation with insufficient muscularity and leanness that led individuals to engage in compulsive behaviours such as excessive exercise, use of anabolic-androgenic steroids, and adherence to a strict diet regimen (Olivardia, Pope, & Hudson, 2000).

In subsequent studies, Pope and his colleagues reconceptualised reverse anorexia as MD, a form of Body Dysmorphic Disorder (BDD), and developed tentative diagnostic criteria for the condition (see Box 1; Pope, Gruber, Choi, Olivardia, & Phillips, 1997). This reclassification was informed by clinical observations leading to the judgment that eating-related pathology was a secondary and superfluous symptom of the disorder (Pope et al., 1997). Despite disagreement regarding this reclassification (Baghurst et al., 2014; Murray & Baghurst, 2013), Pope and colleagues’ criteria have been widely adopted in both clinical and research settings (Sandgren & Lavalee, 2018).
2.2 Prevalence, risk factors and the burden of muscle dysmorphia

2.2.1 Prevalence

An accurate understanding of the population prevalence of mental disorders is important for public policy, research priorities, service planning, understanding mental health trends and reducing the stigma associated with mental illness (Insel & Fenton, 2005; Smink, van Hoeken, & Hoek, 2012). Consequently, well-designed prevalence research is critical to the progression of novel psychiatric diagnoses such as MD.

However, to our knowledge there are no community-based prevalence studies of MD. Although this may be attributed to the relatively recent interest in MD, there are multiple examples of other recently introduced diagnoses that have been the focus of considerable research in the epidemiological field including binge eating disorder (Cossrow et al., 2016), internet gaming disorder (Petry, Rehbein, Ko, & O’Brien, 2015), and hoarding disorder (Nordslettern et al., 2013). An alternative hypothesis for the absence of prevalence research into MD,
may be the relatively high cost associated with epidemiological research in the context of the low presumed prevalence of MD as well as the inconsistent measurement and diagnostic criteria used for MD.

In the absence of direct epidemiological studies, our current knowledge of MD prevalence is largely restricted to research into at-risk samples and by studies of BDD population prevalence. A recent systematic review examining BDD prevalence across both community and clinical settings estimated that BDD has a weighted prevalence of 1.9% in community adult populations (Veale, Gledhill, Christodoulou, & Hodsoll, 2016). Further, preliminary research has identified that between 10 and 30% of individuals with full or partial BDD exhibit muscularity concerns (Mollmann, Dietel, Hunger, & Buhlmann, 2017; Schneider, Turner, Mond, & Hudson, 2016). Although these data may provide some indication of the population prevalence of MD, it should be noted that these studies did not include a diagnostic measure of MD and the clinical populations sampled typically targeted individuals with non-muscularity related BDD such as those seen in rhinoplasty and dermatology clinics (Mollmann et al., 2017; Schneider et al., 2016; Veale et al., 2016). Research that has sampled from populations at risk for MD, such as weightlifters and fitness professionals, have yielded estimated prevalence that typically range between 10 and 25% (Campagna & Bowsher, 2016; Diehl & Baghurst, 2016; Pope & Katz, 1994). Although each of these estimates of prevalence contributes to our knowledge about MD prevalence, larger population-based studies utilising techniques that ensure representative samples, clear diagnostic criteria and psychometrically sound instruments are required to improve our understanding of MD.
2.2.2 Risk factors

The aetiology of MD has most commonly been conceptualized within a biopsychosocial framework (Grieve, 2007; Ricciardelli & McCabe, 2004). Biopsychosocial risk factors have broadly paralleled those reported for traditional eating disorders including neuroticism, low intrinsic self-worth, negative affect, perfectionism, and internalised appearance ideals (Cafri et al., 2005; Davis, Karvinen, & McCreary, 2005; Grieve, 2007). It has also been found that individuals with MD are more likely to report a history of eating pathology and weight/shape concerns (Cafri, Blevins, & Thompson, 2006; Mosley, 2008; Pope et al., 1993; Pope et al., 1997), with males exhibiting MD symptomology also commonly reporting stress related to their masculinity (Murray & Griffiths, 2013; Ung, Fones & Ang, 2000), and females with MD symptoms reporting a history of physical or sexual assault (Gruber & Pope, 1999; Pope et al., 1997).

2.2.3 Burden

Although the predisposing and precipitating mechanisms of MD remain somewhat unclear, more is known about the physical and psychological outcomes of those with MD. There is consistent research evidence that MD is associated with a range of deleterious consequences including significantly greater rates of lifetime mood, anxiety or eating disorders as well as concomitant internalising and pathological eating symptomology (Cafri et al., 2006; Olivardia et al., 2000). Behaviours associated with MD, such as exercise pathology, disordered eating symptomology, and APED use, are associated with adverse health outcomes including increased likelihood of renal failure, physical injury, sudden cardiac death (Cafri et al., 2005). Further impairment arises from the pathological engagement in these behaviours (such as the excessive time spent
exercising and anxiety in many social situations) leading to negative sequelae including the breakdown of romantic relationships, loss of employment, and overall poorer quality of life (Gruber & Pope, 1999; Pope et al., 1997). Finally, due to the increased likelihood of polysubstance abuse in MD populations, individuals are more likely to exhibit clinical aggression, mania, and/or psychosis (Hildebrandt et al., 2011; Pope & Katz, 1988).

2.3 Symptomatology and classification of muscle dysmorphia

MD has recently been officially recognised as a specifier of BDD by the American Psychiatric Association (2013), with its inclusion in the Diagnostic and Statistical Manual of Mental Disorders – 5th edition (DSM-5). The DSM-5 recommends that individuals with MD must meet the criteria for BDD, as outlined in Box 2, with the primary preoccupation related to their inadequate muscularity. Despite this official stance, researchers have questioned the classification of MD as a specifier of BDD, with a number arguing - based on empirical evidence - that it may be more accurately classified as an eating/body image disorder or an addiction (Chung, 2001; Foster et al., 2013; Murray et al., 2012). Furthermore, a recent systematic review of MD research conducted since the release of DSM-5 has revealed that MD researchers have continued to employ Pope and colleagues’ (Pope et al., 1997) criteria with no identified studies utilising the DSM-5 manual’s criteria (Sandgren & Lavalee, 2018).
Two recent reviews have utilised Blashfield and colleagues’ guidelines (1990) for evaluating psychiatric diagnoses (dos Santos Filho, Passarelli Tiricomm Stefano, Touyz, & Mederios Claudino, 2016; Nieuwoudt, Zhou, Coutts, & Booker, 2012) to investigate the adequacy of the current formal classification of MD. These guidelines are intended to objectively examine the adequacy of a psychiatric diagnosis based upon the existence of valid reliable criteria for a set of co-occurring symptoms that can be differentiated from existing diagnoses as well as sufficient amounts of empirical literature. Based on Blashfield’s guidelines (1990), both reviews as well as a later study by Nieuwoudt and colleagues (2016) have concluded that there is currently insufficient support for MD; that Pope’s diagnostic criteria lack reliability; and that there is inadequate empirical support for MD as a co-occurring syndrome that can be delineated from existing diagnoses.

In view of the strong critiques of both the Pope and American Psychiatric Association’s criteria for MD, the overview of MD in the following section employs a bottom-up approach. This approach examines MD components at the symptomatic level to identify key areas that will serve as the foci of this thesis. We will consider body image disturbance symptomology first (Section 2.3.1) followed by behavioural symptomology (Section 2.3.2).


1. The person has a preoccupation with one or more perceived defects in their physical appearance.
2. The preoccupation causes clinically significant distress or impairment in social, occupational, or other important areas of functioning.
3. Preoccupations trigger excessive and repetitive behaviors.
4. The primary focus is not related to body fat or weight as in an eating disorder

MD specifier: The preoccupation is focused on physique being too small or insufficiently muscular.
2.3.1 Body image disturbance and MD

MD has been conceptualised as a disorder of body image disturbance where individuals exhibit a chronic preoccupation with inadequate muscularity (Pope et al., 1997). Body image is a multifaceted construct comprising an individual’s cognitive, affective, and perceptual experience of one’s body (Cash & Pruzinsky, 2002). Disordered body image arises when cognitions and emotions associated with an individual’s body image perception and satisfaction begin to detrimentally affect their self-worth or body esteem, or result in clinical distress and/or dysfunctional behaviours (Barlett, Harris, Smith, & Bonds-Raacke, 2005).

Historically, research examining body image focused on the drive for thinness in females, however, the late 20th century saw a shift towards better understanding the body image aspects specific to males (Mishkind, Rodin, Silberstein, & Stiegel-Moore, 1986; Parent, Schwartz, & Bradstreet, 2016). Body image research in male populations identified more diverse body image concerns than were originally found in female populations, with males encouraged to achieve an increasingly elusive combination of high muscularity and leanness (Cafri & Thompson, 2004; Corson & Andersen, 2002). Research examining muscularity-oriented body image concerns has identified a phenomenon termed the ‘drive for muscularity’ (McCreary & Sasse, 2000). The drive for muscularity, a construct similar to the drive for thinness, has been described as a preoccupation with one’s muscularity and behavioural attempts to increase muscle mass (McCreary & Sasse). Although this drive does not solely occur in the presence of distorted body image or MD, it is supported as a risk factor for MD (Robert, Munroe-Chandler, & Gammage, 2009). A recent meta-analytic review of the drive for muscularity indicated that, whilst further
research is required, preliminary research supports the relationship between MD and the drive for muscularity (Edwards, Tod, & Molnar, 2014).

A further implicated component of body image in MD research is body image distortion, a disturbance in the perception of one's appearance (Rosen, 1992). Body image distortion has been conceptualised as a core symptom of MD. This suggestion is implicit in the description of MD as a specifier of BDD, with MD behaviours assumed to be largely driven by this distorted perception of muscularity (American Psychiatric Association, 2013). This distorted perception of musculature has been reported in both Pope's original paper (Pope et al., 1993) and in subsequent case studies (Mosley, 2008; Pope et al., 1997; Ung et al., 2000). These studies report that individuals with MD commonly describe themselves as small or puny despite their objectively large musculature (Hernandez-Martinez, Gonzalez-Marti, & Jordan, 2017; Pope et al., 1993; Pope et al., 1997). Despite this evidence, there is currently an absence of quantitative data examining the role of body image distortion in MD populations, with an over-reliance on anecdotal reports.

2.3.2 Behavioural symptomology of MD

Consistent with its current classification within the obsessive-compulsive and related disorders of the DSM-5 (American Psychiatric Association, 2013), MD comprises both preoccupations and/or obsessions with insufficient muscularity as well as compulsive or repetitive behaviours to modify or avoid exposing one's physique. Body modification behaviours for the purpose of enhancing muscularity have demonstrated associations with deleterious physical and psychological outcomes such as heightened depressive symptomology, poorer cardiac health, and decreased self-esteem (Cafri et al.,
2005). The following section will review core behavioural symptomology of MD including body image-related behaviours, exercise, use of appearance- and performance-enhancing supplements and substances, and disordered eating.

**Body image-related behaviours.** The compulsive and repetitive behaviours central to BDD include frequent mirror use or checking of the area of concern, attempts to hide or fix the area, and excessive reassurance seeking about the area of concern (American Psychiatric Association, 2013). Although no body-image-related behaviours are described in the DSM-5 criteria that are exclusive to the MD specifier, Pope’s criteria (Pope et al., 1997, p.556.) describes “attempts to hide the area of concern” in their avoidance of physique criterion.

Avoidance of physique exposure behaviours were initially identified in studies by Pope and colleagues (1993, 1997), with individuals describing mirror avoidance, obscuring of muscularity by wearing baggy, ill-fitting clothes, and avoidance of public change rooms and sexual intercourse (as this would require exposure of their perceived-to-be inadequate muscularity). These reports have been supported by further clinical accounts of MD (Mosley, 2008; Murray et al., 2011) as well as in researcher Timothy Baghurst’s first person account of his lived experience of MD (Baghurst, 2012). Despite case study support for avoidance of physique exposure as a criterion of MD, there has been mixed support for it within the quantitative research literature with suggestions that the criterion possesses limited reliability and is not specific to MD populations (Nieuwoudt et al., 2016). In a recent review of Pope’s criteria, Nieuwoudt and colleagues (2016) reported that the avoidance of exposure criterion demonstrated poor inter-rater reliability and was the least common of the criteria.
Although no other compulsive or repetitive behaviours are defined in the DSM-5 specifier for MD or Pope’s criteria (1997), clinical case studies and empirical research have described other body-image protection behaviours occurring in MD (Mosley, 2008; Murray et al., 2011; Murray & Griffiths, 2015). These have included reports of: both avoidance of and frequent mirror checking or use of scales; attempts to enhance muscularity (as will be highlighted in the upcoming sections); and reassurance seeking from others regarding muscularity. These seemingly contrasting behaviours of frequent checking and avoidance are consistent with Fairburn’s transdiagnostic model of eating disorders (Fairburn, Cooper & Shafran, 2003). As both checking and avoidance increase the preoccupation and emotional investment in body image, Fairburn (2003) argues that these activities contribute to increased investment in self-worth and overevaluation of weight/shape. There has been some preliminary support for the use of this model in predicting MD symptomatology (Murray, Rieger, Karlov & Touyz, 2013), however, further examination of these behaviours in clinical MD populations are required.

**Exercise.** Exercise behaviours are reported in almost all cases of MD; in fact, a description in the literature on MD that did not discuss exercise engagement was not found. Although exercise engagement is commonly encouraged as a positive health behaviour, individuals with MD have a relationship with exercise that is marked by rigid inflexibility, a resultant dysfunction in social and occupational domains, and clinical distress. While there is clear evidence for the presence of exercise symptomatology in MD, this is perhaps a consequence of methodological choices that have relied on MD research sampling from weightlifting, bodybuilding, and athlete populations.
Case reports of individuals with MD have described meticulously kept workout records, excessive time spent at gymnasiums, continued exercise in conflict with medical recommendations and despite injury, and high levels of distress if a workout is missed (Mosley, 2008; Murray et al., 2011; Pope et al., 1997; Gruber & Pope, 1999). This relationship with exercise has also been suggested to result in significant preoccupation, with individuals spending excessive time consuming fitness and health magazines as well as online material to learn new techniques and strategies to improve their gym routine (Murray, Rieger, & Touyz, 2011). Indeed, the maintenance of these dysfunctional exercise behaviours and their associated preoccupation has led to individuals leaving successful careers to become personal trainers (Gruber & Pope, 1999).

**Use of appearance and performance-enhancing drugs.** Although the use of anabolic androgenic steroids (AAS) and other appearance- and performance-enhancing supplements and substances (APED) originated in elite athletes, increasingly APED use has been adopted for the purpose of body modification (Millman & Ross, 2003). APED usage has been defined as the use of pharmacological substances for the purpose of modifying one’s appearance and/or to aid in dietary control, exercise, or sporting performance (Hildebrandt et al., 2011). Similar to the consumption of other illicit substances, the pathological use of illicit APED can lead to adverse psychiatric and medical complications including cardiomyopathy, cancer, comorbid psychiatric symptomology, and premature death (Kanayama, Pope, & Hudson, 2001). Conversely, the use of APED can also lead to desirable physical and psychological outcomes including euthymic mood, increased muscle mass, reduced adiposity, increased confidence, and improved athletic performance (Hildebrandt,
Langenbucher, Carr, & Sanjuan, 2007; Vassalo & Olrich, 2010). Furthermore, the level of lean muscularity gain that can be achieved through AAS use in particular is beyond that achievable through regular body change methods such as dieting and exercise (Rohman, 2009).

It was originally hypothesised that MD may be a precipitant and/or a maintaining factor in the use of AAS (Pope et al., 1993). Indeed, AAS usage is significantly higher in MD populations when compared with both individuals with non-muscularity related BDD and healthy populations (Gonzalez-Martí, Fernandez-Bustos, Contreas Jordan, & Sokolova, 2017; Pope et al., 2005). Although the use of AAS in MD is well established, there is a paucity of evidence examining the prevalence of other APED use in MD populations. One recent study that thoroughly examined APED and MD was Hildebrandt and colleagues’ (2011) investigation of illicit-APED users in the community. This study explored associations between MD symptomology and use of nutritional supplements, pro-hormones, steroids, and other ancillary APEDs. The study found a small but significant relationship between extent of APED use and overall MD symptomology as well as a moderate relationship with MD-related functional impairment. These results, as well as suggestions that body image disturbance is most severe in long-term APED users (Kanayama, Barry, Hudson & Pope, 2006), highlight the importance of exploring APED use in clinical MD populations.

Dietary and nutritional supplements (or licit APED as we will refer to them) are an alternative to illicit APED usage and are often posited as a more natural and safe option for achieving similar ergogenic benefits (Carvey, Farina, & Lieberman, 2012; Hartmann & Siegrist, 2006). Prevalence of licit APED use has been estimated to be as high as 90% among elite athletes and approximately
40% of the amateur athlete and gym-using population (Barkoukis, Lzarus, Lucidi, & Tsorbatzoudis, 2015; Burns, Shciller, Fada, Merrick & Wolf, 2004). Licit APED are commonly marketed as increasing the user’s speed, endurance, performance, and strength, as well as reducing recovery time and modifying body shape/size (Muller, Gorrow, & Schneider, 2009). The use of these licit APED, however, is frequently undertaken without the supervision of or consultation with medical or relevant health professionals (Kao, Deuster, Burnett & Stephens, 2012; Mooney et al., 2017). This absence of supervision and high prevalence of licit APED use is particularly concerning in the context of findings that these products typically lack support for their purported benefits, may be mislabeled, and can also contain illicit and harmful ingredients (Abbate et al., 2014; Kohler, Thomas, Petrou, Schanzaer, & Thevis, 2009; Nissen & Sharp, 2003). Furthermore, use of licit APED has been linked to deleterious physical consequences similar to those seen in illicit APED including rhabdomyolysis, cardiac syncope and arrhythmias, seizures, hospitalisation, and premature death (Flanagan, Kaesberg, Mitchell, Ferguson, & Haigney, 2008; Nordt et al., 2012; Stahl, Borlongan, Sterlip, & Sterlip, 2006) as well as greater levels of intention to use illicit APED (Hildebrandt, Harty, & Langenbucher, 2012).

The measurement of both licit and illicit APED usage in MD populations has been hindered by a lack of psychometrically supported assessments. Large-scale population studies have commonly relied upon individual questions of unknown psychometric validity and reliability that commonly fail to capture the breadth of APED usage (Ntoumanis, Ng, Barkoukis, & Backhouse, 2014; Piacentino et al., 2015; Sagoe et al., 2015). While there are a few examples of APED measurement based on interview schedules (such as that employed by
Hildebrandt and colleagues, 2011) or medical examination, such measures are
typically costly to administer and may be unsuitable for field investigations.
Thus, while the use of APED is thought to be a feature of MD that differentiates it
from other psychiatric diagnoses, there is at present no psychometrically valid
survey tool to assess APED usage in large samples.

**Eating pathology.** In their seminal article first identifying MD, Pope and
colleagues (1993) labeled the syndrome as “reverse anorexia.” Reverse anorexia
was described as akin to anorexia nervosa except with body image disturbance
and related behaviors that are reversed. The clearest example of ‘symptom
reversal’ is seen in the domain of eating pathology. Whereas individuals with
anorexia nervosa use dieting methods in an attempt to reduce their size/shape,
individuals with MD are more likely to use these behaviors in an attempt to
increase their size/shape.

Although eating pathology is typically viewed as a core symptom in
anorexia nervosa, the role of eating pathology in MD is much less clear.
Following their reconceptualisation of MD to a form of BDD, Pope and colleagues
(1997) reported that dieting behaviours were not a central nor necessary feature
of MD. This conceptual shift away from dieting and eating pathology was based
on the authors’ anecdotal research experience and has been further supported
by research emphasising that disordered exercise is more common than eating
pathology in individuals with muscularity-oriented body dissatisfaction
(Chittester & Hausenblas, 2009; Robert et al., 2009).

A significant limitation of previous investigations has been the selection
of assessment measures focused largely on the drive for thinness (Cafri &
Thomson, 2004; McCreary & Sasse, 2000). Eating attitudes and behaviours
motivated by the drive for thinness, and related concepts measured in these investigations, include fasting, skipping meals, low calorie/kilojoule diets, and purging (Grigg, Bowman, & Redman, 1996; Savage & Birch, 2010). While thinness-oriented disordered eating is one category of eating pathology, disordered eating practices for the purpose of appearance modification can also include behaviours to gain weight/muscularity (McCreary, Hildebrandt, Heinberg, Boroughs, & Thompson, 2007). Overuse of thinness-oriented disordered techniques would likely lead to a loss of musculature and undermine behaviours to gain weight/muscularity. As MD behaviours are driven by the desire to increase musculature, many thinness-oriented disordered eating behaviours are inconsistent with this goal.

Consequently, researchers examining MD have shifted their focus to muscularity-oriented disordered eating (MDE; Griffiths, Murray, & Touyz, 2013; Murray, Griffiths, & Mond, 2016). MDE is a cluster of pathological eating attitudes and behaviours motivated by the pursuit of muscularity, which may include both eating to increase muscularity as well as leanness (Griffiths et al., 2013). MDE behaviours may include increased caloric intake, bingeing, overconsumption of protein-based foods, and restriction of fats and carbohydrate-based foods (Griffiths et al., 2013; McCreary & Sasse, 2002; Murray, Griffiths, & Mond, 2016; Petrocelli, Oberweis, & Petrocelli, 2008). These behaviours have been observed in children as young as 8 years old (Eisenberg, Wall, & Neumark-Sztainer, 2012; Ricciardelli & McCabe, 2004) and are reported to be as common in males as thinness-oriented dieting behaviours are in females (Cafri, 2005; Nowell & Ricciardelli, 2008). Case reports of individuals with MD have described dieting practices with a range of negative sequelae including
impairment in social and occupational functioning, comorbid eating disorder symptomology, and heightened distress related to delayed eating (more than 2-3 hours) (Mosley, 2008; Murray et al., 2011; Murray et al., 2012).

Research examining the rates of MDE among MD populations has continued to be hindered by a lack of relevant, psychometrically sound measures. One noteworthy investigation by Murray and colleagues (2012) compared rates of eating pathology in a sample of males diagnosed with MD and a sample of males with anorexia nervosa (AN). This study incorporated both a standard measure of eating pathology (Eating Disorders Examination Questionnaire; Fairburn & Beglin, 2008) and a modified version of the measure to better reflect muscularity concerns. Traditional eating pathology, consistent with previous research, was significantly higher in individuals with AN than in those with MD; although, males with MD reported significantly higher traditional eating pathology than healthy controls (Murray et al., 2012). Notably, when the modified version of the eating pathology measure was employed, rates of MDE were similar across both clinical groups. Furthermore, as would be expected, the MD group reported higher drive for bulk than the AN group. While these results are preliminary, they support the suggestion that eating pathology may be a more important feature of MD than previously proposed on the basis of traditional measures of eating pathology. They also highlight the need for appropriate measurement in assessing the role of eating pathology in MD.
2.5 Assessment and treatment of MD

2.5.1 Assessment of MD.

Males with MD are less likely than males with traditional eating disorders to seek treatment for body image concerns (Leone, Sedory, & Gray, 2005). More commonly, these individuals may present to primary care and mental health settings for substance use, self-esteem difficulties, aggressive or violent behaviours, fatigue, mood or anxiety-related symptomology, and physical health consequences of APED use (Leone et al., 2005; Parent, 2013). Consequently, there is a need for accurate MD instruments for screening at-risk individuals presenting to clinicians with these conditions. The development of valid and reliable measurement tools is also key for our understanding of the aetiology, prevalence, and treatment outcomes for individuals with MD.

At present, a number of measures exist for the diagnosis and assessment of MD including the Muscle Appearance Satisfaction Scale (Mayville, Williamson, White, Netemeyer, & Drab, 2002), the Muscle Dysmorphia Symptom Questionnaire (Olivardia et al., 2000) and the Muscle Dysmorphic Inventory (Rhea, Lantz & Cornelius, 2004). However, concerns have been raised regarding the reliability and validity of these measures (Suffolk, Dovey, Goodwin, & Meyer, 2013). As with Pope’s initial MD criteria, these measures have predominantly been developed and validated in male weightlifting and bodybuilding samples, and, perhaps consequently, have demonstrated poorer psychometric properties in female populations. One exception, the Muscle Dysmorphic Inventory (Rhea et al., 2004), was validated in a mixed gender sample; nevertheless, the generalisability of a scale developed in a weightlifting sample to community and clinical populations must be questioned. There has been further concerns about
the use of these measures in assessing MD as many have not been investigated among clinical populations with known-group validity. Rather, studies of these scales have typically involved comparisons between frequent weightlifters or exercisers and less frequent weightlifters or exercisers (McCreary & Sasse, 2002; Rhea et al.).

2.5.2 Treatment of MD.

To date, there have been no interventions developed with the specific purpose of preventing or treating MD. In the absence of such protocols, it has been proposed that interventions commonly used for the treatment of BDD, OCD, and/or eating disorders may be indicated for the treatment of MD. Case studies have demonstrated promising results for the use of pharmacological interventions (Phillips, O'Sullivan, & Pope, 1997; Ung et al., 2000), cognitive therapy (Ung et al., 2000) and an adapted protocol for family-based therapy for anorexia nervosa with individuals with MD (Murray & Griffiths, 2013). Further recommendations have included the use of antidepressants including fluoxetine and sertraline as well as motivational interviewing (Grieve et al., 2009; Leone et al., 2005; Olivardia, 2007). Additionally, Grieve and colleagues (2009) recommended that treatment may include: psychoeducation, harm minimisation, behavioural strategies to limit reassurance seeking and mirror checking, mindfulness to encourage acceptance of current body shape, and living a valued life despite perceived inadequacies. However, none of these proposed treatments for MD have been subjected to efficacy testing in a clinical trial.

Two clear paths exist in establishing efficacious treatments for MD: (i) the development and evaluation of innovative treatment protocols specifically for individuals with MD; and (ii) the adaptation and evaluation of existing protocols
from similar diagnoses for MD populations. Both of these paths rely upon having a clear clinical picture for MD. While new protocols require this picture to identify appropriate treatment targets, the accurate classification of MD establishes its relatedness to similar psychiatric diagnoses, thus, supporting the extrapolation of existing protocols to MD populations. Evaluation of the efficacy of these treatments also requires the existence and/or development of suitable measures to assess treatment outcomes.

2.6 Statement of the problem

The development of new diagnostic categories can stimulate empirical, aetiological, and treatment studies in new branches of psychopathology. However, a haste to define new psychiatric conditions has the potential to lead to a proliferation of inadequately investigated diagnoses, over-pathologising of normal behavior, and redundancy in diagnosis (Blashfield, Sprock & Fuller, 1990). MD is a relatively new psychiatric diagnosis; however, it has already been linked to adverse psychological and physical sequelae including increased risk of suicide, substance dependence, comorbid depression, and sudden cardiac death (Baghurst & Kissinger, 2009; Cafri et al., 2005; Pope et al., 2005). Considering the importance of accurate conceptualisation of psychiatric diagnoses for early screening and identification of symptomology, as well as effective treatment and prevention, there is an urgent need to formulate an accurate clinical picture of MD.

This chapter has highlighted several key concerns with the processes utilised in the development of the existing Pope (1997) and DSM-5 criteria (American Psychiatric Association, 2013) for MD. Issues of primary concern
include the overuse of male weightlifting and bodybuilding samples as well as a lack of meaningful delineation from existing diagnoses. Although this delineation has been highlighted in previous reviews of MD criteria, there is also a need for further research examining the clinical significance of MD symptomatology. Finally, we have identified that research examining the clinical utility of MD has been hindered by a lack of psychometrically valid tools and an uncertain clinical picture.

3. The current research

3.1 The aims of this project

This project sought to develop a clear clinical picture of MD, including its classification, to improve the assessment of MD and to extend our understanding of the nature of MD to inform future research and treatment protocols. These objectives were pursued by addressing the following research questions:

RQ1: Is there currently sufficient evidence to differentiate MD from healthy individuals?

RQ 2: Is there currently sufficient evidence to differentiate MD from existing psychiatric diagnoses?

RQ 3: Can we accurately measure symptoms of MD and, where required, develop psychometrically sound instruments to measure these symptoms?

RQ 4: Can we identify further factors to differentiate MD from similar psychiatric diagnoses?

Four studies were formulated to answer these questions with each study presented in the form of a manuscript in separate chapters of this thesis as follows:
(i) Study 1 (Chapter 2): a systematic and meta-analytic review of MD using Robins and Guze’s method (1970) for validating psychiatric diagnoses. These guidelines were used to examine whether MD could be differentiated from healthy populations (RQ1) and existing diagnoses (RQ2);

(ii) Study 2 (Chapter 3): developed and evaluated the psychometric properties of an instrument measuring appearance- and performance-enhancing supplement and substance use, one of the core symptoms of MD (RQ3);

(iii) Study 3 (Chapter 4): developed and evaluated the psychometric properties of an instrument assessing muscle-ori- nted eating pathology, a symptom of MD (RQ3); and

(iv) Study 4 (Chapter 5): assessed the role of body image disturbance in MD as a potential factor to differentiate MD from existing psychiatric diagnoses (RQ4).

The concluding chapter (Chapter 6) of this project will provide an overall summary of the findings of this project, analyse how the results of each study contribute to furthering our understanding of MD, and discuss the implications of these findings within clinical and research contexts.

3.2 Method

The project involved a systematic review of the literature and three empirical studies employing community-based samples.

Our systematic and meta-analytic review compared groups of individuals exhibiting MD symptomology with healthy controls and individuals with similar
psychiatric diagnoses using Robins and Guze’s method (1970) for validating psychiatric diagnoses. By evaluating MD populations, not limited to evaluating a specific set of a priori diagnostic criteria, we proposed to assess the validity and reliability of MD symptomatology. This methodology allowed for inferences to be made about the validity and clinical distinctiveness of this clinical population as opposed to testing the accuracy of existing diagnostic criteria such as those by Pope and colleagues’ (1997), which has already been the subject of previous research (dos Santos Filho et al., 2016; Niewoudt et al., 2012).

Our empirical investigations were undertaken using the online administration of a battery of assessments in a community sample (Study 2, 3, 4) and a frequent exercising sample (Study 3). These samples were selected with the intention of recruiting samples with more equal gender representation and to reduce the limitations and bias from the over-reliance on weightlifting samples in previous MD research. Additionally, the measures in our battery were specifically developed or selected for their focus on muscularity-oriented body concerns and behaviours, with the aim of avoiding the inaccuracies in MD measurement associated with the use of thinness-oriented survey tools. Finally, our choice of online administration facilitated the collection of a large sample adequately powered to address our core research concerns.

3.3 Overview of the studies

Study 1 addresses the first two research questions (RQ 1 & RQ 2) regarding the validity and utility of MD as a psychiatric diagnosis. This involves a systematic review of research into MD utilising Robins and Guze’s (1970)
method and conducting a meta-analytic review of data where possible. This study proposed to answer the following questions:

1) Is MD a valid psychiatric diagnosis?

2) Are the current diagnostic criteria for MD, either those formulated by Pope et al. (1997) or the DSM-5 criteria (American Psychiatric Association, 2013), consistent with the clinical picture described in the literature?

3) Can individuals with MD be differentiated from healthy controls, selected populations (such as gymnasium users, bodybuilders, or athletes), and individuals with another psychiatric diagnosis (such as anorexia nervosa, steroid abuse disorder, or BDD)?

4) What do we know regarding the course, prognosis, and heritability of MD to date?

Following on from Study 1’s focus on the demarcation of MD from existing psychiatric conditions, Study 2 explores the use of appearance and performance enhancing substance and supplements, which has previously demonstrated potential for delineating MD and BDD. This study aimed to answer RQ 3 by seeking to develop a brief and psychometrically valid tool to assess the use of appearance and performance enhancing supplements and substances. Study 2 presents the methodology utilised to develop this measure and reports the psychometric properties of the measure.

As muscularity-oriented eating pathology has also been identified as a factor that may differentiate MD from existing diagnoses, Study 3 addresses RQ 3 by developing and evaluating the psychometric properties of an instrument designed to evaluate attitudes and behaviours related to muscularity-oriented
disordered eating. Study 3 details the multi-stage process used to develop this measure as well as the psychometric properties of the measure.

Finally, as the role of body image distortion in MD remains unclear, **Study 4** examines the role of body image in MD symptomology. This study sought to address RQ 4 by investigating further factors for delineating MD and existing diagnoses. Study 4 uses a cross-sectional design to examine the roles of body image dissatisfaction and distorted body image in predicting MD behaviours in a community sample.

### 3.4. Anticipated contributions and limitations of the research

The empirical investigations in Studies 2 to 4 in this project have been undertaken in community populations, albeit at-risk groups, with only Study 1 incorporating clinical samples. We believe this broad sampling method is an integral first step to improving our knowledge and understanding of MD. We do, however, acknowledge a priori that there are limitations to the inferences that can be drawn about clinical MD populations from such samples. Despite this, examination of MD in community settings is a key public health issue in its own right. As most individuals with MD do not directly seek treatment (Leone, Sedory, & Gray, 2005), it is important to undertake studies that will inform our knowledge of the symptomatology and demographics of this important target group to facilitate future identification and treatment of this population. Bearing in mind the limitations of the use of highly biased samples in previous research, we anticipate that the selection of an at-risk- and mixed-gender sample may improve the representativeness and generalisability of our results.
This project aims to make several key contributions to the clinical and research domains. Firstly, this is the first project to examine the current clinical significance of individuals experiencing MD, as opposed to assessing either the Pope (1997) or DSM-5 (American Psychiatric Association, 2013) criteria for MD. This allows for the exploration of different clinical pictures without the constraints of potentially flawed diagnostic criteria. This project will also contribute two novel measures of MD symptomology that assess APED use and muscul arity-oriented disordered eating, for which we argue that there are no currently available psychometrically sound measures. Finally, we will provide evidence regarding the role of body image disturbance as a factor that may delineate MD from existing psychiatric diagnoses.

It is anticipated that these contributions will provide clear guidance on the current nosological classification and future assessment of MD. This has important implications for both clinicians and researchers. Firstly, to the benefit of both groups, the psychometrically sound assessment tools developed within this study will support accurate assessment of MD symptomology as required to evaluate treatment outcomes, aetiological mechanisms, and prevalence rates. For clinicians, further diagnostic clarification will support clinical decision-making processes in the absence of current evidence-based interventions. Finally, for researchers, improved diagnostic clarification of MD as well as identification of factors to demarcate MD from similar psychiatric diagnoses will inform the development of future research. Through accurate nosological classification, it is hoped that we can improve research, treatment, and assessment outcomes for those who experience or are at-risk of MD.
References


Cossrow, N., Pawaskar, M., Witt, E. A., Ming, E. E., Victor, T. W., Herman, B. K.,


Suffolk, M. T., Dovey, T. M., Goodwin, H., & Meyer, C. (2013). Muscle dysmorphia:


CHAPTER 2: Study 1

Foreword

It has been suggested that the formation of new diagnostic categories encourages empirical research into burgeoning areas of psychopathology (Blashfield, Sprock, & Fuller, 1990). However, it has also been argued that the acceptance of new diagnostic categories without the support of sufficient existing literature and field trials can lead to poor reliability and validity of psychiatric diagnoses (Blashfield et al., 1990; Frances, 2009). A strong criticism of the DSM-5 (American Psychiatric Association, 2013) has been its inclusion of new psychiatric presentations that are not supported by a strong evidence base (Frances, 2009). The inclusion of muscle dysmorphia (MD) into the DSM-5 has been marked by controversy regarding its categorisation as a form of body dysmorphic disorder, conflict over specific criteria, and a concern that the MD diagnosis may simply represent the overpathologisation of the bodybuilding and weightlifting communities (Baghurst et al., 2014; Chung, 2001; Vanderecyken, 2011). Whilst two systematic reviews (dos Santos Filho et al., 2016; Nieuwoudt, Zhou, Coutts, & Booker, 2012) have utilised Blashfield and colleagues’ (1990) criteria for reviewing the current evidence for recognising psychiatric diagnoses, as yet there has been no systematic exploration of the validity and clinical utility of MD as a clinical diagnosis.

The aim of the review reported in this chapter was to address the following research questions of this project:

RQ1: Is there currently sufficient evidence to differentiate MD from healthy individuals?
RQ 2: Is there currently sufficient evidence to differentiate MD from existing psychiatric diagnoses?

To investigate these questions we undertook a systematic and meta-analytic review of research into MD using Robins and Guze’s (1970) method for the validation of psychiatric diagnoses. This method represented a shift from the previous DSM-II diagnostic classification system (American Psychiatric Association, 1968) that was informed by clinical judgment and experience to a focus on evidence, including data from follow-up and family studies (Feighner et al., 1972). The application of the Robins and Guze method was intended to expedite the development of clear psychiatric diagnoses and to permit improved treatment planning (Feighner et al., 1972) and has been applied to disorders such as Non-Fat-Phobic Anorexia Nervosa (Becker, Thomas, & Pike, 2009) and schizophrenia (Feighner et al.). The Robins and Guze approach involves assembling evidence relevant to each of the following five domains:

1) Clinical picture of disorder;

2) Laboratory studies;

3) Delineation from other disorders;

4) Follow-up studies; and

5) Family studies

As noted earlier, previous studies have utilised Blashfield and colleagues’ criteria for examining MD, which assess the validity and reliability of an a priori set of diagnostic criteria. However, a strength of Robins and Guze’s method is their openness to redefining a clinical picture based upon the current state of evidence. As such, this paper aims to undertake a qualitative and quantitative review of the existing research into MD using Robins and Guze’s approach to test
the validity of the disorder as a psychiatric diagnosis. In addition to addressing RQ 1 and RQ 2, we sought to provide an evidence-based clinical picture of MD, as well as to identify gaps in the literature that could inform our subsequent investigations.

The review identified two key factors to guide these investigations:

1) Individuals with MD likely represent a clinically relevant sample with more significant mental ill-health, APED use, eating and exercise pathology, impairment, and obsessive-compulsive symptomology than healthy controls and frequent exercisers. They were also found to report significantly lower wellbeing.

2) There is presently insufficient evidence to delineate MD from similar existing diagnoses including anorexia nervosa, bulimia nervosa, and body dysmorphic disorder; however, eating pathology and body change behaviours may be potential factors for future investigation.
Muscle dysmorphia: Refining diagnostic criteria and improving nosological clarity
By Marita Cooper, Kamryn Eddy, Jennifer Thomas, Debra Franko, Kathleen M. Griffiths, & Bradley Carron-Arthur

Abstract

Objective: Muscle dysmorphia (MD) is a new and controversial addition to DSM-5 as a specifier of body dysmorphic disorder (BDD). Research has critiqued the reliability and validity of MD as a clinically relevant condition. The current study aimed to assess the validity of the diagnosis of MD based on a systematic and meta-analytic review of the MD literature using Robins and Guze’s 5-phase approach. Method: We performed a systematic search of the ProQuest, PsycInfo and PubMed databases for the period prior to July 2017 using search terms relevant to MD. This resulted in a final set of 34 relevant papers. Data relevant to Robins and Guze’s five phases (clinical picture of disorder; laboratory studies; delineation from other disorders; follow-up studies; family studies) were analyzed using quantitative meta-analyses, cluster analysis, and qualitative syntheses. Results: Cluster analysis revealed two distinct symptomatic presentations of MD. Quantitative meta-analyses differentiated MD populations from controls; however, results were inconclusive in delineating MD from existing disorders. The symptomatic profiles, treatment response, and familial links for MD were similar to those for eating disorders and BDD. Discussion: This is the first study to review MD using Robins and Guze’s (1970) method. Whilst individuals with MD reported higher psychopathology than healthy comparisons, there was insufficient evidence that MD can be meaningfully differentiated from BDD or traditional eating disorders. Additional studies are
required to further determine the course, prognosis and heritability of this disorder to advance research and clinical care for this patient group.
1. Introduction

Muscle dysmorphia (MD) was recently included as a specifier of Body Dysmorphic Disorder (BDD) in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013). Since this addition, there has been disagreement over the classification of the disorder (Chung, 2001; Foster, Shorter, & Griffiths, 2013), its proposed diagnostic criteria (Baghurst, et al., 2014; Murray & Baghurst, 2013), and whether MD is a condition manufactured by the popular media (Vanderecyken, 2011). Perhaps reflecting these concerns, a recent meta-analysis found that no published research conducted on MD since the release of the DSM-5 has utilized the manual’s criteria (Sandgren & Lavellee, 2018). The limited researcher and evidentiary support for its current classification, raises the question of the reliability and validity of MD as a clinically relevant, distinct, and reliable condition.

Pope and colleagues (Pope, Katz, & Hudson, 1993) first identified MD in their study of affective and psychotic symptomology in anabolic-androgenic steroid users. Within this sample, Pope and colleagues (1993) discovered a subset of bodybuilders who were preoccupied with thoughts of their musculature and engaged in dysfunctional behaviors to both increase their musculature and hide their perceived inadequate physique. The condition was, at the time, termed “reverse anorexia,” since the observed body image disturbances and related behaviors resembled, but were the reverse of, those seen in anorexia nervosa. While individuals with anorexia nervosa are typically characterized by a preoccupation or fear of weight gain and exhibit restrictive dieting or compensatory behaviors, individuals with reverse anorexia were observed to

NB: This study is currently under review. Please refer to the final published material when citing.
have a preoccupation of being insufficiently muscular and lean, and exhibit
muscle-building behaviors such as excessive exercise, using anabolic-androgenic
steroids, and adhering to a strict diet regimen (Olivardia, Pope, & Hudson, 2000).

Subsequent to Pope and colleagues’ (1993) seminal article, “reverse
anorexia” was categorized as a form of BDD, with the tentative diagnostic criteria
detailed in Box 1 (Pope, Gruber, Choi, Olivardia, & Phillips, 1997). This shift of
MD to a subtype of BDD was informed by observations that eating-related
pathology was a secondary and superfluous symptom of the disorder (Pope et
al., 1997). MD has since been officially recognized by the American Psychiatric
Association (2013) with its inclusion in the DSM-5 for individuals meeting
criteria for BDD who are preoccupied with concerns that their physique is too
small or insufficiently muscular.

<table>
<thead>
<tr>
<th>1. The person has a preoccupation with the idea that one’s body is not sufficiently lean and muscular. Characteristic associated behaviors include long hours of lifting weights and excessive attention to diet.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. The preoccupation causes clinically significant distress or impairment in social, occupational, or other important areas of functioning, as demonstrated by at least two of the four following criteria:</td>
</tr>
<tr>
<td>e. The individual frequently gives up important social, occupational, or recreational activities because of a compulsive need to maintain his or her workout and diet schedule;</td>
</tr>
<tr>
<td>f. The individual avoids situations where his or her body is exposed to others, or endures such situations only with marked distress or intense anxiety;</td>
</tr>
<tr>
<td>g. The preoccupation about the inadequacy of body size or musculature causes significant distress or impairment in social, occupational, or other important areas of functioning; and</td>
</tr>
<tr>
<td>h. The individual continues to work out, diet or use ergogenic substances despite knowledge of adverse physical or psychological consequences.</td>
</tr>
<tr>
<td>3. The primary focus of the preoccupation and behaviors is on being too small or inadequately muscular, as distinguished from fear of being fat, as in anorexia nervosa, or a primary preoccupation only with other aspects of appearance, as in other forms of BDD.</td>
</tr>
</tbody>
</table>

Box 1. Pope and colleagues’ criteria for MD. From Pope, Gruber, Choi, Olivardia, & Phillips, 1997, p.556

Despite this official recognition of MD, researchers have continued to
question both the shift in conceptualization of MD as well as its current
diagnostic criteria (Sandgren & Lavallee, 2018). Supporting these concerns has

NB: This study is currently under review. Please refer to the final published material when citing.
been research demonstrating that MD has etiological similarities, gender disparity, age of onset, and treatment response more consistent with typical eating disorder diagnoses than with BDD (Lamanna, Grieve, Pitt Derryberry, Hakmann & McClure, 2010; Murray et al., 2012). Furthermore, it has been suggested that the low reported prevalence of eating pathology in MD populations may be explained by the use of assessments that focused solely on the drive for thinness, rather than the drive for bulk that characterizes MD (Murray & Touyz, 2016).

Researchers have not only raised concern over the nosological classification of MD, but have also questioned the clinical validity and reliability of this diagnosis. Two recent reviews, one prior to the DSM-5 release and one since, have concluded that there is currently insufficient support for MD as a distinct diagnosis (dos Santos Filho, Passarelli Tirico, Stefano, Touyz, & Mederios Claudino, 2016; Nieuwoudt, Zhou, Coutts, & Booker, 2012). Each of these reviews examined Pope and colleague’s (1997) criteria using Blashfield, Sprock, and Fuller’s (1990) guidelines for assessing if the extant empirical evidence supports the addition of a condition to the psychiatric classification system. They concluded that research on MD to date has showed inconsistent inter-rater reliability (kappa=.39-.83) for MD, inadequate support for symptoms as a co-occurring syndrome, and insufficient evidence that MD is a distinct clinical entity (dos Santos Filho et al., 2016; Nieuwoudt et al., 2012; Nieuwoudt et al., 2016). Whilst some of these criticisms are related to the paucity of evidence to date, they are also consistent with previous suggestions that Pope’s (1997) criteria are overly subjective, with terms such as “preoccupation,” and “excessive,” open to

NB: This study is currently under review. Please refer to the final published material when citing.
diverse clinical interpretations (Nieuwoudt et al.). It should be noted that as these studies were conducted with a focus on Pope’s initial criteria for MD, there has not yet been an evaluation of MD as a construct or of the DSM-5 criteria for MD.

The development of new diagnostic categories can stimulate empirical, etiological and treatment studies in new branches of psychopathology. However, a haste to define new psychiatric conditions has the potential to lead to a proliferation of inadequately investigated diagnoses and over-pathologizing of normal behavior. With a lack of support for the DSM-5 criteria (Sandgren & Lavallee, 2018) and the description by Pope and colleagues (1997) of their own diagnostic criteria for MD as “tentative” based on “preliminary observations,” the clinical validity of MD requires urgent attention.

Robins and Guze’s (1970) 5-phase method has been widely used to determine diagnostic validity in the extant psychiatric literature. The application of this system was intended to expedite the development of clear psychiatric diagnoses and to permit improved treatment planning (Feighner et al., 1972). The method involves a review of five domains of the research literature as follows: (1) clinical picture of disorder; (2) laboratory studies; (3) delineation from other disorders; (4) follow-up studies; and (5) family studies. Robins and Guze (1970) assert that this systematic method for studying psychiatric diagnoses is essential to establishing the diagnostic validity of a condition. While other approaches may assess the validity and reliability of an a priori set of diagnostic criteria, a strength of Robins and Guze’s guideline is their openness to redefining a clinical picture based upon the current state of evidence.

*NB: This study is currently under review. Please refer to the final published material when citing.*
Accordingly, the current paper sought to undertake both a qualitative and quantitative review of the existing research into MD using Robins and Guze’s method for testing the validity of the disorder as a psychiatric diagnosis.

2. Method

2.1. Identification of studies and study selection

We followed the PRISMA guidelines (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009) for reporting this meta-analysis and the protocol was registered with PROSPERO (Protocol no: 42017070127). We conducted comprehensive systematic searches, during the period June to July 2017 for articles published prior to these dates, of the following electronic databases: ProQuest, PsycInfo and Pubmed. Search terms included: “muscle dysmorphia,” “muscle dysmorphic,” “reverse anorexia,” and bigorexia. A supplementary examination of reference sections for additional papers was also conducted.

Inclusion criteria were as follows:

1) Studies must be from a peer-reviewed journal or dissertation/thesis;
2) Studies must be written in English; and
3) Studies must include either a participant/group of participants identified as experiencing heightened symptoms of or diagnosed with muscle dysmorphia, muscle dysmorphic disorder, reverse anorexia or bigorexia.

Studies that did not include empirical data, with the exception of case studies, were excluded. Two authors (MC and BC) initially screened all titles and publication types to remove articles that could immediately be excluded. Further review of abstracts and full text manuscripts was then conducted independently.
to determine eligibility for inclusion. Any differing opinions between the raters were discussed to reach a consensus decision.

2.2. Data extraction and management

The first author extracted and recorded all data for Robin and Guze’s 5-phase method (P1=clinical picture of disorder; P2=laboratory studies; P3=delineation from other disorders; P4=follow-up studies; and P5=family studies) in an Excel spreadsheet. Data for studies relevant to P1 were then transferred into IBM SPSS Statistics version 20 (IBM Corporation, 2011), whilst studies related to P2 and P3 were transferred into OpenMeta[Analyst] (Wallace et al., 2012).

2.3. Study variables

We extracted the following information from each study: study design; type and size of total sample; clinical group and comparison group; MD prevalence; demographic status and physical measurements (e.g. reported height, weight, body fat percentage, and fat-free mass index) of the clinical group; variables related to MD, including age of onset, precipitant, antecedent, assessment tool utilized, comorbidity, and treatment; symptoms described by MD participants; and outcome measure scores. In addition, means and standard deviations of the following measures were extracted from the laboratory studies: wellbeing, mental ill-health, APED use, eating pathology, exercise pathology, impairment, body dissatisfaction, and obsessive-compulsive pathology. Where studies utilized multiple measures of a variable, for example depression or anxiety for mental ill-health, the means of each individual variable (where
appropriate) or effect size were averaged prior to calculating mean difference scores and pooled standard deviations were calculated.

2.4. Statistical analyses

The methodological quality of the studies included in this review was assessed using the National Mental Health Research Council of Australia (NHMRC; 2009) guidelines. We assessed risk of bias using the framework of the Cochrane Collaboration’s tool for assessing risk of bias (Cochrane, 2011). We investigated P1 by examining both the clinical picture data (terms related to symptomatology extracted from each case study, e.g. muscularity, exercise, dissatisfaction) as well as a qualitative review of etiological mechanisms. The clinical picture data was further examined via a two-step cluster analysis. Data for P2 and P3 were subjected to quantitative meta-analysis using a random effects model comparing standardized mean differences between MD and comparison groups. Q and I² statistics were computed to evaluate heterogeneity between studies, with the option of conducting a fixed-effects model in the instance of homogenous studies and a random-effects model if heterogeneity was found. We had planned to employ funnel plots to evaluate risk of publication bias; however, the use of funnel plots is not recommended when the number of included studies is less than 10 and in the presence of heterogeneity between studies (Lau, Ioannidis, Terrin, Schmid, & Olkin, 2006). Separate meta-analyses were undertaken for each of the variables of wellbeing, mental ill-health, APED use, eating pathology, exercise pathology, impairment, body dissatisfaction, muscle dissatisfaction, body change behaviors, and obsessive-compulsive

NB: This study is currently under review. Please refer to the final published material when citing.
pathology. Due to the small number of papers found for P4 and P5, results from these studies were reviewed qualitatively and will be presented individually.

3. Results

The PRISMA flow diagram depicting the identification and inclusion of studies is represented in Figure 1. Database searches and other sources yielded a total of 655 titles. Following a title review, 210 full text papers were retrieved for further examination. A total of 34 papers involving 31 distinct samples satisfied the inclusion criteria and were included in the final review of which 17 papers (laboratory data and delineation) were included in the quantitative meta-analysis (see Table 1).

<INSERT TABLE 1 HERE>

3.1. Description of studies included

Using Robins and Guze's (1970) method, 24 studies contained data relevant to the clinical picture of MD (P1), 15 were laboratory studies (P2), five studies related to delineation from other disorders (P3), five involved follow-up studies (P4), and two articles were a family studies (P5).

Study characteristics. Of the 34 papers included in the overall review, 20 were observational, and 10 were case reports. Thirty of the 34 studies included in this review were ranked as the lowest level of evidence (Grade IV; NHMRC, 2008). Although 22 studies (65% of total) described their MD sample as “diagnosed,” the method of establishing MD diagnosis ranged from the use of pilot screening items (18%), unstructured or structured clinical interviews (60%), musculature-related measures (14%), BDD measures assessing area of concern (1%), and latent class analysis (1%).
Populations. The studies included in this review typically sampled participants from selected populations. These were most commonly weightlifting/bodybuilding groups, followed by steroid-using and frequent-exercising populations; with the remaining studies being from community populations and five of the case reports utilizing a clinically presenting sample. Studies specifically examining MD were predominantly undertaken in male populations (70%), with eight studies incorporating mixed gender populations and only one article focused specifically on female participants (Gruber & Pope, 1999). Overall, the MD groups typically comprised Caucasian, heterosexual male participants, aged approximately 26 years old, currently engaging in frequent weightlifting/bodybuilding (4 times/week or more), and whom commonly reported lifetime psychiatric comorbidity.

3.2. Risk of bias

We assessed risk of bias at both the individual study and overall level, identifying key areas of bias at both levels. Firstly, there was a high level of comorbidity within the MD samples, in particular steroid abuse which is known to lead to a range of deleterious physical and psychological sequelae (Kanayama, Hudson, & Pope, 2008). Despite this, the potential confounding effects of comorbid disorders, substance usage or the use of frequent exercising samples was rarely accounted for. Additionally, the risk of bias may be greater among studies that did not utilize psychometrically sound diagnostic measures. Due to the small number of studies identified, we were unable to address these concerns as moderator variables. It should also be noted that multiple studies were excluded from analysis due to incomplete reporting of outcome data for the
MD group in some published articles. We attempted to overcome this bias by requesting data from the original authors; however, such data were not always made available to us. Finally, we acknowledge the potential bias in our pragmatic decision to exclude articles not written in English.

Figure 1. PRISMA flow diagram of selected articles

NB: This study is currently under review. Please refer to the final published material when citing.
3.2. Clinical picture (Robins and Guze – Phase 1)

P1 of Robins and Guze’s (1970) method requires a description of both the cluster of clinical symptoms associated with the disorder as well as demographic and etiological factors related to the disorder (Feighner et al., 1972). Fifteen cases were included in this analysis from 10 individual studies.

Clinical symptoms. A two-step cluster analysis yielded a two-cluster model (see Box 1). The model demonstrated fair quality (Silhouette measure = 0.3) and an acceptable cluster ratio of 1.5. We named these clinical pictures the “purging disorder presentation,” where the picture focused on maladaptive behaviors to change one’s physique, and the “body image presentation,” where individuals experience a preoccupation and dissatisfaction with their muscularity and leanness resulting in dysfunctional body change behaviors, mental acts, and checking behaviors.
Cluster 1 – “Purging disorder presentation”

A. Recurrent inappropriate behaviors to increase shape/size including:
   1. Use of appearance/performance enhancing substances.
   2. Unhealthy dieting practices.
   3. Dysfunctional exercise behavior, commonly weightlifting.
B. Behaviors cause impairment in social, occupational or other important areas of functioning or lead to negative physical or mental health outcomes.

Cluster 2 – “Body image presentation”

A. Preoccupation and dissatisfaction with the inadequacy of their musculature, leanness, and shape/size.
B. Individual performs repetitive behaviors (e.g. mirror checking, weighing, reassurance seeking), mental acts (e.g. comparing his or her appearance with that of others), or engages in the following behaviors:
   1. Use of appearance/performance enhancing substances and supplements.
   2. Avoidance of situations where the individual exposes their body in front of others.
   3. Engaging in exercise for the purpose of increasing musculature, leanness, and shape/size.
   4. Rigidly adhering to dietary rules for the purpose of increasing musculature, leanness, and shape/size.
C. Behaviors cause impairment in social, occupational or other important areas of functioning, require excessive amounts of the individual’s time, or lead to negative physical or mental health outcomes.

Box 2. Cluster analysis findings for clinical pictures of MD

Etiology. This review yielded preliminary information about the development and onset of MD. Average age of MD onset was reported to be between 18 and 20 years (Cafri, Olivardia, & Thompson, 2008; Murray et al, 2012; Olivardia et al., 2000). Multiple studies reported a history of psychiatric diagnoses (Cafri, Blevins, & Thompson, 2006; Mosley, 2008; Phillips, O’Sullivan, & Pope, 1997; Pope, Katz, Hudson, 1993; Pope et al., 1997; Ung, Fones, & Ang, 2000), with common lifetime diagnoses including eating disorders (anorexia and bulimia nervosa), mood disorders (major depressive and manic/hypomanic episodes), anxiety disorders, and substance abuse and dependence; although, there were

NB: This study is currently under review. Please refer to the final published material when citing.
also studies that reported individuals did not exhibit prior mental health concerns (Gruber & Pope, 1999; Murray, Maguire, Russell, & Touyz, 2011a; Murray, Reiger, & Touyz, 2011b; Pope et al., 1997). Other important etiological mechanisms identified were a history of feeling weak and/or emasculated (Murray & Griffiths, 2013; Ung et al), restrictive dieting (Mosley, 2008; Murray et al, 2011a; Murray et al, 2011b; Olivardia et al., 2000), and weight/shape concerns related to being either much smaller or larger than peers (Cafri et al., 2006; Mosley, 2008; Murray et al., 2011a; Murray et al., 2011b; Murray & Griffiths, 2015; Pope et al, 1993; Pope et al., 1997; Ung et al). Direct precipitants of the onset of MD symptomology included use of anabolic-androgenic steroids (Pope et al., 1993; Pope et al., 1997) as well as experiences of physical or sexual assault (Gruber & Pope, 1999; Pope et al., 1993; Pope et al., 1997).

3.3 Laboratory studies (Robins and Guze – Phase 2)

The MD group was compared with both bodybuilding and community groups using a quantitative meta-analysis that yielded pooled standardized mean effects (see Table 2). Examination of Q and I² statistics indicated significant heterogeneity among studies on all summary measures. This heterogeneity would be expected due to the differences in both the criteria utilized for the MD group (some studies comprising clinically diagnosed individuals and others high-risk university students) as well as the diversity in healthy control groups (community individuals as well as amateur bodybuilders). While there are differing views on addressing heterogeneity in meta-analyses, we managed this by employing a random-effects model and exercising caution in the interpretation of results. The MD group demonstrated consistently higher scores.
on all measures of psychopathology, symptomology, and impairment, and indicated fewer wellbeing factors than comparison groups.

Table 2.
Results of continuous random-effects analyses for laboratory studies

<table>
<thead>
<tr>
<th>Variable</th>
<th>Included studies</th>
<th>SMD estimate</th>
<th>95% CI, lower</th>
<th>95% CI, upper</th>
<th>Std. err</th>
</tr>
</thead>
<tbody>
<tr>
<td>APED use</td>
<td>9</td>
<td>-2.81*</td>
<td>-4.81</td>
<td>-0.81</td>
<td>1.02</td>
</tr>
<tr>
<td>Body dissatisfaction</td>
<td>6</td>
<td>-1.82**</td>
<td>-2.42</td>
<td>-1.21</td>
<td>0.31</td>
</tr>
<tr>
<td>Eating pathology</td>
<td>8</td>
<td>-1.72**</td>
<td>-2.65</td>
<td>-0.80</td>
<td>0.47</td>
</tr>
<tr>
<td>Exercise pathology</td>
<td>6</td>
<td>-2.78*</td>
<td>-4.89</td>
<td>-0.66</td>
<td>1.07</td>
</tr>
<tr>
<td>Impairment</td>
<td>3</td>
<td>-3.23**</td>
<td>-4.79</td>
<td>-1.67</td>
<td>0.80</td>
</tr>
<tr>
<td>Mental ill-health</td>
<td>8</td>
<td>-0.84**</td>
<td>-1.23</td>
<td>-0.45</td>
<td>0.20</td>
</tr>
<tr>
<td>Obsessive-compulsive pathology</td>
<td>7</td>
<td>-1.24*</td>
<td>-2.12</td>
<td>-0.36</td>
<td>0.45</td>
</tr>
<tr>
<td>Wellbeing</td>
<td>5</td>
<td>0.82**</td>
<td>0.40</td>
<td>1.24</td>
<td>0.21</td>
</tr>
</tbody>
</table>

*p<.01, **p<.001

NB: Positive SMD estimates indicate higher scores in the healthy population group and negative SMD estimates indicate higher scores in the MD group.

3.4. Delineation from other disorders (Robins and Guze – Phase 3)

Five studies were identified that related to Robins and Guze’s (1970) third phase, delineation from other disorders. Three of these studies related to differentiating MD and non-muscularity related BDD (Nieuwoudt, Zhou, Coutts, & Booker, 2015; Pope et al., 2005; Schneider, Mond, Turner, & Hudson, 2017) with a further three studies examining differences between MD and eating disorders (Compte, Sepulveda, & Torrente, 2015; Murray et al., 2012; Niewoudt et al., 2015). No studies were found to delineate MD from other existing diagnoses such as obsessive-compulsive disorder.

MD vs. Eating Disorders

Three studies compared individuals with eating disorders and those with MD (Compte et al., 2015; Murray et al., 2012; Niewoudt et al., 2015). Murray
and colleagues (2012) investigated differences between males diagnosed with MD and males diagnosed with anorexia; while the comparison group in the two more recent papers comprised individuals diagnosed with EDNOS (sub-threshold anorexia nervosa and bulimia nervosa; Compte et al.) and those labeled as ‘having an eating disorder,’ identified by an elevated EAT-26 scores (Niewoudt et al., 2015).

Results from our meta-analytic comparison are presented in Table 3. As there was evidence of heterogeneity among studies on all summary measures - eating pathology ($I^2=90.2$, $Q = 20.42$), exercise ($I^2=96.87$, $Q = 63.83$), and muscle dissatisfaction ($I^2=65.59$, $Q = 5.81$) – these results should be interpreted with caution. Similar to P2 this heterogeneity is likely due to the diverse clinical comparison groups, which varied across studies from individuals engaged in ED treatment for anorexia nervosa (Murray et al., 2013) to university students with sub-threshold diagnoses (Compte et al., 2005) and bodybuilders (Niewoudt et al., 2015). The MD group reported significantly lower levels of traditional eating pathology than the ED groups. No differences were found between the ED and MD groups for either exercise pathology or muscle dissatisfaction.

Table 3. Results of continuous random-effects analyses for delineation from ED

<table>
<thead>
<tr>
<th>Variable</th>
<th>Included studies</th>
<th>SMD estimate</th>
<th>95% CI, lower</th>
<th>95% CI, upper</th>
<th>Std. err</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional eating pathology</td>
<td>3</td>
<td>-9.98*</td>
<td>-17.85</td>
<td>-2.11</td>
<td>4.02</td>
</tr>
<tr>
<td>Exercise pathology</td>
<td>3</td>
<td>2.70</td>
<td>-3.25</td>
<td>8.64</td>
<td>3.03</td>
</tr>
<tr>
<td>Muscle dissatisfaction</td>
<td>3</td>
<td>1.146</td>
<td>-0.26</td>
<td>0.72</td>
<td>0.715</td>
</tr>
</tbody>
</table>

*p<.05

NB: Positive SMD estimates indicate higher scores in the ED group and negative SMD estimates indicate higher scores in the MD group.
**MD vs. BDD**

Three studies compared individuals with MD with individuals with non-muscularity-related BDD (Niewouldt et al. 2015; Pope et al., 2005; Schneider et al., 2017). It should be noted that in Pope and colleagues’ study, of the 14 men in the MD sample, 12 reported additional non-muscle-related lifetime BDD, predominantly exhibiting concerns about hair and skin.

Results from our meta-analytic comparison are presented in Table 4. Similar to the previous two sections, there was evidence of heterogeneity among studies on all summary measures, so results should be interpreted with caution. The MD group reported significantly higher muscularity-oriented body change behaviors than the BDD group; however, MD was unable to be differentiated from BDD on the factors of eating pathology and obsessive compulsive pathology.

**Table 4. Results of continuous random-effects analyses for delineation from BDD**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Included studies</th>
<th>SMD estimate</th>
<th>95% CI, lower</th>
<th>95% CI, upper</th>
<th>Std. err</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eating pathology</td>
<td>3</td>
<td>-0.12</td>
<td>-1.53</td>
<td>1.28</td>
<td>0.72</td>
</tr>
<tr>
<td>Muscularity-oriented body change behaviors</td>
<td>3</td>
<td>-0.96*</td>
<td>-1.65</td>
<td>-0.28</td>
<td>0.49</td>
</tr>
<tr>
<td>Obsessive compulsive pathology</td>
<td>3</td>
<td>0.12</td>
<td>-0.33</td>
<td>0.56</td>
<td>0.23</td>
</tr>
</tbody>
</table>

*p<.05

NB: Positive SMD estimates indicate higher scores in the BDD group and negative SMD estimates indicate higher scores in the MD group.

**3.5. Follow-up studies (Robins and Guze – Phase 4)**

No naturalistic follow-up or longitudinal studies were found. We were able to identify three treatment case studies (Murray & Griffiths, 2015; Phillips et al., 1997; Ung et al., 2000) and two clinical trials of BDD, in which a subset of
participants reported preoccupation with muscularity/physique (Wilhelm et al., 2014; Wilhelm, Phillips, Fama, Greenberg, & Sketee, 2011).

The three case studies (Murray & Griffiths, 2015; Phillips et al., 1997; Ung et al., 2000) described the presentation and treatment of three young males with MD (15 years, 23 years, and 22 years, respectively). While one study reported the use of both pharmacological and therapeutic intervention (Ung et al.), the other reports described the use of pharmacological intervention alone (Phillips et al.) and therapeutic intervention alone (Murray & Griffiths, 2015). Both reports (Phillips et al.; Ung et al.) using pharmacological intervention prescribed antidepressant medications (Clomipramine and Sertraline); while, one therapeutic study applied a cognitive behaviour therapy approach (Ung et al.) and the final study (Murray & Griffiths, 2015) used an adapted Family Based Therapy (Lock & Le Grange, 2013) developed for the treatment of eating disorders in adolescents. Post-treatment symptomology differed across the three studies including: significant reductions in symptomology (Murray & Griffiths), decreases in exercise urges and distress (Phillips et al.), and symptom improvement (decreased preoccupation and rigidity) with some residual symptomatology (Ung et al.). Clinical trials of treatment effectiveness for BDD have also included individual participants reporting a preoccupation with physique. These studies have suggested efficacy for cognitive behavioral approaches (Wilhelm et al., 2011; Wilhelm et al., 2014), although it should be noted that results have not been presented by area of concern (i.e. concern related to hair or nose as opposed to muscularity).
3.6. Family studies (Robins and Guze – Phase 5)

No studies were identified that specifically explored heritability or family history in MD populations. However, one epidemiological study from Finland noted a potential familial transmission between MD symptomatology (undiagnosed) and anorexia nervosa in three sets of male twins (Raevouri et al., 2009). A further study with the same group assessed muscle dissatisfaction and use of muscle-enhancing substances (Raevouri, Keski-Rahkonen, Rose, Rissanen, & Kaprio, 2006). This study found that while muscle dissatisfaction was likely a combination of genetic and unique environment effects, use of muscle-enhancing substances was predicted by common and unique environment components (Raevouri et al., 2006).

4. Discussion

This study was the first to evaluate the clinical utility and validity of MD as a construct using Robins and Guze’s method (1970). Quantitative analysis identified two prospective clinical pictures of MD, which we termed a "purging disorder presentation" and a "body image disorder presentation." The meta-analytic results must be interpreted with caution due to high heterogeneity between studies; however, our results differentiated MD from healthy controls on measures of psychopathology and wellbeing, indicating it likely is a clinically relevant condition. Our attempts to delineate MD from existing diagnoses resulted in mixed findings. Individuals with MD were found to report lower traditional eating pathology than individuals with an ED and higher muscularity-oriented body change behaviors than individuals with BDD. Despite this, there were no differences between individuals with MD and ED on measures of muscle

NB: This study is currently under review. Please refer to the final published material when citing.
dissatisfaction or exercise pathology, nor between individuals with MD and BDD on measures of obsessive-compulsive pathology or eating pathology. Considering this, consistent with the two previous reviews using Blashfield’s guidelines (dos Santos Filho et al., 2016; Nieuwoudt et al., 2016), we did not find sufficient, meaningful evidence to delineate MD from existing psychiatric conditions. This review highlighted that the clinical presentation, classification, course, and prognosis of MD is still largely under-researched. Thus, we offer the following options for the future development of MD as a diagnosis.

4.1. MD as a purging disorder

The first clinical picture we described was a purging disorder presentation of MD. While there was a large overlap between the behavioral symptomatology included in both clusters, this presentation does not include preoccupation or body image disturbance described previously nor the mental acts and repetitive behavior criterion. To our knowledge, this is the first study to relate MD to purging disorder. Although, our delineation results suggested possible symptom overlap between MD and eating disorders, no individuals with purging disorder were identified in these studies.

Purging disorder has typically been conceptualized from a thinness-oriented perspective (Lydecker, Shea, & Grilo, 2017). Consequently, its symptomatology comprises the use of behaviors that aim to decrease adiposity or weight (Keel & Striegel-Moore, 2009). In a systematic review of the validity and reliability of purging disorder, the behavioral symptomatology of purging disorder is reported to include self-induced vomiting, and/or the misuse of laxatives, diuretics, or enemas (Keel & Striegel-Moore). Despite this, there has
recently been support for expanding compensatory behaviors conceptualized underneath this disorder to include excessive exercise. In a large-scale study, Lydecker and colleagues (2017) found that individuals with purging disorder symptomatology via driven exercise exhibited similar levels of weight/shape concern as individuals purging through traditional methods. As most research examining purging disorder has focused on self-induced vomiting and laxative use (Keel & Striegel-Moore), these results represent a promising step in increasing our understanding of more diverse compensatory behaviors and their related impairment and distress. Further research, especially in mixed gender samples (as with the study by Lydecker et al.), is required to better conceptualize purging/compensatory behaviors and how these may relate to MD symptomatology.

4.2. MD as a body image disorder

Our second and alternative clinical picture presented MD as a body image disorder. Laboratory studies provided inconsistent evidence for MD as a distinct entity from traditional eating disorders. While individuals diagnosed with MD reported lower levels of traditional eating disorder symptomology than individuals with an ED, no significant differences were found between the groups on the measures of exercise pathology or muscle dissatisfaction. Interestingly, while our findings indicated lower traditional eating pathology in MD populations, Murray and colleagues’ (2012) reported that self-reported symptoms of muscularity-oriented disordered eating did not differ between males with MD and males with anorexia nervosa. There was also further support for comparing MD and anorexia nervosa from an included clinical case study.

NB: This study is currently under review. Please refer to the final published material when citing.
(Murray & Griffiths, 2015). This study described the successful administration of Family-Based Therapy (Lock & LeGrange, 2013), a leading intervention for anorexia nervosa in adolescents, in the treatment of MD in a young adolescent male. Finally, a family study identified in our review described three males with anorexia nervosa who had a monozygotic or dizygotic twin with MD symptomology (Raevouri et al., 2009).

According to the DSM-5 (American Psychiatric Association, 2013), MD must be differentiated from eating disorders by the absence of concerns related to body fat/weight. However, we question whether the direction of an individual’s body dissatisfaction (i.e. muscularity-oriented or thinness-oriented) is meaningful in delineating diagnoses. Furthermore, multiple case studies have highlighted the importance of a “lean, muscular ideal,” in MD, with individuals engaging in strategies to decrease their body fat in an effort to highlight muscularity (Mosley, 2008; Murray et al., 2012; Murray & Griffiths, 2015). Our findings are preliminary and our inferences are limited due to an absence of research with strong methodology. However, they do highlight a pattern that, despite differences in thinness-oriented symptomology, there appears to be symptomatic overlap between MD and eating disorders, once the directionality of weight/shape preoccupation and related behaviors are removed.

4.3. MD as a BDD specifier

Our third suggestion would be to retain MD as a BDD specifier. As with both of the earlier recommendations, we found mixed support for this proposal. Firstly, our cluster analysis did not indicate body image distortion in either presentation. As distorted body image is a core component in BDD, these
findings may suggest body image distortion as a factor for demarcating MD from BDD. Furthermore, our delineation found differences between MD and BDD on measures of musculature-oriented body change behaviors. Despite this, our meta-analytic results indicated that MD was indistinguishable from BDD on measures of eating pathology and obsessive-compulsive symptomology. Our mixed findings are also reflected in the literature. In support of differentiating MD from BDD, individuals with MD have exhibited poorer quality of life, increased likelihood of a substance use disorder and previous suicide attempt relative (Pope et al.) as well as greater bodybuilding dependence, injury risk, and overall MD symptomatology (Niewoudt et al.) than individuals with non-musculature BDD. Conversely, these same researchers identified no significant differences between the groups on measures of BDD severity, non-muscle related preoccupation and behaviors, BDD-related impairment, lifetime psychiatric disorder, or delusionality (Pope et al.,) as well as eating pathology, muscle checking, or substance use (Niewoudt et al.). Upon examination of these studies, as well as research by Schneider and colleagues (2017), we noted that in nearly all cases individuals with MD had comorbid non-musculature oriented BDD. Consequently, it is possible that the identified differences – poorer quality of life, increased likelihood of a substance use disorder and previous suicide attempt as well as greater MD symptomatology – simply indicate greater impairment related to MD and concurrent BDD, as opposed to differences between the disorders. Retaining MD as a BDD specifier is further supported by clinical trials for BDD that have successfully included individuals with MD, potentially indicating similar treatment responses for these disorders (Wilhelm...
et al., 2011; Wilhelm et al., 2014). Although laboratory studies demonstrated that
individuals with MD are a clinically relevant population, this option raises the
question as to whether there is a need for MD to be differentiated from non-
muscularity orientated BDD.

4.4. Limitations and future research

Despite its strengths, there were a number of limitations of this research. As meta-analyses and systematic reviews only produce results as informative and important as the studies they incorporate, we recommend interpreting the results of this paper with caution. All meta-analyses conducted in this paper violated the assumption of homogeneity of studies. As such, ongoing research is required to delineate MD populations with healthy controls and individuals with similar existing diagnoses. We recommend that studies aiming to further examine the clinical validity of MD source sample from the general population (i.e. not bodybuilders and frequent exercisers) and utilize valid diagnostic measures. Furthermore, studies attempting to demarcate MD from existing diagnoses should include measures of exercise pathology, APED use, obsessive-compulsive pathology, and functional impairment as well as both muscularity and thinness-oriented measurements of eating and body image pathology. Finally, we recommend assessing delusionality in individuals with MD, as this has been recognized as a factor delineating BDD and obsessive-compulsive disorder (Phillips et al., 2012).

This review highlighted, and was itself limited by, inconsistent measurement and classification of MD in the extant literature. Despite the existence of multiple measures of MD symptomology, there appeared to be no
consensus on the most accurate method of determining clinical status. Furthermore, consistent with the findings of a recent systematic review by Sandgren and Lavalee (2018), we found that researchers are continuing to utilize Pope’s original diagnostic criteria rather than the DSM-5 diagnosis. Although, in itself, this is not problematic, this factor likely limits the inferences that can be drawn from our findings about the clinical utility of the DSM-5 MD diagnosis. Further research is required to assess the reliability of the syndrome of symptoms outlined in the DSM-5. Future studies would also benefit from comparing the two symptom presentations identified in our research.

Finally, our findings were mostly based on studies of low methodological quality (NMRHC, 2009). Although, it is common for early explorations of diagnoses to predominantly comprise case reports and observational studies, there is a clear need for increased research with rigorous methodology examining MD. The currently available data precludes determining a clear clinical picture of MD or better understanding the clinical course, treatment, heritability, and prognosis of the disorder. Considering this, we hope that our findings supporting MD as a clinically relevant psychiatric condition will encourage further research in this area. Studies assessing the heritability and temporal stability of MD are required to understand more about its prognosis and course, while we require greater understanding of the clinical picture of MD to develop effective treatment protocols.

4.5. Conclusions

Current research indicates that MD is likely a psychiatric condition distinct from a healthy pursuit of muscularity. Despite this, our study revealed
insufficient evidence for MD as a clinically distinct entity or as exhibiting a clearly defined clinical picture. Proposals of retaining MD as a BDD specifier as well as recategorizing it as a body image disorder or purging disorder are appealing in that they may allow for the wealth of knowledge and expertise in these fields to be extrapolated to MD. However, we acknowledge that our study was the subject of a number of limitations and that our conclusions were restricted by the small number of available research studies and the high heterogeneity between studies in the meta-analysis. It is hoped that further research that will provide further evidence for delineating these disorders, for developing more valid and reliable nosological systems, and for permitting the development of efficacious treatment protocols for this population.

NB: This study is currently under review. Please refer to the final published material when citing.
NB: * - studies included in systematic review only, ** studies included in meta-analytic and systematic review

References


*NB: This study is currently under review. Please refer to the final published material when citing.*


Dos Santos Filho, C. A., Passarelli Tirico, P., Stefano, S. C., Touyz, W. W., &


NB: This study is currently under review. Please refer to the final published material when citing.


NB: This study is currently under review. Please refer to the final published material when citing.


NB: This study is currently under review. Please refer to the final published material when citing.


*NB: This study is currently under review. Please refer to the final published material when citing.*
Table 1

Description of studies included in review based on Robins and Guze’s five-phase method

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study design</th>
<th>Sample</th>
<th>Clinical picture details (P1)</th>
<th>Laboratory comparisons (P2)</th>
<th>Discrimination from other disorders (P3)</th>
<th>Course, treatment, heritability, and prognosis (P4&amp;P5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babusa, Czegledi, Tury, Mayville &amp; Urban, 2015</td>
<td>Cross-sectional</td>
<td>304 male weightlifters</td>
<td>High risk MD = 46, Moderate risk MD = 158, Low risk MD = 100</td>
<td>High risk MD group significantly higher bodybuilding dependence, muscle checking, substance use, injury risk, exercise dependence, quantity-frequency of exercise, supplement use, current AAS use, and lifetime AAS use. NB: As the above factors were used in LCA to determine groups they were not used in analysis.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bo et al., 2014</td>
<td>Cross-sectional</td>
<td>N=440 students in Dietetics, Exercise and Sports Science and Biology. (males = 202, females = 218) MD traits =26 (MDDI&gt;39) Co = 414 (MDDI &lt;39)</td>
<td>• 73% male • Aged 20.2 • 23% reported concurrent disordered eating pathology • 42% exhibited concurrent orthorexia symptomology</td>
<td>MD traits group reported significantly more exercise, supplement use, dieting behaviors, than students not exhibiting MD traits.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NB: This study is currently under review. Please refer to the final published material when citing.
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Study Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cafri, Blevins &amp; Thompson, 2006</td>
<td>Case study</td>
<td>Male diagnosed with MD using structured clinical interview.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Age = 20 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Single</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Hispanic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- H/o preoccupation with leanness, substance dependence, panic disorder &amp; depressive episode.</td>
</tr>
<tr>
<td>Cafri, Olivardia &amp; Thompson, 2008</td>
<td>Cross-sectional</td>
<td>N=51 male weightlifters</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Current MD=15, Past MD=8, No current/past MD=28.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Age = 25 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 80% Caucasian</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 93% single</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Mean age of MD onset 19.17 yrs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 100% heterosexual</td>
</tr>
</tbody>
</table>

Current MD group had significantly higher BDD symptom severity, checking behaviours, preoccupation w/muscularity, bodybuilding dependence, muscle dissatisfaction, and functional impairment than individuals with past or no current/past MD. No differences on substance use and injury risk. Additionally, those with current/past MD had significantly greater lifetime prevalence of mood & anxiety disorder than no h/o MD group. No significant differences on lifetime diagnosis of eating disorder, non-muscle BDD, steroid use or steroid abuse/dependence.

NB: This study is currently under review. Please refer to the final published material when citing.
Chaney, 2008  Cross-sectional  N=304 bisexual & homosexual males, Low MD = 56 (25%ile MASS)  High MD = 55 (75%ile MASS)  

High MD group reported significantly higher loneliness & lower self-esteem.

Compte, Sepulceda, & Torrente, 2015  Cross-sectional  N=472 male university students  High MD (DMS-S>=31) = 200 (n=15 also had ED risk), Low MD (DMS-S<31) =272 (n=3 had ED risk)

Follow-up: 4 subthreshold AN 5 subthreshold BN MD (DMS-S>=52) = 33  

At risk for MD group reported significantly higher eating pathology, anxiety, use of supplements, exercise, and body dissatisfaction as well as lower self-esteem than not at risk group.

At risk MD group exhibited lower body dissatisfaction and exercise than concurrent MD and ED group. No differences on body dissatisfaction between MD only and ED only groups. No significant differences on self-worth, anxiety, worry or supplement use from MD and ED, or ED only groups.

Follow-up: MD group had significantly lower eating pathology and higher drive for masculinity than ED groups. No differences in body dissatisfaction, exercise or supplement use.
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Findings</th>
</tr>
</thead>
</table>
| Danilova, Diekhoff & Vandehey, 2013       | Quasi-experimental | N=77 males recruited from universities and gyms  
High risk MD (MDI>84) = 16  
Low risk MD (MDI<84) = 41  
MDI•Age=26.6 yrs  
316 mins/week exercise  
1.8 days/week diet  
3.9 days/week supplement use | High risk MD group reported lower body satisfaction and higher discrepancy between actual and ideal body than low risk group.  
No significant differences in dieting, supplement use, exercise frequency between high and low risk groups. |
| Fabris, Longobardi, Prino & Settani, 2017 | Cross-sectional | N=170 male bodybuilders  
High risk MD=44  
(MDDI >39)  
Co=126 | MD group reported significantly higher levels of insecure attachment, lower confidence, higher discomfort with intimacy, relations as secondary, need for approval and preoccupation with relationship than comparison group.  
No significant differences between 25%ile and 75%ile for involvement in competitive bodybuilding or use of anabolic substances. |
Weekly training 8.22 hrs in males & 6.56  
44% of males with MD were suspected of AAS use, with 47% of females with MD suspected of AAS use. | MD group reported significantly lower sport ability, physical condition, and strength than comparison group. MD group also rated themselves as significantly less attractive, having poorer physical self-concept and general self-concept than comparison group. MD |
**Gonzalez-Marti, Fernandez-Bustos, Contreas Jordan & Sokolova, 2017**

Cross-sectional

N= 323 (113 male, 210 female)  
High MD =25 (MMDS >= 24)  
Low MD=28 (MMDS < 4)

- Age = 21.6 yrs

NB: As the above were included in the LCA to group MD participants, they were not included in analyses.

**Goodale, Watkins, & Cardinal, 2001**

Cross-sectional

N= 323 (113 male, 210 female)  
High MD =25 (MMDS >= 24)  
Low MD=28 (MMDS < 4)

- Age = 21.6 yrs  
- High MD group reported significantly greater ED pathology, depression, sociocultural influences, distress when missing exercise, & less social support than Low MD group.

**Gruber & Pope, 1999**

Case report

N=75 female bodybuilders  
Cl = 10 females with h/o experiencing rape.  
Co = 56 females without reported h/o rape.

- Age = 31 yrs  
- Caucasian female. No h/o psychiatric diagnosis. Reported h/o of rape and physical assault precipitating MD.  
- 50% of Cl group reported history of ED, 80% reported PTSD post-rape, 70% use of AAS post-sexual assault.

**Hernandez-Martinez, Gonzalez-Cross-sectional**

N=32 male weightlifters  
At risk MD

- FFMI=22.38

No inferential statistics comparing groups. Correlational analyses showed significant

**NB: This study is currently under review. Please refer to the final published material when citing.**
<table>
<thead>
<tr>
<th>Authors</th>
<th>Study Type</th>
<th>Sample Size</th>
<th>Participants</th>
<th>Associations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marti, Contreas &amp; Jordan, 2017</td>
<td></td>
<td>(ESM &gt;=3) = 2 Co = 30</td>
<td>• Associations between ESM category and bodybuilding dependence, muscle checking, substance, injury risk and muscle dissatisfaction.</td>
<td></td>
</tr>
<tr>
<td>Hildebrandt, Schlundt, Langenbuche &amp; Chung, 2006</td>
<td>Cross-sectional</td>
<td>N=237 male gymnasium users, MD = 64, Muscle concerned =63, = Fat concerned=66, Normal bodybuilders=38, Co=30</td>
<td>• Age 31.83 yrs MD group had higher body dissatisfaction, obsessive-compulsive symptomatology, and physique anxiety than non-bodybuilders and normal bodybuilders, Higher drive for bulk and bulimia than fat concerned, normal bodybuilders, and non-bodybuilders. Higher bulimia than all others. No significant differences in training history.</td>
<td></td>
</tr>
<tr>
<td>Hitzeroth, Wessels, Zungu-Dirwayi, Oosthuizen &amp; Stein, 2001</td>
<td>Cross-sectional</td>
<td>N = 28 amateur bodybuilding competitors (males = 24, females = 4) MD = 15 (males</td>
<td>• Age = 25.47 yrs • 73% single MD group reported significantly higher supplement use, BDD preoccupations other than muscularity, and previous medical consultations than comparison group. No</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Sample Size</td>
<td>Characteristics</td>
<td>Findings</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>-------------</td>
<td>----------------</td>
<td>---------</td>
</tr>
<tr>
<td>Longobardi, Prino, Fabris &amp; Settani, 2017</td>
<td>Cross-sectional</td>
<td>N=145 male bodybuilders (MDDI&gt;39) = 36 Co (MDDI&lt;39) = 109</td>
<td>Age = 26.5 yrs</td>
<td>MD group had significantly higher levels of psychological symptomology and dissociative experiences than comparison group. No significant differences in participation of bodybuilding competitions or AAS-use.</td>
</tr>
<tr>
<td>Macik &amp; Kowalska-Dabrowska, 2015</td>
<td>Cross-sectional</td>
<td>N=85 regular gym users (Novel screening checklist&gt;26) = 30 12 women, 18 men</td>
<td>Training 4 or more times per week. Attend gym for physique and fitness reasons as well as improving mood, health, advice, and spend free time. Use gym classes for confidence, physique, mood, dissociation, attractiveness, friendships. Report feeling satisfaction, tired, strong, check, impatience, quite, insufficient post-workout. 60% described poor social life.</td>
<td>Significant differences in gym hours/week, psychiatric history, use or substance use.</td>
</tr>
</tbody>
</table>

NB: This study is currently under review. Please refer to the final published material when citing.
<table>
<thead>
<tr>
<th>Reference</th>
<th>Study Type</th>
<th>Case Summary</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mosley, 2009</td>
<td>Case study</td>
<td>Male bodybuilder meeting Pope et al. (1997) criteria</td>
<td>Age = 27 yrs, FFMI = 28, H/o depression &amp; bulimia nervosa, H/o restrictive dieting, weight concerns (“scrawny”)</td>
</tr>
<tr>
<td>Murray &amp; Griffiths, 2015</td>
<td>Case study</td>
<td>Male with MD diagnosed with Pope criteria</td>
<td>Age = 15 yrs, Reported experiences of bullying, feeling weak (self-comparison with “more manly” boys).</td>
</tr>
<tr>
<td>Murray, Maguire, Russell &amp; Touyz, 2011</td>
<td>Case study</td>
<td>Male with MD using unstructured clinical interview</td>
<td>Age = 20yrs, FFMI = 20.5, H/o childhood obesity, bullying related to weight, Precipitated by severe dietary restriction and intense exercise.</td>
</tr>
<tr>
<td>Murray, Rieger, Hildebrandt, Karlov, Russell, Boon, Dawson &amp; Touyz, 2012</td>
<td>Cross-sectional</td>
<td>N=60 males Males with MD = 21 Males with AN = 24 Co = 15 gymnasium-using males</td>
<td>MD group reported significantly higher eating pathology, Muscularity-oriented eating pathology, MD symptomology, Compulsive exercise, and APED use than healthy controls. MD group reported higher MD pathology, APED use, and lower eating pathology than AN group. No differences in Muscularity-oriented eating pathology, or Compulsive exercise between MD and AN group.</td>
</tr>
</tbody>
</table>

*NB: This study is currently under review. Please refer to the final published material when citing.*
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participant Characteristics</th>
<th>Findings</th>
</tr>
</thead>
</table>
| Murray, Rieger, & Touyz, 2011             | Case study      | Male with MD using semi-structured clinical interview                                       | • Age = 32 yrs  
• FFMI = 28  
• No previous psychiatric history.  
• H/o weight concern (“chubby”) and dietary restraint. |
| Nieuwoudt, Zhou, Coutts & Booker, 2015    | Cross-sectional | N=648 male weightlifters  
MD (75%ile MASS)=110 (n=36 with comorbid MD & BDD, n=60 with comorbid MD & eating disorder)  
ED = 219 (n=43 with comorbid BDD)  
BDD =69  
Co=250                      | Means and standard deviations were reported for each group on measures of MASS and EAT-26; however, statistical test results comparing no disorder to MD group were not presented.  
MD only group significantly higher than BDD only group on bodybuilding dependence, injury risk, and total MD symptoms.  
MD only group significantly higher than ED only group on bodybuilding dependence, muscle checking, substance use, injury risk, muscle satisfaction. No difference on ED pathology between all groups. |
| Olivardia, Pope, & Hudson, 2000  
And Choi, Pope & Olivardia, 2002 | Cross-sectional | N=54 male weightlifters  
MD= 24 men with, Co = 30 male weightlifters  
• FFMI = 25.2  
• Single 100%  
• Caucasian 87.5%  
• Age of onset = 19.4 yrs  
• 42% demonstrated excellent insight; 50% fair or poor insight | MD group reported significantly greater likelihood of past or current psychiatric diagnosis (MD, AD, ED), increased obsessive compulsive symptomology, reported poorer body image, increased eating pathology, lower appearance |
| Phillips, O'Sullivan & Pope, 1997 | Case report | Male with MD diagnosed with structured clinical interview | • Onset of steroids at least 1 year following MD 73% antecedent family relationship difficulties  
  • Age = 23 yrs  
  • H/o BDD, depression. Family h/o panic disorder and possible motor tics.  
  • Phillips, O'Sullivan & Pope, 1997 Case report  
  Male with MD diagnosed with structured clinical interview  
  • Age = 23 yrs  
  • H/o BDD, depression. Family h/o panic disorder and possible motor tics.  
  • Phillips, O'Sullivan & Pope, 1997 Case report  
  Male with MD diagnosed with structured clinical interview  
  • Age = 23 yrs  
  • H/o BDD, depression. Family h/o panic disorder and possible motor tics.  
  • Phillips, O'Sullivan & Pope, 1997 Case report  
  Male with MD diagnosed with structured clinical interview  
  • Age = 23 yrs  
  • H/o BDD, depression. Family h/o panic disorder and possible motor tics.  |
|---|---|---|---|
| Pope, Gruber, Choi, Olivardia & Phillips, 1997 | Case reports | 2 males and 2 females with MD | • 27 year old heterosexual, Caucasian male, FFMI =28. H/o substance dependence & depression.  
  • 22 year old homosexual Caucasian male, FFMI 25.9. H/o bipolar disorder, panic disorder, & OCD.  
  • 29 year old female bodybuilder. H/o of AN, BN, & physical assault.  
  • 29 year old Caucasian female with no psychiatric history.  
  • 27 year old heterosexual, Caucasian male, FFMI =28. H/o substance dependence & depression.  
  • 22 year old homosexual Caucasian male, FFMI 25.9. H/o bipolar disorder, panic disorder, & OCD.  
  • 29 year old female bodybuilder. H/o of AN, BN, & physical assault.  
  • 29 year old Caucasian female with no psychiatric history.  
  • 27 year old heterosexual, Caucasian male, FFMI =28. H/o substance dependence & depression.  
  • 22 year old homosexual Caucasian male, FFMI 25.9. H/o bipolar disorder, panic disorder, & OCD.  
  • 29 year old female bodybuilder. H/o of AN, BN, & physical assault.  
  • 29 year old Caucasian female with no psychiatric history.  |

*NB: This study is currently under review. Please refer to the final published material when citing.*
| **Pope, Katz & Hudson, 1993** | Case reports | 108 male body builders (51% AAS use) MD = 9 Case report described 3 males with MD | - 19 year old Caucasian male. H/o AN & hypomania.  
- 27 year old Caucasian male. H/o restrictive dieting, hypomania, weight concern & AN.  
- 26 year old Caucasian male with AAS dependence. |
| **Pope, Pope, Menard, Fay, Olivardia & Phillips, 2005** | Cross-sectional | N = 63 males with lifetime BDD MD = 14 Non-MD BDD = 49 | - Age = 36 yrs  
- 100% Caucasian  
- 85.7% single  
- 86% also had other BDD areas of concern  
MD group was significantly more likely to lift weights excessively, had more areas of concern, higher lifetime suicide attempts, substance use disorders poorer QOL, and mental health than non-MD BDD group. No significant differences in BDD severity, delusionality, functional impairment, social functioning. |
| **Raevouri, Hoek, Susser, Kaprio, Rissanen & Keski-Rahkonen, 2009** | Observational | N = 2122 male twins born in Finland Cl = 5 individuals with lifetime AN (3 of whom had | Pair I – Monozygotic Male-Male, both twins exhibited MD symptomology.  
Pair 2 – Dizygotic Male-Male, co-twin exhibited MD symptomology.  
Pair 4 – Dizygotic Male- |
<table>
<thead>
<tr>
<th>Study</th>
<th>Methodology</th>
<th>Sample Description</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raevouri, Keski-Rahkonen, Rose, Rissanen &amp; Kaprio, 2009</td>
<td>Observation</td>
<td>N = 319 male twin pairs (636 twins) and 85 twin individuals, total 721, born in Finland Cl = 220 individuals high muscle dissatisfaction</td>
<td>Male, co-twin exhibited MD symptomology. Higher correlations of muscle dissatisfaction and APED use were found between monozygotic than dizygotic twins. The contribution of both genetic and unique environment effects predicted dissatisfaction, however, environmental factors best predicted APED use.</td>
</tr>
<tr>
<td>Schneider, Mond, Turner, &amp; Hudson, 2017.</td>
<td>Cross-sectional</td>
<td>N = 162 adolescents (male=93, female = 69) with full or partial BDD</td>
<td>Males with MD had higher drive for muscularity, and more areas of concerns. No differences in QOL, BDD symptom severity, or comorbid symptomology.</td>
</tr>
<tr>
<td>Ung, Fones &amp; Ang, 2000</td>
<td>Case study</td>
<td>MD = 41 males Male diagnosed with MD.</td>
<td>Tx=Sertraline and CBT Post-tx BDD-YBOCS = 30 Post-tx – decreased preoccupation over 2 years. Increased flexibility w/ workout and diet. Residual dysphoria and negative cognitions.</td>
</tr>
</tbody>
</table>

*NB: This study is currently under review. Please refer to the final published material when citing.*
and work stress.

| Wilhelm, Phillips, Didie, Buhlmann, Greenberg, Fama, Keshaviah & Steketee, 2014 | Randomized Controlled Trial | N = 36 individuals with BDD (2 exhibited primary concerns related to body build) | Tx – 22 sessions of CBT-BDD Post tx participants reported decreased BDD symptomology, impairment and depressive symptoms as well as improved insight. Results were maintained at 6mth f/up. |
| Wilhelm, Phillips, Fama, Greenberg & Steketee, 2011 | Pilot trial | N = 12 individuals with BDD (2 exhibited primary concerns of body shape/size) | Tx – 18 sessions of modular CBT for BDD Post tx – participants reported decreased BDD symptomology & depression as well as improved insights. Tx gains were maintained at 3 & 6mth f/up |

Table notes: AAS; anabolic androgenic steroids; AD – anxiety disorder; AN – anorexia nervosa; APED – appearance/performance-enhancing drugs; BDD – body dysmorphic disorder; BDD-YBOCS – Yale Brown Obsessive-Compulsive Scale Modified for Body Dysmorphic Disorder; BN – bulimia nervosa; Cl – clinical group; Co – comparison group; CBT – Cognitive behavior therapy; DMS – Drive for Muscularity Scale; DMS-S – Drive for Muscularity – Spanish version; EAT-26 – Eating Attitudes Test; ED – eating disorder; ESM – Muscle Appearance Satisfaction Scale – Spanish version; FFMI – Fat Free Mass Index; LCA – Latent Class Analysis; MASS – Muscle Appearance Satisfaction Scale; MD – muscle dysmorphia MDDI – Muscle Dysmorphia Diagnostic Inventory; MDDS – Modified Muscle Dysmorphia Scale; MDI – Muscle Dysmorphia Inventory; QO – quality of life; Tx – treatment.

NB: This study is currently under review. Please refer to the final published material when citing.
CHAPTER 3: Study 2
Foreword

The findings from the Study 1 review suggest that the use of appearance and performance enhancing substances/supplements (APED) represents a differentiating factor between muscle dysmorphia (MD) and healthy controls. APED include the use of both illicit substances such as anabolic androgenic steroids, peptide hormones, growth factors, Beta-2 agonists, hormone and metabolic modulators, diuretics, stimulants, narcotics, cannabinoids, and glucocorticoids as well as licit supplements such as caffeine, protein supplements, nitric oxide, and glutamine (World Anti Doping Agency, 2018). As MD was originally identified in a study of anabolic steroid users (Pope, Katz, & Hudson, 1993), the use of ergogenic substances has been considered a common feature in its clinical picture. Consistent with this conclusion, a previous systematic review of research investigating appearance and performance enhancing use of APEDs reported that increased APED usage was associated with overall MD symptomology and MD-related impairment (Hildebrandt et al., 2011). Furthermore, individuals reporting the longest durations of APED usage have been found to report more severe body image disturbance than short-term APED users (Kanayama, Barry, Hudson, & Pope, 2006). Despite awareness of APED use as a key public health concern, there is currently a lack of psychometrically sound survey measures to assess APED usage in large populations. Consequently, this chapter aimed to address the following research question:

NB: This study is currently under review. Please refer to the final published material when citing.
RQ 3: Can we accurately measure symptoms of MD and, where required, develop psychometrically sound instruments to measure these symptoms?

The APED measure, the Appearance and Performance-Enhancing Supplement and Substance Use Scale (APES), was initially developed as a pilot project in a previous unpublished thesis (Cooper, unpublished dissertation). This study was limited to a small sample size and did not examine the psychometric properties of the instrument with the exception of its internal reliability. The paper in this chapter reports a more comprehensive psychometric evaluation of the APES using a large community sample.

**Reference not in article**

Preliminary psychometric characteristics of a measure for appearance and performance-enhancing supplement and substance usage
By Marita Cooper, Richard Hicks & Kathleen M Griffiths

Abstract

Background: The use of both licit and illicit appearance and performance-enhancing drugs (APED) has been recognized as a key public health concern. While it is estimated that rates of APED usage in non-athlete populations is increasing, true prevalence studies are limited by a lack of psychometrically reliable and valid measures. This paper describes the development and psychometric evaluation of a survey tool for assessing APED usage. Method: Initial items for the Appearance and Performance Enhancing Supplement/Substance Use Scale (APES) were 7 supplements and 13 substances. Following a preliminary exploration, we evaluated the factor structure, concurrent validity, internal consistency, and test-retest reliability of the APES in a community sample of 265 participants. Results: The APES demonstrated strong internal consistency, medium to strong concurrent validity, and good to excellent test-retest reliability. A two-factor model of supplements and substance use items was supported for the data. Conclusions: These results provide preliminary support for the reliability and the validity of APES in assessing both licit and illicit APED usage. Future studies are recommended to replicate and extend the psychometric investigations reported here to encompass additional validation studies and populations.

Highlights
- Appearance- and performance-enhancing drug (APED) use is a key public health issue.
- Appearance and Performance-Enhancing Supplement Use Scale (APES) surveys licit & illicit APED use

NB: This study is currently under review. Please refer to the final published material when citing.
• Principal component analysis supported the two-factor model of supplement and substance use
• Psychometrics suggest the APES possesses strong validity and reliability

Keywords
Appearance and performance enhancing drug use; Psychometrics; Dietary supplements; Anabolic-androgenic steroids; Polysubstance use; Muscularity

NB: This study is currently under review. Please refer to the final published material when citing.
Development and validation of a measure for appearance and performance-enhancing supplement and substance usage

1. Introduction

Use of appearance and performance enhancing drugs (APED) is steadily increasing, with population lifetime steroid usage estimated to be 3.3% worldwide.\(^1\) Though the rise in use of these substances began initially with elite athletes using anabolic-androgenic steroids (AAS) to boost performance,\(^2\) the profile of AAS users has now broadened to include amateur bodybuilders and gym-users.\(^3\) Additionally, musculature enhancement and self-esteem are now recognized as common motivations for APED usage.\(^3\)-\(^5\) The proliferation of AAS use is occurring despite well-documented adverse psychological and physical effects of AAS including heightened risk for cardiomyopathy, cancer, development of a major mood disorder and substance dependence.\(^6\) As such, developing a clearer understanding of the etiology, prognosis, and population engaging in AAS use, as well as other APED, is a key public health concern.\(^6\)

Although AAS abuse has been the key focus of much research into APED usage, users often utilize AAS in conjunction with other illicit substances to increase leanness, support recovery, improve endurance, enhance sexual performance, reduce inflammation, and/or reduce side effects of other APED.\(^7\) The World Anti-Doping Agency also refers to the use of peptide hormones, growth factors, Beta-2 agonists, hormone and metabolic modulators, diuretics, stimulants, narcotics, cannabinoids, and glucocorticoids on their most recent list of prohibited substances.\(^8\) The effects of APED polypharmacy are at present under-researched; however, initial findings suggest that abuse of multiple substances is associated with cardiovascular disease, increased violence, poorer psychopathology and sudden death.\(^9,\,10\)

Although little is known about the etiology of APED use, preliminary research has indicated that use of nutritional supplements may precede use of illicit substances.\(^11\) Although the use of temporal precedence to indicate a “gateway” association has been questioned,\(^12\) researchers have also found that supplement users are more likely than non-supplement users to indicate positive attitudes towards doping and to perceive APED usage as safe.\(^11,\,12\) Nutritional supplements are typically posited as a more natural and safe alternative to the

\(\text{NB: This study is currently under review. Please refer to the final published material when citing.}\)
illicit substances outlined earlier, whilst still achieving similar appearance and performance enhancing benefits. Pharmacological examination of nutritional supplements, however, has found that products: frequently contain different ingredients to those labeled; can be contaminated with banned substances as well as harmful toxins; and some lack support for their claimed benefits.19

Despite attempts to apply systematic techniques to the understanding of APED usage (from hereon referring to both illicit substance and licit supplement usage), we must be cautious in interpreting population findings due to a lack of psychometrically supported measures of APED use (see reviews by Ntoumanis et al., Piacentino et al., and Sagoe et al.). In all but two studies identified in these reviews, authors utilized interview schedules or medical examinations inappropriate for large population studies and/or individual questions without psychometric integrity to identify APED usage. For example, Kanayama and colleagues24 reported on APED usage in a sample of 511 gym users. However, the questionnaire only asked participants about use of protein, creatine, androstenedione, ephedrine and anabolic steroids, a list that failed to encompass the rich array of APED available and commonly utilized. Such measurements of APED usage that are based on a small selection of APED, combining substances into broad categories (e.g. dietary supplements, muscle-building products), or use open-ended questions where individuals may fail to recognize products as an APED (e.g. caffeine, diuretics), may lead to underestimations of APED use.

Considering the proliferation of both licit and illicit APED usage, improved understanding of predictors, long-term effects, and efficacious treatments in large samples is key for this public health issue. This study aimed to develop a psychometrically valid and reliable survey tool for assessing APED usage.

2 Materials and method

2.1. Participants

The sample comprised 265 participants, aged between 18-73 years ($M = 32.65$ years, $SD = 9.44$ years). There were 85 males (32%) and 176 females (68%), with a mean Body Mass Index (BMI) of 25.2 ($SD=5.11$) across the sample. Follow-up testing was undertaken on a self-selected subset of the sample, comprising 61 participants, aged between 19 and 61 years ($M= 31.14$, $SD= 7.62$)
of whom 18 (30%) were males and 42 (70%) were females (gender data was missing for one participant). The BMI of the follow-up sample was 25.0 (SD=4.70). There were no significant differences in age, gender or BMI between the participants who participated in the follow-up survey and those who completed the first survey only.

Participants were recruited utilizing a snowball method via Facebook as part of a larger study into muscle dysmorphia symptomatology. Three key methods of recruitment were utilized: 1) a study Facebook group containing information about the study and an invitation to participate; 2) paid Facebook advertisements inviting individuals to participate targeting individuals interested in Wellbeing, Nutrition, Crossfit, Gym-users, and Bodybuilding (no country restriction); and 3) posts about the study on Facebook groups comprising potentially at-risk participants (such as Wellbeing, Nutrition, Gym interest groups) and by Lifestyle/Wellness professionals. All participants were offered the opportunity to enter a competition to win a $100 gift voucher for an online food delivery service.

Participation was voluntary and confidential. Consent was requested at the beginning of the study. The study was approved by the Australian National University Human Research and Ethics Committee (Protocol No.: 2016/585).

2.2 Appearance and Performance Enhancing Supplement/Substance Use Scale.

The APES was developed to evaluate individual use of supplements and substances for the purpose of body modification and enhanced performance. It was intended for administration in general population epidemiological studies, in at-risk populations such as athletes and gym users, and in clinical settings. Using a wording format developed by Kanayama and colleagues in their study of AAS and supplement use among gymnasium clients, respondents are provided with both the brand names and street names of common supplements and asked to indicate on a 4-point Likert scale “how much time you have used each of these supplements in the past 3 years” (0=never, 1= a little (0-1 month), 2=moderate (1-6 months), 3=a lot (6 months or more). Supplements and substances were selected from several sources including: Kanayama and colleagues measure; an empirical review of commonly used muscle enhancing substances; the

NB: This study is currently under review. Please refer to the final published material when citing.
World Anti-Doping Agency’s 2012 List of Prohibited Substances and Methods; and research from Australian Supplement supplier websites. This resulted in an item pool comprising 7 supplements and 13 substances (see Appendix A). Participants were also provided with the option of including other supplements or substances used as appropriate. Items were worded so that higher scores indicated higher levels of substance and supplement use.

Prior to the current study, we evaluated the APES items in a pilot sample comprising 67 participants aged 17 years and over who were recruited from undergraduate psychology courses, local gymnasiums and supplement stores in the Gold Coast region of Queensland, Australia (Cooper, 2013, unpublished thesis). This pilot study was approved by the Bond University Human Research and Ethics Committee (Protocol No: RO1087). Participants comprised 48 males and 19 females aged 18-64 years (M = 27.8 years, SD = 10.1 years). The pilot study demonstrated strong internal reliability for the APES Supplement use items (Cronbach’s α = .84), the Substance use items (Cronbach’s α = .85), and the overall APES items (Cronbach’s α = .87). To determine whether removal of any items would improve reliability, item-total correlations were inspected. Three items (caffeine, non-steroidal anti-inflammatories, and amphetamines) showed item-total correlations below .40 (Cronbach’s α = .89 Supplement use; .86 Substance use). However, the increase in Cronbach’s alpha was marginal following their removal. Accordingly, all items were retained in the current study.

2.3. Measures of concurrent validity

Three measures were included to assess concurrent validity. It was predicted that they would show a moderate to strong association with the APES.

2.3.1. Drive for muscularity. The Drive for Muscularity (DMS) is a 15-item measure developed to assess attitudes and behaviors related to the pursuit of muscularity. The DMS has demonstrated strong internal consistency and test-retest reliability.

2.3.2. Muscularity-oriented disordered eating. The Eating for Muscularity Scale (EMS) is a 30-item instrument that measures pathological eating attitudes and behaviors driven by a desire to enhance muscularity. The EMS has high internal
consistency, moderate to strong concurrent validity with associated measures, good construct validity and strong test-retest reliability.\textsuperscript{29}

**2.3.3. Obsessive exercising.** Obsessive exercising was evaluated using the Obligatory Exercise Questionnaire (OEQ).\textsuperscript{30} The OEQ is a brief 20-item measure comprising three subscales - exercise fixation, exercise frequency and exercise commitment - which aim to identify individuals at risk for an eating disorder.\textsuperscript{31} The instrument employs a 4-point Likert scale where 1 represents ‘Never’ and 4 represents ‘Always.’ The OEQ has demonstrated strong internal consistency as well as strong concurrent validity with other measures of exercise dependence and addiction.\textsuperscript{32, 33}

**2.4 Other measures**

Participants completed demographic questions (age, gender, height, weight, ethnicity, and sexual orientation). The survey also contained other measures that are not the subject of the current paper.

**2.4 Procedure**

Interested participants accessed the Information Sheet and survey through Google Docs. The Information Sheet provided participants with an overview of the research purpose and procedure as well as any risks or benefits associated with participation. At the end of the questionnaire, participants were offered an option to “opt-out” of receiving the follow-up questionnaire. The latter was administered 2 weeks after the submission of the initial survey data.

**2.5 Statistical analyses**

Data were analyzed using IBM SPSS Statistics v25.\textsuperscript{34} Item 14 was removed from the DMS Attitude score as well as the overall score as this question (related to steroid use) was deemed to overlap with the APES. Descriptive statistics were calculated for the demographic characteristics of the sample and the APES items. Visual inspection of box plots on all subscale and score level variables indicated the presence of outliers in the data but as extreme scoring is expected on such scales all data were retained. The APES subscales and total scores violated normality assumptions ($p < .001$). The homogeneity of variances assumption was met for the Substance subscale ($p = .11$), but, was violated for both the Supplement and total score levels. However, given its larger sample size,
parametric testing was deemed viable. Gender differences in supplement and substance use were examined using independent t-tests and repeated-measure t-tests were used to explore differences in mean supplement and substance use. Since the APES items met assumptions for both the Kaiser-Meyer-Olkin (KMO= .78) measure of sampling adequacy and Bartlett’s Test of Sphericity ($\chi^2(190)=3947.89$, $p<.001$), a principal component analysis was conducted to examine the factor structure of the APES in which coefficients below .4 were suppressed. Items were retained if they met at least a .4 factor loading on their first factor and did not load .4 or higher on any other factor. The total number of components retained was determined using the Kaiser criterion, scree test and Horn’s parallel analysis. Cronbach’s coefficient alpha was again used to evaluate internal reliability. Finally, two-tailed Pearson’s r correlations were conducted to assess concurrent validity and temporal stability.

3. Results

3.1 Descriptive results

Descriptive results are shown in Table 1. There were higher levels of supplement use than substance use for both males and females ($t(79) = 9.67$, $t(166) = 9.81$ respectively; $p<.001$). Although supplement use frequency was higher in males than females ($t(250) = 3.95$, $p<.001$), there was no significant difference in substance use frequency between men and women ($t(252)= 0.99$, $p = .33$). As can be seen in Table 2, the most popular supplements reported were protein, caffeine, non-steroidal anti-inflammatoryatories, and Branch Chain Amino Acids (BCAA) and amino acids in both males and females. Table 2 also shows the most commonly endorsed substances were Diuretics, Beta-2 Agonists, Antiestrogenics, and Amphetamines in males; females most frequently used Diuretics and Beta-2 Agonists. Overall, 70% of the sample reported using at least one supplement and 7% of the sample reported using one or more illicit APEDs.
Table 1
*Use of APES by gender*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean (SD)</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>APES – Supplement use</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females (n=176)</td>
<td>3.05 (4.71)</td>
<td>2.00</td>
<td>0-22.00</td>
</tr>
<tr>
<td>Males (n=85)</td>
<td>6.39 (5.64)</td>
<td>6.00</td>
<td>0-22.00</td>
</tr>
<tr>
<td>Total (n=261)</td>
<td>4.53 (5.03)</td>
<td>3.00</td>
<td>0-21.00</td>
</tr>
<tr>
<td><strong>APES – Substance use</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females (n=176)</td>
<td>0.19 (1.66)</td>
<td>0.00</td>
<td>0-21.00</td>
</tr>
<tr>
<td>Males (n=85)</td>
<td>0.40 (1.67)</td>
<td>0.00</td>
<td>0-13.00</td>
</tr>
<tr>
<td>Total (n=261)</td>
<td>0.35 (1.85)</td>
<td>0.00</td>
<td>0-22.00</td>
</tr>
<tr>
<td><strong>APES - Overall</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females (n=176)</td>
<td>3.96 (5.26)</td>
<td>2.00</td>
<td>0-34.00</td>
</tr>
<tr>
<td>Males (n=85)</td>
<td>6.79 (6.21)</td>
<td>6.00</td>
<td>0-27.00</td>
</tr>
<tr>
<td>Total (n=261)</td>
<td>4.89 (5.74)</td>
<td>3.00</td>
<td>0-34.00</td>
</tr>
</tbody>
</table>

Table 2
*Frequency of participants reporting one month or more APED usage by gender*

<table>
<thead>
<tr>
<th>Item</th>
<th>Females</th>
<th>Males</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>APES – Supplement use</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Protein</td>
<td>104 (58.8%)</td>
<td>63 (75%)</td>
<td>168 (63.6%)</td>
</tr>
<tr>
<td>2. Creatine</td>
<td>27 (15.3%)</td>
<td>29</td>
<td>56 (21.2%)</td>
</tr>
<tr>
<td></td>
<td>(34.5%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Glutamine</td>
<td>20 (11.4%)</td>
<td>22</td>
<td>42 (16%)</td>
</tr>
<tr>
<td></td>
<td>(26.2%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. BCAA and amino acids</td>
<td>34 (19.3%)</td>
<td>33</td>
<td>68 (25.5%)</td>
</tr>
<tr>
<td></td>
<td>(39.3%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Caffeine</td>
<td>60 (34.1%)</td>
<td>48</td>
<td>109 (41.4%)</td>
</tr>
<tr>
<td></td>
<td>(57.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Nitric Oxide</td>
<td>16 (9.1%)</td>
<td>13</td>
<td>29 (11%)</td>
</tr>
<tr>
<td></td>
<td>(15.5%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Non-steroidal anti-</td>
<td>34 (19.3%)</td>
<td>35</td>
<td>71 (27%)</td>
</tr>
<tr>
<td>inflammatories</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NB: This study is currently under review. Please refer to the final published material when citing.
**APES – Substance use**

<table>
<thead>
<tr>
<th>Substance Type</th>
<th>1 (0.6%)</th>
<th>2 (1.1%)</th>
<th>3 (1.1%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testosterone precursors or prohormones</td>
<td>2 (1.1%)</td>
<td>9 (10.8%)</td>
<td>11 (4.2%)</td>
</tr>
<tr>
<td>Ephedrine</td>
<td>3 (1.7%)</td>
<td>5 (6%)</td>
<td>8 (3%)</td>
</tr>
<tr>
<td>Diuretics</td>
<td>5 (2.8%)</td>
<td>3 (3.6%)</td>
<td>8 (3%)</td>
</tr>
<tr>
<td>Anabolic-androgenic steroids</td>
<td>1 (0.6%)</td>
<td>2 (2.4%)</td>
<td>3 (1.1%)</td>
</tr>
<tr>
<td>Beta-2 agonists</td>
<td>3 (1.7%)</td>
<td>3 (3.6%)</td>
<td>6 (2.3%)</td>
</tr>
<tr>
<td>Human-growth hormones</td>
<td>2 (1.1%)</td>
<td>1 (1.2%)</td>
<td>3 (1.1%)</td>
</tr>
<tr>
<td>Human chorionic gonadotropin</td>
<td>1 (0.6%)</td>
<td>1 (1.2%)</td>
<td>2 (0.8%)</td>
</tr>
<tr>
<td>Dehydroepiandrosterone</td>
<td>1 (0.6%)</td>
<td>1 (1.2%)</td>
<td>2 (0.8%)</td>
</tr>
<tr>
<td>Insulin</td>
<td>1 (0.6%)</td>
<td>2 (2.4%)</td>
<td>3 (1.1%)</td>
</tr>
<tr>
<td>Antiestrogenic drugs</td>
<td>1 (0.6%)</td>
<td>3 (3.6%)</td>
<td>4 (1.5%)</td>
</tr>
<tr>
<td>Amphetamines</td>
<td>1 (0.6%)</td>
<td>4 (4.8%)</td>
<td>5 (1.9%)</td>
</tr>
<tr>
<td>Cocaine</td>
<td>1 (0.6%)</td>
<td>2 (2.4%)</td>
<td>3 (1.1%)</td>
</tr>
<tr>
<td>Gamma-hydroxybutyric acid (GHB)</td>
<td>2 (1.1%)</td>
<td>2 (2.4%)</td>
<td>4 (1.5%)</td>
</tr>
</tbody>
</table>

### 3.2 Construct validity

A principal component analysis was conducted using an Oblimin rotation as factors were expected to be correlated. Visual inspection of the scree plot (Figure 1) revealed a steep drop-off after the second component and a further drop-off after the fourth component. The first two robust factors identified had eigenvalues of 7.46 and 3.36, accounting for 37% and 16% off the variance in the data. The third and fourth factor had eigenvalues of 1.26 and 1.07 and together accounted for a further 11.6% of the total variance. The eigenvalues of the first two components were greater than their corresponding randomly generated data eigenvalues derived from the parallel analysis, supporting a two-component solution. Considering these results, a two-component solution was selected. Items 10-20 loaded on to Factor 1 as the Substance subscale, and items 1-7 loaded on to Factor 2 as the Supplements subscale. Item 9 did not load adequately on to either scale and was removed, after which a second analysis was conducted in which item 8 was also removed (see Table 3).
Figure 1. Scree plot of Eigenvalues from initial principal component analysis

Table 3

Factor loadings of two-component APES

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1 - Substances</th>
<th>Factor 2 - Supplements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Protein</td>
<td>.73</td>
<td></td>
</tr>
<tr>
<td>2. Creatine</td>
<td>.74</td>
<td></td>
</tr>
<tr>
<td>3. Glutamine</td>
<td>.70</td>
<td></td>
</tr>
<tr>
<td>4. BCAA and amino acids</td>
<td>.78</td>
<td></td>
</tr>
<tr>
<td>5. Caffeine</td>
<td>.66</td>
<td></td>
</tr>
<tr>
<td>6. Nitric Oxide</td>
<td>.62</td>
<td></td>
</tr>
<tr>
<td>7. Non-steroidal anti-inflammatories</td>
<td>.43</td>
<td></td>
</tr>
<tr>
<td>10. Diuretics</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>11. Anabolic-androgenic steroids</td>
<td>.67</td>
<td></td>
</tr>
<tr>
<td>12. Beta-2 agonists</td>
<td>.72</td>
<td></td>
</tr>
<tr>
<td>13. Human-growth hormones</td>
<td>.89</td>
<td></td>
</tr>
<tr>
<td>14. Human chorionic gonadotropin</td>
<td>.93</td>
<td></td>
</tr>
<tr>
<td>15. Dehydroepiandrosterone</td>
<td>.95</td>
<td></td>
</tr>
<tr>
<td>16. Insulin</td>
<td>.44</td>
<td></td>
</tr>
</tbody>
</table>
17. Antiestrogenic drugs          .89
18. Amphetamines               .74
19. Cocaine                     .90
20. Gamma-hydroxybutyric acid (GHB) .85

3.3 Internal consistency

Item-total correlations (see Table 4) showed strong internal consistency in the Supplements subscale of the APES (Cronbach’s α = .82), the Substances subscale of the APES (Cronbach’s α = .90), and the overall APES scale (Cronbach’s α = .78). The removal of the three lowest correlating items (5, 7, and 18) did not increase Cronbach’s alpha upon removal. They were therefore retained in the instrument.

Table 4

APES item reliability analysis

<table>
<thead>
<tr>
<th>Item</th>
<th>Item-Total Correlation</th>
<th>Cronbach’s α if deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Supplements subscale</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Protein</td>
<td>.60</td>
<td>.79</td>
</tr>
<tr>
<td>2. Creatine</td>
<td>.60</td>
<td>.79</td>
</tr>
<tr>
<td>3. Glutamine</td>
<td>.58</td>
<td>.80</td>
</tr>
<tr>
<td>4. BCAA and amino acids</td>
<td>.68</td>
<td>.78</td>
</tr>
<tr>
<td>5. Caffeine</td>
<td>.61</td>
<td>.79</td>
</tr>
<tr>
<td>6. Nitric Oxide</td>
<td>.56</td>
<td>.80</td>
</tr>
<tr>
<td>7. Non-steroidal anti-inflammatories</td>
<td>.39</td>
<td>.82</td>
</tr>
<tr>
<td><strong>Substances subscale</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Diuretics</td>
<td>.49</td>
<td>.91</td>
</tr>
<tr>
<td>11. Anabolic-androgenic steroids</td>
<td>.59</td>
<td>.90</td>
</tr>
<tr>
<td>12. Beta-2 agonists</td>
<td>.65</td>
<td>.90</td>
</tr>
<tr>
<td>13. Human-growth hormones</td>
<td>.84</td>
<td>.89</td>
</tr>
<tr>
<td>14. Human chorionic gonadotropin</td>
<td>.90</td>
<td>.90</td>
</tr>
<tr>
<td>15. Dehydroepiandrosterone</td>
<td>.92</td>
<td>.89</td>
</tr>
</tbody>
</table>
16. Insulin & .38 & .91 \\
17. Antiestrogenic drugs & .86 & .89 \\
18. Amphetamines & .70 & .89 \\
19. Cocaine & .84 & .89 \\
20. Gamma-hydroxybutyric acid (GHB) & .78 & .89 \\

3.4 Concurrent validity

As hypothesized, the APES demonstrated medium to strong correlations with each of the following scales: Drive for Muscularity – Behavior subscale ($r = .72, p < .001$), Drive for Muscularity – Attitude subscale ($r = .42, p < .001$), Eating for Muscularity Scale ($r = .46, p < .001$), and the Obligatory Exercise Questionnaire ($r = .53, p < .001$). This suggests that the APES is positively associated with both a drive for muscularity as well as engagement in other muscularity-enhancing behaviors.

3.5 Test-retest reliability

Of the original participant pool, 61 participants completed the APES at the second data point to determine test-retest reliability. Results indicated good reliability for the Supplements subscale ($r = .88, p < .001$) and excellent reliability for the Substances subscale and total APES score ($r = .99, p < .001; r = .93, p < .001$ respectively).

4. Discussion

The aim of this study was to develop and provide initial psychometric data for a survey tool measuring APED usage. The APES was developed based upon a thorough search of current literature on licit and illicit substances used for the purpose of enhancing appearance and performance. Following a principal components analysis, the initial 20 items in the tool were reduced to 18 items comprising two subscales: supplement use (7 items) and substance use (11 items). Testing in an at-risk community sample supported the internal consistency and temporal stability of the measure. Further, significant correlations at both the subscale and total score levels with other measures of attitudes and behaviors related to the pursuit of muscularity suggest strong concurrent validity of the measure.

NB: This study is currently under review. Please refer to the final published material when citing.
To date, research into APED use has relied upon semi-structured interviews, individual survey questions or measures only evaluating AAS use. It is likely that this inconsistency in measurement approaches as well as an absence of valid and reliable instruments has contributed to highly varied population estimates of APED use. Additionally, while semi-structured and clinical interview methods have demonstrated validity, these are costly and time-consuming to administer in large epidemiological samples. The current findings present a first step towards the development of a psychometrically supported survey tool for assessing the use of both appearance/performance enhancing supplements and substances.

Use of illicit substances to enhance appearance/performance in this sample (7%) was consistent with other studies that have estimated use between 3-15% of the population.\(^{12,38}\) By contrast, this study found that usage of APED supplements was relatively common with 70% of participants endorsing use of licit APED in the past three years. This was higher than previously found levels of supplement usage in surveys of other at-risk populations such as gym users and athletes where prevalence has typically ranged from 36-45\%.\(^{12,38,39}\) The current estimated prevalence is, however, similar to that reported by Hildebrandt and colleagues',\(^{11}\) in which 85% of the sample reported supplement use for the purposes of muscle-building and/or weight loss in a university sample. One key difference between other studies and both our study and that reported by Hildebrandt and colleagues\(^ {11}\) was the use by the former of a simple dichotomous question rather than measures comprising prompts or listed supplements for participants. This would fit with our earlier suggestion that overly broad clusters of APEDs or open ended-questions may yield underestimates of APED prevalence where individuals do not recognize a product as an APED.

Significantly, the high rates identified appear in the context of the well-documented and publicized adverse effects of both licit and illicit APED usage.\(^ {6,16-18}\) This is consistent with other risky health behaviors such as tobacco use, unprotected sexual intercourse, and alcohol use during pregnancy in which knowledge of associated risks has not predicted positive attitudes and behaviors.\(^ {40-42}\) The APES has the potential for use in research to further elucidate
etiological and risk factors for APED usage, the identification of which are key for the development of both treatment and prevention programs for APED use.

Additionally, the associations of APED usage with other muscularity-enhancing behaviors including obligatory exercise and muscularity-oriented disordered eating, suggests that further research is required to explore this cluster of maladaptive behaviors. Muscle dysmorphia, a recent addition to the Diagnostic and Statistical Manual of Mental Disorders,\(^4^3\) includes a preoccupation with one’s muscularity and dysfunctional or impairing behaviors to increase muscularity including excessive exercise, avoidance of physique exposure, use of ergogenic substances, and disordered eating.\(^4^4\) The APES may be a useful tool in assessing the levels of both licit and illicit APED usage in these populations given that previous research has focused almost exclusively on AAS usage in this group.\(^4^5\)

Despite the strengths of the current study, there were also several limitations. Firstly, the study employed a largely female, community sample, albeit using a sampling method targeting individuals interested in health/wellbeing, fitness, and bodybuilding. The gender imbalance, which may have arisen as a consequence of the use of snowball method of recruitment, raises questions about the relevance of the findings to males. Further research is required to evaluate the APES among men. While supplement use is common in community samples, it is expected that illicit APED usage will be lower in the community than at-risk and clinical groups as well as in female populations. These expectations were evident in the relatively skewed distribution in particular of the substances subscale. Future research should further explore the reliability and validity of the APES in other populations including clinical populations (such as individuals with muscle dysmorphia, anorexia and bulimia nervosa, as well as individuals with AAS-related substance abuse disorders) and at-risk populations such as amateur athletes and bodybuilders. A further limitation of this study was the large attrition between the test-retest stages. Approximately three-quarters of the participants opted out of or did not complete the survey measure at follow-up. This was likely due to the large battery of questionnaires required at time 1 and that the reward incentive was

\(NB: \) This study is currently under review. Please refer to the final published material when citing.
only relevant for people completing time 1 with no further incentive offered for the follow-up study. Nevertheless, the demographic and BMI characteristics of those completing the re-test were similar to those who completed only the first survey and at minimum the findings indicate that the measure has a high degree of stability over time among highly motivated participants. Further limitations of the current research include the lack of measures to assess divergent validity and the absence of empirical data on actual supplement and substance abuse against which to validate the measure. Future research is required to replicate and extend the psychometric investigations reported here to encompass additional validation studies and populations.

5. Perspectives

The current study provides preliminary data on the psychometric characteristics of the APES. The data is encouraging, suggesting that the tool warrants further research. The APES has the potential to be used to create a clearer picture of APED usage including an enhanced understanding of the etiology, prevalence, prognosis and treatment outcomes for those using APED. The tool may also support clinicians in gaining a clear understanding of the APED usage in patients presenting with musculature-oriented concerns.
References


*NB: This study is currently under review. Please refer to the final published material when citing.*


NB: This study is currently under review. Please refer to the final published material when citing.


*NB: This study is currently under review. Please refer to the final published material when citing.*


34. IBM Corp. IBM SPSS Statistics for Mac (Version 25.0) [Computer Software]. 2017. Armonk, NY: IBM Corp


NB: This study is currently under review. Please refer to the final published material when citing.


APPENDIX A

Please read the following questions carefully and indicate how much time you have used each of these supplements in the past 3 years for the purpose of body modification (e.g. weight loss, weight gain, muscle change, etc.) by putting a circle around the number on the following scale:

0 = Never
1 = A little (0-1 months)
2 = Moderate (1-6 months)
3 = A lot (6 months or more)

<table>
<thead>
<tr>
<th>Supplement</th>
<th>Never</th>
<th>A little</th>
<th>Moderate</th>
<th>A lot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein (includes protein powders, protein drinks, protein bars, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creatine (includes Amplified Creatine 189, Creatine Muscle Stack, Cell-Tech, MM1000 Creatine, Animal Max, BetaCene, Create, JetMass, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glutamine (includes Micronized Glutamine, Glutamine caps, Glutamine powder, GL-3, Engultarade, LG Pro, Magnum G, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCAAs and amino acids (includes L-Cartinine, Infinite Force, Universal Atomic, Xtend, Amino Energy, Atomic 7, Superior Amino, Recon, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caffeine (includes coffee, NoDoze, OxyELITE Pro, HydroxyStim, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitric Oxide (Jack3D, MM250 Pre-Workout, Juggernaut, ZizzZazz, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-steroidal anti-inflammatories (includes Aspirin, Ibuprofen, Aristocort, Aristospan, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Testosterone precursors or prohormones (includes Andro*GEN, Andro Stack, Androstat, THG methyl-1-testosterone, androstenedione, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ephedrine (includes Ephedra, Animal Cuts, Hydroxycut, Ripped Fuel, Cuts III, Diet Fuel, Dynabutus, Metabolife, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diuretics (includes Acetazolamide, Furosemide, Thiazide, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anabolic-androgenic steroids (includes Tetrahydrogestrinone, Norbolethone, Desoxymethyltestosterone, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta-2 agonists (includes Clenbuterol, Salbutamol, Terbutaline, Formoterol, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
13. Human-growth hormones 0 1 2 3
14. Human chorionic gonadotropin 0 1 2 3
15. Dehydroepiandrosterone 0 1 2 3
16. Insulin 0 1 2 3
   Antiestrogenic drugs (includes Tamoxefin, Aromasin/Exemestane, Femara/Letrozole, Abraxane etc.) 0 1 2 3
17. Amphetamines (includes Adderall, Dextedrine, Dextrostat, Desoxyn, ProCentra, Vyvanse, Speed, Ice, etc.) 0 1 2 3
18. Cocaine 0 1 2 3
20. Gamma-hydroxybutyric acid (GHB) 0 1 2 3
Other (please list): 0 1 2 3
Other (please list): 0 1 2 3
Other (please list): 0 1 2 3
Other (please list): 0 1 2 3

Supplemental table

Table 1
APES total score correlations with related measures

<table>
<thead>
<tr>
<th></th>
<th>APES-supp</th>
<th>APES-sub</th>
<th>APES-total</th>
<th>DMS-behav</th>
<th>DMS-att</th>
<th>EMS</th>
<th>OEQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>APES-supp</td>
<td>1</td>
<td>.23**</td>
<td>.95**</td>
<td>.73**</td>
<td>.46**</td>
<td>.45**</td>
<td>.53**</td>
</tr>
<tr>
<td>APES-sub</td>
<td>1</td>
<td>.53**</td>
<td>.21*</td>
<td>.17*</td>
<td>.14*</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>APES-total</td>
<td>1</td>
<td>.72**</td>
<td>.45**</td>
<td>.46**</td>
<td>.52**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DMS-behav</td>
<td>1</td>
<td>.63**</td>
<td>.43**</td>
<td>.62**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DMS-att</td>
<td>1</td>
<td>.37**</td>
<td>.39**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OEQ</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**p<.001, *p<.05 APES-supp = Appearance and Performance Enhancing Supplement/Substances Scale – Supplement subscale, APES-sub = Appearance and Performance Enhancing Supplement/Substances Scale – Substances subscale, APES-total = Appearance and Performance Enhancing Supplement/Substances Scale – total score, EMS = Eating for Muscularity Scale, DMS-behav = Drive for Muscularity Scale – Behavior subscale, DMS-att = Drive for Muscularity Scale – Attitudes subscale, and OEQ = Obligatory Exercise Questionnaire

NB: This study is currently under review. Please refer to the final published material when citing.
CHAPTER 4: Study 3

Foreword

As well as the use of ergogenic substances, the findings from Study 1 suggested that a more relevant consideration of eating pathology may be important in improving the classification and understanding of MD (Pope et al., 1997; Murray et al., 2012). This paper presents the development and psychometric evaluation of a second novel measure of MD symptomology, the Eating for Muscularity Scale (EMS).

The role of eating pathology in MD has been a contentious issue within the literature. Whilst Pope and colleagues’ (1997) initially described dieting behaviour as secondary and superfluous to the clinical picture of MD, it has also been suggested that eating pathology in MD has been underrecognised due to an emphasis on disordered eating behaviours promoting thinness, rather than muscularity (Murray et al., 2012).

Muscularity-oriented disordered eating has been described as pathological eating attitudes and behaviors driven by the pursuit of muscularity (Griffiths, Murray, & Touyz, 2013). This drive, and thus its related behaviours, differs from thinness-oriented disordered eating in that it aims to increase muscularity and often simultaneously decrease body fat. Muscularity-oriented disordered eating behaviors include overconsumption of protein-based foods, frequent eating (every 2 to 3 hours), restriction of fats and carbohydrate-based foods, liquefying/blending ingredients to increase caloric intake, and compensation-related to workout/gym regimes that lead to impairments in social, academic, or occupational functioning and/or negative physical and mental health outcomes (Griffiths et al., 2013; Murray, Griffiths & Mond, 2016). As traditional eating
pathology assessment tools have been developed within the framework of thinness-oriented body dissatisfaction, the relationship between muscularity-oriented eating pathology and MD is yet to be effectively explored.

In response, this chapter aimed to address the following research question:

RQ 3: Can we accurately measure symptoms of MD and, where required, develop psychometrically sound instruments to measure these symptoms?

This question will be addressed by the development and psychometric evaluation of the Eating for Muscul arity scale (EMS). The EMS was developed to assess muscularity-oriented disordered eating based on a rigorous a priori process prior to its evaluation in two community samples.
Getting shredded: Development and validation of a measure of muscularity-oriented disordered eating
By Marita Cooper and Kathleen M. Griffiths

Abstract

Objective: This study reports on the development and initial psychometric evaluation of the Eating for Muscularity Scale (EMS), a measure of muscularity-oriented disordered eating (MDE) attitudes and behaviors. Method: A literature review was undertaken to define the construct of MDE and relevant subdomains. A large pool of items was developed for these subdomains and then reduced based on feedback from subject-matter experts. A Principal Components Analysis (PCA) (n=265) and Principal Axis Factoring analysis (PAF) (n=204) were conducted to refine the measure and identify latent factors in two samples of participants recruited online. The factor structure was evaluated using Confirmatory Factor Analysis in a final group of participants (n=222). Preliminary psychometric evaluation of the EMS was conducted in the original sample (n=266). Results: PCA and a priori rules reduced items prior to undergoing PAF where a four-factor model was recommended. A CFA resulted in a 30-item measure supporting the four subscales of: Dieting Behaviors, Functional Impairment, Compensatory Behaviors and Negative Affect. The EMS demonstrated high internal consistency (Cronbach’s α = .96), moderate to strong concurrent validity with related measures (r = .43-.78), good construct validity using known-group assessment, and strong test-retest reliability (r = 0.92). Discussion: These findings provide preliminary support for the EMS as a reliable and valid measure of MDE. The instrument has the potential to support future research to improve clinical understanding and assessment of MDE and its etiology, prevalence and treatment outcomes. Further studies are required to evaluate the clinical and research utility of the scale.

NB: This study is currently under review. Please refer to the final published material when citing.
Introduction

Our society’s focus on increasing muscularity is neither new nor culturally bound to western society (Gray & Ginsburg, 2001). Rather, it has been recorded that men in medieval times would attempt to appear more muscular by stuffing their shirts with hay (Pope, Phillips, & Olivardia, 2000) and the pursuit of muscularity has been found to span Caucasian, Asian, African, and Hispanic populations (Jung, Forbes, & Chan, 2010; Hong & Ennis, 2014; Swami, 2016). Recently though, muscular body ideals are becoming unattainable through regular weight management strategies and, as a result, individuals are increasingly using unhealthy body-change behaviors including compulsive exercise, use of appearance- and performance-enhancing substances, as well as muscularity-oriented disordered eating to pursue mesomorphic ideals (Hildebrandt et al., 2011; McCabe & Ricciardelli, 2001; McCabe & Ricciardelli, 2003; Petrocelli, Oberweis, & Petrocelli, 2008; Pope & Katz, 1988). As these behaviors become more prevalent, there is a growing need for accurate assessment measures to increase our understanding of their etiology and prevalence, and support the development of efficacious interventions.

Muscularity-oriented disordered eating (MDE) has been described as pathological eating attitudes and behaviors driven by the pursuit of muscularity (Griffiths, Murray, & Touyz, 2013). While disordered eating has historically been conceptualized using the lens of the thin body ideal, MDE incorporates a cluster of behaviors motivated by the pursuit of a muscular body ideal. As the direction of actual to ideal body shifts from losing weight to gaining muscle, understandably the dieting strategies required to achieve this ideal also change. Consequently, MDE behaviors include overconsumption of protein-based foods, frequent eating (every 2 to 3 hours), restriction of fats and carbohydrate-based foods, liquefying/blending...
ingredients to increase caloric intake, and compensation-related to workout/gym regimes that lead to impairments in social, academic, or occupational functioning and/or negative physical and mental health outcomes (Griffiths et al., 2013; Murray, Griffiths & Mond, 2016).

Initial research exploring disordered eating pathology in individuals exhibiting a high drive for muscularity relied on traditional eating disorder measures such as the Eating Disorders Examination Questionnaire (EDE-Q, Fairburn & Beglin, 1994) or the Eating Attitudes Test (Garner, Olmstead, Bohr, & Garfinkel, 1982). However, these measures examine concerns and behaviours related to weight loss, shape concern, and caloric restriction, all of which characterize the drive for thinness rather than the drive for muscularity. In fact, such behaviors are typically incongruent with MDE as they would likely result in a decrease in muscularity. Consequently, this approach to measurement has likely led to an underestimation of MDE behaviors and, in particular, disordered eating symptoms in male populations who are typically more invested in muscular ideals as opposed to thinness (Murray, Griffiths, & Nagata, 2018).

More recently, there have been attempts to create or adapt standard disordered eating measures to better evaluate MDE, either in the form of individual items (Carter & Rudd, 2005; Eisenberg, Wall & Neumark-Sztainer, 2012) or MDE-related subscales (Forbush et al., 2013; Ricciardelli & McCabe, 2002; Stanford & Lemberg, 2012). However, these measures commonly limit their assessment of MDE to the use of supplements, preoccupation or dissatisfaction with muscularity, or the use of anabolic androgenic steroids. One exception is Murray and colleagues’ (Murray et al., 2012) adaptation of the EDE-Q to focus on muscularity concerns, which adapted items assessing fear of weight gain/fat to fear of weight loss/inadequate muscularity,
those asking about decreasing food amounts to increasing food amounts, and added preoccupations with protein consumption. Despite such changes, rudimentary revisions to standard eating disorder measures that attempt to replicate the original scale cannot capture the true scope of MDE behaviors. Consequently, although these instruments each represent constructive attempts to better understand MDE, they lack the richness of information that is gained from assessments into thinness-oriented disordered eating and fail to encompass the specific rituals and routines of those engaging in MDE. These limitations have likely contributed to the underestimation of the prevalence of MDE, hindering our understanding of such behaviors and further marginalizing those struggling with MDE and its related sequelae.

In summary, despite longstanding acknowledgement of the pursuit of muscularity and the use of dieting strategies to achieve greater muscularity, we are currently limited by inadequate assessment of MDE. Improved understanding of MDE is key to developing efficacious prevention and treatment protocols as well as improving our overall knowledge in this area. The purpose of the current study is to develop and assess the psychometric properties of a new measure of MDE, the Eating for Muscularity Scale (EMS). This paper reports the rigorous process employed to develop the EMS and establish its factor validity, followed by an examination of convergent validity and reliability in an at-risk community sample.

**Method**

*Scale development*

This section describes (i) the development of the item pool for the EMS; (ii) item reduction using Principal Components Analysis (PCA), factor extraction using Principal Axis Factoring (PAF); and (iii) the finalization of the EMS using a confirmatory factor analysis (CFA).

NB: This study is currently under review. Please refer to the final published material when citing.
(i) **Construct and Item Development**

The first author conducted a narrative review of the literature to develop an overarching definition of the construct “Muscularity-Oriented Disordered Eating.” Based on this review, MDE was defined as “eating and eating-related behaviors and attitudes aimed to increase muscularity and leanness. Engaging in these behaviors leads to distress, dysfunction, and/or negative health outcomes.” Subdomains of MDE were then defined through exploration of relevant models (Cafri et al., 2005; Pope et al., 1997; Ricciardelli & McCabe, 2004) as detailed in Table 1. A large item pool was generated to tap each of these subdomains. Sample items for each domain are listed in Table 1. Further items were adapted from the Eating Attitudes Test (Garner & Gardinkel, 1979), Eating Disorder Diagnostic Scale (Stice, Telch, & Rizvi, 2000), EDE-Q (Fairburn & Beglin, 2008), and Multidimensional Eating Disorder Inventory for Anorexia Nervosa and Bulimia (Garner, Olmstead, & Polivy, 1983) resulting in a total of 131 items.

<INSERT TABLE 1 HERE>

Content validity and initial item reduction was established through consultation with subject matter experts. A total of 12 experts were contacted via email and provided with a link to review the scale online via Google Forms. Five authors responded, with four completing the full-scale review and one providing general feedback. Experts were asked to review each item on a 5-point Likert scale (1-Poor, 5-Excellent) indicating the extent to which the item contributed to the validity of the proposed construct. Inter-rater reliability was assessed using Perreault and Leigh’s (1989, p141) variation of Cohen’s kappa. Evaluation of inter-rater reliability with different combinations of experts demonstrated that use of three judges yielded a fair score (IR = .35). Accordingly, items were retained when at least

*NB: This study is currently under review. Please refer to the final published material when citing.*
three out of four judges rated the item 4 or 5 (Very good to Excellent). Based on this algorithm, a total of 84 items were retained for pilot testing (comprising between 3 and 19 items for each domain, with one domain, Media Messages, completely removed). These items were then ordered by subdomain in an alternating format. The response format for the EMS replicated the 7-point Likert scale and scoring of the EDE-Q (Fairburn & Beglin, 2008, rating the number of days in the previous four weeks in which individuals have engaged in each item (0 = no days, 1= 1-5 days, 2 = 6-12 days, 3 = 13-15 days, 4 = 16-22 days, 5 = 23-27 days, 6 = every day). A final review of face validity was conducted to remove redundant items resulting in a final 79 items for analysis.

(ii) Item reduction and factor validity

Participants

Development and validation of the EMS was undertaken using two separate community-based samples: the Facebook sample and the MTurk sample. The samples were recruited using different approaches to ensure testing on a diverse pool of participants. As the levels of eating pathology were lower in the Facebook sample, (n=265) we conducted initial item reduction using PCA in this sample. The MTurk sample was then split into two using the SPSS random sampling option, producing MTurk Sample 1 (n = 204) data from which was employed in the PAF, and MTurk Sample 2 (n= 222) which provided the data for the CFA.

The Facebook sample. This group comprised 266 participants, aged between 18-73 years (M = 32.58 years, SD = 9.48 years). There were 85 males (32%) and 176 females (68%), with a mean Body Mass Index of 25.2 (SD=5.10) across the sample. Participants were recruited utilizing a snowball method via Facebook as part of a larger study into muscle dysmorphia symptomatology. Three key methods of

NB: This study is currently under review. Please refer to the final published material when citing.
recruitment were utilized: 1) a study Facebook group containing information about the study and an invitation to participate; 2) paid Facebook advertisements inviting individuals to participate targeting individuals interested in Wellbeing, Nutrition, Crossfit, Gym-users, and Bodybuilding (no country restriction); and 3) posts about the study on Facebook groups comprising potentially at-risk participants (such as Wellbeing, Nutrition, Gym interest groups) and by Lifestyle/Wellness professionals. We determined that this recruitment methodology was likely to incorporate both community and at-risk group of participants. All participants were offered the opportunity to enter a competition to win a $100 gift voucher for an online food delivery service. Participation was voluntary and confidential.

The MTurk sample 1. This group comprised 204 participants aged between 18 and 69 years ($M = 36.56$ years, $SD = 10.52$ years). There were 127 males (58%), 92 females (42%), with one individual identifying as transgender. Of the sample, 61% identified as Caucasian and 24% as Asian or South Asian. The mean Body Mass Index (BMI) of the sample was 24.9 ($SD=5.78$) and 89% identified as heterosexual.

The MTurk sample 2. This group comprised 222 participants aged between 19 and 69 years ($M = 38.11$ years, $SD = 10.82$ years). There were 119 males (59%), 84 females (41%), with 62% of the sample identifying as Caucasian and 24% as Asian or South Asian. The mean Body Mass Index (BMI) of the sample was 26.3 ($SD=5.55$) and 88% identified as heterosexual.

MTurk participants were recruited via Amazon Mechanical Turk, a crowd-sourced online labor market widely used by social scientists and psychology researchers (Chandler & Shapiro, 2016). There is evidence that psychometric properties of data generated by MTurk workers is of high quality and similar to that seen for other community populations, yielding high scale reliability and equivalent

NB: This study is currently under review. Please refer to the final published material when citing.
factor solutions (Chandler & Shapiro). We specifically targeted at-risk individuals engaging in frequent exercise (4 times per week or more). Participants were advised that participation was voluntary, although they were paid for their participation.

Participation in this study was anonymous and consent was requested at the beginning of the study.

**Procedure**

All participants completed the survey online via Google Forms. MTurk participants were advised that the completion time for the task was estimated at 8 minutes and participants received $1.50USD upon completing the survey. To check for adequate respondent attention, two screening questions were inserted into the battery (Berinsky, Margolis & Sances, 2014). These questions required participants to follow an instruction in responding such as “To demonstrate your attention please change the previous three responses to not applicable.” The study was approved by the Australian National University Human Research and Ethics Committee (Protocol Number: 2016/585).

**Statistical Analyses**

Data were exported as a CSV file directly from the Google Forms website. Analyses were performed using IBM SPSS Statistics version 20 (IBM Corporation, 2011). Ten percent of the MTurk sample (n=48) failed both attention-screening items and were removed from the analysis. Boxplots were scrutinized for outliers in age and BMI and outlying scores were replaced with scores two standard deviations above or below the mean. Independent t-tests were utilized evaluate differences in demographics (gender, age, ethnicity, BMI, and sexual orientation) between samples. While there were no significant differences found in demographics between the two
MTurk samples, the Facebook sample had significantly more females, greater Caucasian representation, and were younger than the overall MTurk sample.

Following assumption testing using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett’s Test of Sphericity, PCA with promax oblique rotation were conducted reduce poorly performing items; items that cross-loaded (greater than .32) and those with component loadings less than .4 were removed. A set of a priori rules for item-reduction was developed based on the scale development recommendations proposed by Clark and Watson (1995), Matsunaga (2010), and Netemeyer and colleagues (Netemeyer, Bearden & Sharma, 2003), as follows: high inter-item correlations ($r>.80$), low item-total correlations ($r<.40$), and low mean inter-item correlations ($r<.30$). Following this item reduction process, we undertook an exploratory factor analysis using PAF with a promax oblique rotation to examine the underlying structure of the EMS. The total number of components retained was determined using the Kaiser criterion, scree test (Cattell, 1966) and Horn’s parallel analysis (Horn, 1965). Overall fit of the model was evaluated with the standardized root mean residual (SRMR; Hu & Bentler, 1998), root mean square error of approximation (RMSEA; Steiger, 1990), and comparative fit index (CFI; Bentler, 1990). The cutoffs for each criteria were: CFI > 0.90 (Kline, 1998), RMSEA < 0.08 (Browne & Cudeck, 1993), and SRMR < 0.08 (Hu & Bentler, 1999). A final CFA was conducted in R Console (R Core Team, 2017) using the lavaan package (Rosseel, 2012). CFA using the Maximum Likelihood estimate examined the overall fit of the four-factor models of the EMS using the SRMR, RMSEA, and CFI as noted earlier.
Results

Item reduction: PCA – Facebook sample

The preliminary PCA was supported by excellent KMO measure of sampling adequacy and Bartlett’s Test of Sphericity (KMO=90; $\chi^2 (2415)= 14,729$, $p<.001$). Communalities of items ranged from .56 to .88. Items with loadings of less than .4 on their primary component (7 items) or items that loaded higher than .32 on a secondary factor (5 items) as well as items that did not load on to core components (5 items) were removed in iterative analyses. A final 8 items were retained for face validity and 4 items were removed that were not interpretable.

Item reduction results

We conducted further item reduction using the a priori strategy in both the Facebook sample and MTurk Sample 1. This led to the refinement of the EMS items as follows: high inter-item correlations (8 items removed), low item-total correlations (2 items removed), and low inter-item means (0 items removed). Two items were retained despite violation of these rules due to perceived strong face validity and highly skewed participant responses.

PAF: MTurk Sample 1

The remaining 38 items underwent a PAF to determine underlying factors in the MTurk Sample 1. The KMO measure of sampling adequacy and Bartlett’s Test of Sphericity were again excellent (KMO=92; $\chi^2 (703)= 5,306$, $p<.001$). Visual inspection of the scree plot (Figure 1) revealed a strong first two factors with a drop off after the third component. The first four factors identified had eigenvalues of 15.03, 3.66, 2.89, and 1.79 and accounted for 61.5% of the variance in the data. The eigenvalues of the first four components were greater than corresponding randomly generated data eigenvalues derived from the parallel analysis.

NB: This study is currently under review. Please refer to the final published material when citing.
Considering these results, a 4-component solution was selected and the PAF was re-run restrained to four factors. Items with limited interpretability (4 items), loadings of less than .4 on their primary factor (2 items), and items that loaded higher than .32 on a secondary factor were removed (4 items) in an iterative process. The final four-factor model resulted in 30-items demonstrating fair fit (CFI = 0.92, RMSEA = 0.08, SRMR = .07). The four factors were defined as follows: Muscularity-oriented Dieting (MD), Functional Impairment (IMP), Compensatory Behaviors (COMP) and Negative Affect (NA).

**CFA: MTurk Sample 2**

The four-factor model demonstrated adequate fit within this sample (CFI = 0.87, RMSEA = 0.09, SRMR = 0.07). Table 2 shows the factor loadings of each item.

<INSERT TABLE 2 HERE>

**Psychometric assessment of the EMS**

Based on the promising findings discussed, we sought to further explore the psychometric properties of the EMS including the assessment of: concurrent validity, construct validity, internal consistency, and test-retest reliability.

**Method**

**Participants**

The Facebook sample (n=265), including a subset of participants followed up at 2 weeks, were utilized to further assess the psychometric properties of the EMS. The follow-up group comprised 61 participants, aged between (M= 31.14, SD= 7.62) of whom 18 (30%) were males and 42 (70%) were females (gender status was missing for one participant). The BMI of the follow-up sample was 25.0 (SD= 4.70). There were no significant differences in age, gender or BMI between the participants.
who participated in the follow-up survey and those who completed the first survey only.

**Eating for muscularity**

Muscularity-oriented disordered eating attitudes and behaviors were measured utilizing the 30-item EMS described previously.

**Measures for assessing construct and concurrent validity**

**Drive for muscularity.** The Drive for Muscularity (DMS; McCreary, 2007) is a 15-item measure designed to assess attitudes and behaviors related to the pursuit of muscularity. It has demonstrated strong internal consistency and test-retest reliability (McPherson, McCarthy, McCreary, & McMillan, 2010). A score above 31 on the DMS has previously been used as a cutoff to differentiate individuals exhibiting heightened symptoms of muscle dysmorphia with healthy individuals (Maida & Armstrong, 2005). The DMS possessed excellent internal consistency in both male (Cronbach’s α = .91) and female participants (Cronbach’s α = .90) in our sample.

**Body dissatisfaction.** The Male Body Attitudes Scale (MBAS; Tylka, Bergeron, & Schwatrz, 2005) – Muscle Dissatisfaction subscale was utilized to evaluate muscle dissatisfaction in the sample. The 24-item overall measure employs a 6-point Likert scale ranging from 1 representing ‘Never’ to 6 representing ‘Always.’ Constructed to evaluate aspects of body image preoccupation and dissatisfaction, the MBAS comprises three subscales: Muscularity, Body Fat and Height. The MBAS possesses high internal consistency and strong convergent validity with other measures of body esteem and attitudes towards muscularity (Tylka et al., 2005). The MBAS demonstrated excellent internal consistency in both male (Cronbach’s α = .94) and female participants (Cronbach’s α = .93) in our sample.
Eating disorder symptomology. The EDE-Q (Fairburn & Beglin, 1994) was used to assess attitudes and behaviors underpinning anorexia nervosa and bulimia nervosa. It possesses good to excellent test-retest reliability, good criterion validity and concurrent validity (Luce & Crowther, 1999; Mond, Hay, Rodgers, Owen, & Beumont, 2004). The EDE-Q demonstrated excellent internal consistency in both male (Cronbach’s $\alpha = .95$) and female participants (Cronbach’s $\alpha = .96$) in our sample.

Eating disorder-related impairment. Impairment related to eating disorder symptomology was measured using the 16-item Clinical Impairment Assessment (CIA; Bohn & Fairburn, 2008). The self-report measure evaluates an individual’s level of distress and impairment resulting from their weight/shape concern and behaviors. The CIA possesses excellent internal consistency and good convergent validity with the EDE-Q (Reas, Rø, Kapstad, & Lask, 2010). The CIA demonstrated strong internal consistency in both male (Cronbach’s $\alpha = .97$) and female participants (Cronbach’s $\alpha = .96$) in our study.

Procedure

The Facebook sample completed the survey online via Google Forms as previously described. At the completion of the battery, participants were able to “opt-out” of receiving the follow-up questionnaire. Those who remained in the study were sent the follow-up EMS questionnaire via email two-weeks following the submission of their initial survey.

Statistical Analyses

Mean total and sub-scale EMS scores were computed by summing the ratings across items and dividing by the number of items in the scale or sub-scale. Internal consistency was measured using Cronbach’s alpha coefficient. Concurrent validity,
interfactor correlations and test-retest reliability were each assessed using Pearson’s r correlations. Group differences in subscales scores as well as construct validity with known-groups were assessed using independent t-tests.

Results

Gender effects

Table 3 shows the EMS subscale and total scale scores as a function of gender. Independent t-tests found that females reported higher levels of negative affect ($t(259) = -3.66, p < .001$) and impairment ($t(254) = -2.41, p = .02$); however, there were no other significant gender differences at the EMS subscale or total scale level.

<INSERT TABLE 3 HERE>

Interfactor correlations

All interfactor correlations were in the moderate ($r = .55$) to high ($r = .71$) range. This suggests that subscales are sufficiently independent whilst still likely conforming to a higher order structure (Kline, 2005).

Internal consistency

The EMS demonstrated high internal consistency (Cronbach’s $\alpha = .96$). The subscales also possessed very good to excellent internal consistency as follows: IMP (Cronbach’s $\alpha = .92$), MD (Cronbach’s $\alpha = .92$), NA (Cronbach’s $\alpha = .86$), and COMP (Cronbach’s $\alpha = .95$). All items demonstrated good item-total correlations ($>.60$) and any improvements in reliability following item removal were marginal.

Concurrent validity

Results of correlational analyses can be seen in Tables 4 and 5. The EMS total scale demonstrated moderate to large correlations with each of the related measures suggesting strong concurrent validity. At the subscale level, all relationships were significant. The MBAS muscularity subscale demonstrated weaker, albeit significant,
relationships with the IMP and COMP subscales. The CIA and EDE-Q were found to have strongest relationships with the NA and IMP subscales, although all relationships involving these scales were significant.

<INSERT TABLE 4 HERE>

<INSERT TABLE 5 HERE>

**Construct validity**

Known-validity was assessed using a recommended cutoff of scores greater than 31 on the DMS. Means for each group are shown in Table 3. Results showed significantly higher EMS scores in the group of individuals exhibiting heightened muscle dysmorphia symptomology ($t(242) = -4.73, p < .001$) representing a medium effect size (Hedges’ $g = 0.61$).

**Test-retest reliability**

The EMS demonstrated strong test-retest reliability ($r = .92$). The subscales also possessed very good to excellent temporal stability as follows: IMP ($r = .92$), MD ($r = .89$), NA ($r = .84$), and COMP ($r = .89$).

**Discussion**

This study reported on the development and initial psychometric evaluation of a measure of musculature-oriented disordered eating attitudes and behaviors. To date, assessment of MDE has been hindered by the use of instruments designed to measure attitudes and behaviors motivated by the drive for thinness as well as those with limited exploration of the broad scope of MDE-specific rituals and regimes. To our knowledge, the EMS is the first instrument specifically developed to measure the broad-spectrum of MDE attitudes and behaviors with the equivalent richness of information that is incorporated in traditional eating disorder measures.

*NB: This study is currently under review. Please refer to the final published material when citing.*
The EMS was developed through a rigorous process resulting in a 30-item instrument comprising four subscales. The content validity of the EMS was supported by the initial expert review whilst its factor validity was demonstrated by confirmatory factor analysis. The EMS exhibited moderate-strong concurrent validity with the muscle dissatisfaction and traditional measures of disordered eating and related impairment. Furthermore, examination of the construct validity of the EMS found a moderate effect in differentiating individuals with and without heightened muscle dysmorphia symptomatology. Analyses also demonstrated that the EMS possesses very good to excellent internal consistency and test-retest reliability at both the total and subscale level. Finally, gender discrepancies were observed at the subscale-level only, indicating the EMS may provide a less gender-biased assessment of eating concerns than traditional eating disorder measures.

Our analysis revealed a four-factor model of the EMS, comprising the subscales of Functional Impairment, Muscularity-oriented Dieting, Compensatory Behavior, and Negative Affect. Importantly, each of these factors has been previously described in models of MDE (Cafri et al., 2005; Pope et al., 1997; Ricciardelli & McCabe, 2004) providing support for the interpretation of the EMS factors. The findings also support the need for an instrument such as the EMS that assesses the multifaceted construct of MDE, as opposed to the individual item measures that have previously been used in the field (Carter & Rudd, 2005; Eisenberg et al., 2012).

Eating disorder measures have historically focused on thinness-oriented eating pathology developed and validated in female populations (Murray, Griffiths & Nagata, 2018). This sampling bias has hindered our understanding of the diverse forms of body dissatisfaction that can underpin eating pathology, especially in relation to MDE and muscularity-oriented dissatisfaction. The EMS will require
further psychometric evaluation including in a clinical sample (e.g. individuals with muscle dysmorphia). However, this study has supported the potential of the EMS as a relatively brief and easy to complete measure for use in population-level epidemiological research as well as improved detection of MDE in a clinical setting. These studies can hopefully lead to improved understanding of the prevalence and etiology of MDE as well as the development of efficacious treatments for MDE.

There are several limitations to this study. Firstly, it was conducted in community samples, albeit with both groups at least partially from a selected at-risk community samples. Thus, it is currently unclear how the EMS will perform with individuals in a clinical population. While it is promising that the EMS was able to differentiate individuals with heightened muscle dysmorphia symptomatology, further research is required to evaluate the diagnostic validity of EMS subscales in differentiating clinically presenting individuals with MDE from healthy controls. Future studies should also examine the temporal stability of the EMS over a longer period (6-12 months) to ensure the suitability of the EMS for assessing symptomatic change. Finally, the RMSEA fit indices of 0.8 and 0.9 raised a question regarding the fit of the model. However, it is difficult to achieve low RMSEA values with a participant to variable ratio of the magnitude seen in this study. Further, the indices did not exceed 0.1 above which Brown and Cudeck (1993) suggest that no RMSEA is acceptable and the remaining fit indices were satisfactory.

Additionally, although the moderate to high concurrent validity with other measures is a strength of the EMS, it is important to further examine the nature of this relationship. An association between standard eating disorder measures, such as the EDE-Q and CIA, and the EMS was predicted considering both the conceptual overlap between MDE and eating disorders (the preoccupation with body dissatisfaction and

*NB: This study is currently under review. Please refer to the final published material when citing.*
subsequent maladaptive eating/compensatory strategies) and some behavioural similarities between the two disorder types (rigid eating practices, excessive attention to diet, functional impairment, negative affect related to eating). However, future studies will be required to determine if the EMS adequately discriminates between MDE and thinness-oriented eating pathology.

The gender representation in this study was both a strength and limitation of this paper. It has been reported that less than 1% of all research on anorexia nervosa has been conducted in male populations (Murray et al., 2016). The development of measures and norms in overly homogenous samples can compromise the psychometric performance of these measures in diverse populations and, thus, the suitability of these measures in assessing more diverse body concerns (Darcy & Lin, 2012; Jennings & Phillips, 2017). Thus, the gender-diversity within the samples in both our studies was a strength. However, the number of female participants in the Facebook sample was still twice that of male participants, a less than optimal outcome given that MDE is often more common in male populations (Eisenberg et al., 2012; Heywood & McCabe, 2006).

Despite its limitations, there are several strengths to this paper. The EMS was developed based on a research-driven definition of the MDE construct with individual items generated through theoretical models of MDE. This process of development in conjunction with the engagement of subject-matter expert reviewers to assess the content validity of the items and the systematic assessment of the factor, construct and concurrent validity, as well as reliability of the EMS constituted a robust process for building a clinically and research-backed instrument for assessing MDE.

In conclusion, this study found preliminary support for the EMS as a reliable and valid measure of muscularity-oriented disordered eating pathology. Although
further evaluation of the measure is needed assess its suitability in clinical populations, the EMS provides a promising avenue for assessing MDE. Such research will support the clinical understanding and assessment of MDE as well as research into its prevalence, etiology, treatment and prognosis.

Acknowledgements
Thank you to Dr Liz Reiger for your support and advice with this paper and measure development. Thank you to the following experts who helped in the refinement of this tool: Dr Timothy Baghurst, Dr Scott Griffiths, Dr Deborah Mitchison, Dr Roberto Olivardia, and Mr David Wiss. Additionally, thank you to Mrs Kate Pollard Gough who supported the development of the measure and the collection of data.
References


*NB: This study is currently under review. Please refer to the final published material when citing.*
Perspectives. Arlington, VA: American Psychological Association


NB: This study is currently under review. Please refer to the final published material when citing.


NB: This study is currently under review. Please refer to the final published material when citing.


*NB: This study is currently under review. Please refer to the final published material when citing.*
Muscle dysmorphia: An under-recognized form of body dysmorphic disorder. 

*Psychosomatics, 38*. 548-557.


Retrieved from http://search.proquest.com/docview/220467863?

accountid=2650


*NB: This study is currently under review. Please refer to the final published material when citing.*


*NB: This study is currently under review. Please refer to the final published material when citing.*
# TABLES AND FIGURES

**Table 1**

*Subdomains of MDE and sample items of the EMS*

<table>
<thead>
<tr>
<th>Domain</th>
<th>Sample items</th>
</tr>
</thead>
</table>
| Preoccupation with leanness and muscularity | • I give a lot of time and thought to becoming more muscular  
• Thinking about my body shape and/or size made it very difficult for me to concentrate on things I am interested in  
• I give a lot of time and thought to becoming more lean |
| Dieting to gain weight/increase muscularity | • I must eat mostly protein-based foods (e.g. red meat, fish, chicken, etc.)  
• I have liquefied/blended foods to consume more kilojoules/calories  
• I have forced myself to eat even if I am full |
| Dieting for leanness                        | • I have eaten a low fat diet  
• I have fasted (skipped at least two meals in a row) to reduce body fat  
• I must eat low carbohydrate foods (e.g. eggs, meats, leafy greens, etc.) |
| Dietary restraint                           | • I have specifically avoided processed foods  
• I must restrict high carbohydrate foods (e.g. bread, pasta etc.)  
• I have deliberately limited the amount of ‘junk food’ I have eaten |
| Excessive attention to diet                 | • It’s important for me to be aware of how many kilojoules/calories are in the food I am eating  
• I have weighed ingredients when preparing meals and snacks  
• I give a lot of time and thought to eating and food preparation |
| Functional impairment                       | • My eating regime interferes with my work/studies  
• I have give up important social events to ensure I am eating every 2-3 hours  
• My eating regime interferes with my hobbies and leisure time |
<table>
<thead>
<tr>
<th>Health risk behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td>• I have used dieting methods that I know are unhealthy</td>
</tr>
<tr>
<td>• I have used appearance/performance-enhancing substances (e.g. anabolic androgenous</td>
</tr>
<tr>
<td>steroids, human growth hormones, etc.) that I know are unhealthy</td>
</tr>
<tr>
<td>• I have used nutritional supplements (e.g. protein, l-carnitine, caffeine, etc.)</td>
</tr>
<tr>
<td>that have caused me physical harm/damage</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Messages from the media, peers, family</td>
</tr>
<tr>
<td>• I read health/fitness magazines or websites to find dieting methods to become more</td>
</tr>
<tr>
<td>muscular</td>
</tr>
<tr>
<td>• I have used a diet method I was advised by a fitness instructor/personal trainer/</td>
</tr>
<tr>
<td>coach</td>
</tr>
<tr>
<td>• I have used a diet method I read about online</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Binge eating and compensation</td>
</tr>
<tr>
<td>• I have made myself vomit to reduce body fat</td>
</tr>
<tr>
<td>• I have had periods of eating where I felt I may not be able to stop</td>
</tr>
<tr>
<td>• I add extra exercise into my normal regime if I have eaten too much 'junk food'</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Negative affect</td>
</tr>
<tr>
<td>• I felt anxious if I was not able to prepare my meals and snacks in advance</td>
</tr>
<tr>
<td>• I feel guilty when I eat 'bad/junk foods'</td>
</tr>
<tr>
<td>• I feel uncomfortable eating with family or friends where I do not know the nutritional content of food I am consuming</td>
</tr>
</tbody>
</table>
Figure 1. Scree plot of Eigenvalues from initial principal axis factoring in MTurk sample 1

<table>
<thead>
<tr>
<th>Item</th>
<th>IMP</th>
<th>MD</th>
<th>COMP</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have used dieting methods that have caused me psychological harm/damage</td>
<td>.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have given up important social events to maintain my diet schedule</td>
<td>.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have given up important recreational activities to maintain my diet schedule</td>
<td>.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have given up important work commitments to maintain my diet schedule</td>
<td>.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My eating regime interferes with my hobbies and leisure time</td>
<td>.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My eating regime interferes with my work/studies</td>
<td>.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My eating regime interferes with my social life</td>
<td>.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thinking about increasing my muscularity makes it very difficult for me to concentrate on things I am interested in</td>
<td>.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have used dieting methods that I know are unhealthy</td>
<td>.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel that food controls my life</td>
<td>.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have deliberately limited the amount of 'junk food' I have eaten</td>
<td>.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have attempted to follow definite rules about my eating</td>
<td>.79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2

Factor loadings of EMS items

NB: This study is currently under review. Please refer to the final published material when citing.
I have specifically avoided foods that are high in sugar \[0.78\]
I must eat low carbohydrate foods (e.g. eggs, meats, leafy greens, etc.) \[0.76\]
I have strictly adhered to a rigid eating plan \[0.75\]
I have deliberately increased the amount of protein I have eaten \[0.67\]
I must restrict high fat foods (e.g. fried foods, oils, chocolate, etc.) \[0.65\]
I must eat mostly protein-based foods (e.g. red meat, fish, chicken, etc.) \[0.65\]
I have eaten more for functionality than taste \[0.61\]
I have prepared my meals and snacks in advance \[0.57\]
I have tracked/recorded my daily kilojoule/calorie intake \[0.47\]
I add extra exercise into my normal regime if I feel guilty about what I have eaten \[0.89\]
I add extra exercise into my normal regime if I have eaten 'too much' \[0.87\]
I add extra exercise into my normal regime if I have eaten too many carbohydrates \[0.85\]
I add extra exercise into my normal regime if I have eaten too many high fat foods \[0.85\]
I will increase my exercise if I break a dietary rule \[0.77\]
I get anxious or upset when I eat 'bad/junk foods' \[0.70\]
I feel anxious when I do not know the nutritional content of food I am consuming \[0.59\]
I feel uncomfortable when I eat 'bad/junk foods' \[0.55\]
I would be stressed if I was not able to prepare my meals and snacks in advance \[0.44\]

NB: This study is currently under review. Please refer to the final published material when citing.
Table 3
EMS subscale and total score means and standard deviations by gender

<table>
<thead>
<tr>
<th>Gender</th>
<th>IMP</th>
<th>MD</th>
<th>COMP</th>
<th>NA</th>
<th>EMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female (n=176)</td>
<td>0.70*</td>
<td>1.78</td>
<td>1.36</td>
<td>1.94**</td>
<td>1.33</td>
</tr>
<tr>
<td></td>
<td>(1.10)</td>
<td>(1.58)</td>
<td>(1.49)</td>
<td>(1.85)</td>
<td>(1.25)</td>
</tr>
<tr>
<td>Male (n=85)</td>
<td>0.41*</td>
<td>1.80</td>
<td>1.08</td>
<td>1.19**</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>(0.85)</td>
<td>(1.38)</td>
<td>(1.14)</td>
<td>(1.29)</td>
<td>(1.00)</td>
</tr>
<tr>
<td>Total (n=261)</td>
<td>0.63</td>
<td>1.80</td>
<td>1.28</td>
<td>1.72</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>(1.06)</td>
<td>(1.51)</td>
<td>(1.39)</td>
<td>(1.72)</td>
<td>(1.17)</td>
</tr>
<tr>
<td>Low MD group (n=140)</td>
<td>0.52</td>
<td>1.25</td>
<td>1.07</td>
<td>1.32</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>(1.09)</td>
<td>(1.38)</td>
<td>(1.34)</td>
<td>(1.68)</td>
<td>(1.18)</td>
</tr>
<tr>
<td>High MD group (n=104)</td>
<td>0.72</td>
<td>2.48**</td>
<td>1.50*</td>
<td>2.17**</td>
<td>1.64**</td>
</tr>
<tr>
<td></td>
<td>(1.04)</td>
<td>(1.40)</td>
<td>(1.36)</td>
<td>(1.66)</td>
<td>(1.03)</td>
</tr>
</tbody>
</table>

*p<.05 **p<.001. IMP = Functional Impairment, MD = Muscularity-oriented Dieting, COMP = Compensatory Behaviors, NA = Negative Affect, and EMS = Eating for Muscularity Scale.

Table 4
EMS total score correlations with related measures

<table>
<thead>
<tr>
<th></th>
<th>EMS</th>
<th>CIA</th>
<th>EDE-Q</th>
<th>MBAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMS</td>
<td>1</td>
<td>.72**</td>
<td>.78**</td>
<td>.42**</td>
</tr>
<tr>
<td>CIA</td>
<td>1</td>
<td>.87**</td>
<td>.29**</td>
<td></td>
</tr>
<tr>
<td>EDE-Q</td>
<td></td>
<td></td>
<td>.30**</td>
<td></td>
</tr>
<tr>
<td>MBAS</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

**p<.001 EMS = Eating for Muscularity Scale, CIA = Clinical Impairment Assessment, EDE-Q = Eating Disorder Examination-Questionnaire (EDE-Q), and MBAS - Male Body Attitudes Scale – Muscularity Dissatisfaction Subscale.

Table 5
EMS subscale score correlations with related measures

<table>
<thead>
<tr>
<th></th>
<th>IMP</th>
<th>MD</th>
<th>COMP</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMS</td>
<td>.83**</td>
<td>.88**</td>
<td>.82**</td>
<td>.88**</td>
</tr>
<tr>
<td>CIA</td>
<td>.81**</td>
<td>.49**</td>
<td>.56**</td>
<td>.67**</td>
</tr>
<tr>
<td>EDE-Q</td>
<td>.77**</td>
<td>.61**</td>
<td>.60**</td>
<td>.72**</td>
</tr>
<tr>
<td>MBAS</td>
<td>.25**</td>
<td>.42**</td>
<td>.25**</td>
<td>.36**</td>
</tr>
</tbody>
</table>

*p<.01 **p<.001. IMP = Functional Impairment, MD = Muscularity-oriented Dieting, COMP = Compensatory Behaviors, NA = Negative Affect, and EMS = Eating for Muscularity Scale; EMS = Eating for Muscularity Scale, MBAS – Male Body Attitudes Scale – Muscularity Dissatisfaction Subscale, CIA = Clinical Impairment Assessment, and EDE-Q = Eating Disorder Examination-Questionnaire (EDE-Q).

NB: Assumption testing was undertaken for all analyses. As well as those already included in text, all assumptions were met or parametric testing was still deemed viable due to large sample sizes.
CHAPTER 5: Study 4

Foreword

Body image disturbance is well established as a core symptom and aetiological mechanism in the development and maintenance of both body dysmorphic disorder (BDD) and traditional eating disorders (ED), such as anorexia and bulimia nervosa. Considering the current classification of muscle dysmorphia (MD) as a specifier of BDD, it is implicit in this diagnosis that MD is also driven by body image disturbance and its subcomponents. Whilst the role of body image dissatisfaction in MD is well documented (Diehl & Baghurst, 2016; Olivardia, Pope, & Hudson, 2000), evidence for the presence of body image distortion in MD has been restricted to case studies and insufficiently powered studies (Hernandez-Martinez, Gonzalez-Marti, & Jordan, 2017; Pope et al., 1993; Pope et al., 1997). Accordingly, Study 4 aimed to address the final research question:

RQ 4: Can we identify further factors to differentiate MD from similar psychiatric diagnoses?

As Study 1 highlighted specific concerns regarding the current diagnostic criteria for MD, Study 4 utilised a bottom up approach to understanding MD by exploring behavioural symptomology. For this purpose, we assessed the relative contributions of body image distortion and body image dissatisfaction in predicting attitudes related to the drive for muscularity as well as MD behaviours including obligatory exercise, muscularity-oriented disordered eating (using the measure developed in Study 3), and use of appearance and

NB: This study is currently under review. Please refer to the final published material when citing.
performance-enhancing supplements and substances (using the measure developed in Study 2) in a community sample.
Bodybuilding to the max: Is our pursuit of muscularity driven by body image dissatisfaction or distortion?  
Marita Cooper & Kathleen M. Griffiths

Abstract

Muscle dysmorphia (MD), a specifier of Body Dysmorphic Disorder, is characterized by body image distortion involving an imagined or grossly exaggerated perception of insufficient muscularity. However, research to date understanding body image distortion in MD has relied on case studies and insufficiently powered studies to support this. This study explored the relative contributions of body image dissatisfaction and body image distortion to engagement in MD behaviors. We recruited a community sample of 265 participants online through Facebook. A multivariate multiple regression analysis found that body image dissatisfaction, but not body image distortion, was associated with increased engagement with risky muscle enhancing strategies, uniquely predicting 80% of the variation in the model. Whilst the role of body image dissatisfaction in MD behaviors has been established previously, our results contradict the assumption that perceptual disturbances drive MD symptomology. Additional research is required to further delineate MD from psychiatric diagnoses.

Highlights

- Two key components of body image (BI) are BI distortion and BI dissatisfaction
- It has been assumed that muscle dysmorphia (MD) is driven by muscularity-related distortion
- Although evidence has linked MD with BI dissatisfaction, little research has explored BI distortion
- The current study found that BI dissatisfaction but not distortion was associated with MD behaviors
- Future studies should explore the accuracy of BI distortion in clinical MD samples

NB: This study is currently under review. Please refer to the final published material when citing.
1. Introduction

Muscle dysmorphia (MD), a specifier of Body Dysmorphic Disorder (BDD), is characterized by a preoccupation with insufficient muscularity. Despite its inclusion in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013), researchers continue to disagree over the clinical validity, reliability and diagnostic distinctiveness of MD from other psychiatric diagnoses (Baghurst, et al., 2014; Chung, 2001; Murray & Baghurst, 2013; Foster, Shorter, & Griffiths, 2013). This disagreement is further fueled by recent meta-analyses demonstrating insufficient evidence for the reliability and validity of MD criteria (dos Santos Filho et al., 2016; Nieuwoudt et al., 2016). Nevertheless, there is support for the assertion that those with muscularity concerns represent a population that merits further investigation. In a systematic review of MD research based on Robins and Guze’s (1970) method for evaluating psychiatric diagnoses, we conducted a meta-analysis of studies that compared individuals with MD to healthy controls and frequent exercisers (Cooper et al., in preparation). We found that individuals with MD represent a clinically relevant population, highlighting the importance of developing a clear clinical picture and etiological understanding of MD that can inform future research.

MD was first identified by Pope and colleagues (Pope, Katz, & Hudson, 1993) in a sample of male bodybuilders. Although originally likened to anorexia nervosa, the authors later reconceptualized MD as a form of BDD (Pope, Gruber, Choi, Olivardia, & Phillips, 1997). Based upon their preliminary observations of 81 weightlifters exhibiting this preoccupation with insufficient muscularity, Pope and colleagues (1997) developed the tentative criteria for MD outlined in Box 1. Although

NB: This study is currently under review. Please refer to the final published material when citing.
there has been widespread adoption of these criteria among researchers and clinicians, others have critiqued the usage of the criteria due to their development within a small and unrepresentative sample and a lack of field trials. There have also been concerns that the Pope criteria are overly subjective and open to diverse clinical interpretation (Niuwoudt, Zhou, Coutts & Booker, 2012).

Subsequent to Pope and colleagues’ research, the American Psychiatric Association (2013) officially recognized MD as a specifier of BDD, the latter being classified as a condition in the category of Obsessive-Compulsive and Related Disorders. These formal criteria, represented in Box 2, require that an individual meet all criteria for BDD and, that for MD to be diagnosed, the predominant preoccupation must be related to insufficient muscularity. However, a recent systematic review found that no MD research had utilized the manual’s criteria in the 4 years following its release in 2013 (Sandgren & Lavalee, 2018). Furthermore, recent literature indicates that MD may be more accurately classified as a body image or eating disorder rather than a Body Dysmorphic Disorder (Cooper et al., in preparation; Murray et al., 2012).

A key component of both BDD and traditional eating disorders – anorexia and bulimia nervosa – is the construct of body image disturbance. This disturbance may manifest as: body image distortion, a perceptual disturbance where an individual perceives aspects of their body as different from or exaggerated relative to the objective appearance of their body; body image dissatisfaction, whereby individuals are dissatisfied with the perception of their body or; excessive body image investment, in which an individual’s physique unduly influences their self-worth or

NB: This study is currently under review. Please refer to the final published material when citing.
self-evaluation (Cash & Pruzinsky, 2002). It is implicit in the description of MD, a specifier of BDD, that MD behaviors are driven by a body image distortion, namely a preoccupation with an imagined or grossly exaggerated perception of insufficient musculature. However, research to date has relied on case studies and insufficiently powered studies to support this suggestion (Hernandez-Martinez, Gonzalez-Marti & Jordan, 2017; Pope et al., 1993; Pope et al., 1997).

Understanding the individual roles of body image dissatisfaction and distortion in driving muscularity-oriented body change behaviors is integral to our understanding of MD. This information will contribute not only to the development of accurate criteria for MD but also, consequently, to the development of effective treatment and prevention protocols. The current study will explore the relative contributions of body image dissatisfaction and body image distortion to engagement in muscle enhancing behaviors. As the validity of the current criteria for diagnosing MD has been questioned, in this study we have employed a series of MD symptom/behavior assessment scales rather than a categorical diagnosis based on potentially flawed diagnostic criteria. Based on previous research, we hypothesize that body image dissatisfaction will predict MD behavior. There is an assumption in the literature that MD is associated with a body image distortion related to musculature. However, in the absence of previous research on the role of body image disturbance in muscle enhancing behaviors, we offer no hypothesis concerning its contribution or relative contribution to MD.
2. Materials and method

2.1. Participants

The sample comprised 265 participants, aged between 18-73 years (M = 32.58 years, SD = 9.48 years). There were 85 males (32%) and 176 females (68%), with a mean Body Mass Index (BMI) of 25.2 (SD=5.10) across the sample.

Participants were recruited utilizing a snowball method via Facebook as part of a larger study into muscle dysmorphia symptomatology. Three key methods of recruitment were utilized: 1) a study Facebook group containing information about the study and an invitation to participate; 2) paid Facebook advertisements inviting individuals to participate targeting individuals interested in Wellbeing, Nutrition, Crossfit, Gym-users, and Bodybuilding (no country restriction); and 3) posts about the study on Facebook groups comprising potentially at-risk participants (such as Wellbeing, Nutrition, Gym interest groups) and by Lifestyle/Wellness professionals. All participants were offered the opportunity to enter a competition to win a $100 gift voucher for an online food delivery service.

Participation was voluntary and confidential. Consent was requested at the beginning of the study. The study was approved by the Australian National University Human Research and Ethics Committee (Protocol No.: 2016/585).

2.2. Measures

Participants completed the following battery as well as other questionnaires as part of a larger study into muscularity-oriented attitudes and behaviors.

2.3.1. Drive for muscularity. The Drive for Muscularity (DMS; McCreary, 2007) is a 15-item measure developed to assess attitudes and behaviors related to the pursuit of muscularity. The survey measure asks participants to rate responses on a 6-point Likert scale between Never and Always, where higher scores are indicative of a

NB: This study is currently under review. Please refer to the final published material when citing.
higher level of the Drive for Muscularity. The DMS has demonstrated strong internal consistency and test-retest reliability (McPherson, McCarthy, McCreary, & McMillan, 2010). In the current sample, internal consistency of the DMS was excellent (Cronbach’s alpha = .91).

2.3.2. Muscularity-oriented disordered eating. The Eating for Muscularity Scale (EMS; Cooper & Griffiths, submitted) is a 30-item instrument that measures pathological eating attitudes and behaviors driven by a desire to enhance muscularity. The instrument utilizes a 7-point Likert scale requesting participants to rate the number of days in the previous four weeks in which they have engaged in each item (0 = no days, 6 = every day). The EMS had high internal consistency, moderate to strong concurrent validity with associated measures, good construct validity, and strong test-retest reliability in the current sample (Cooper et al., submitted).

2.3.3. Obligatory exercising. Obligatory exercising was evaluated using the Obligatory Exercise Questionnaire (OEQ; Pasman, & Thompson, 1988). The OEQ is a brief 20-item measure comprising three subscales - exercise fixation, exercise frequency and exercise commitment - which aim to identify individuals at risk for an eating disorder (Ackard, Brehm, & Steffen, 2002). The instrument employs a 4-point Likert scale where 1 represents ‘Never’ and 4 represents ‘Always.’ The OEQ has demonstrated strong internal consistency as well as strong concurrent validity with other measures of exercise dependence and addiction (Terry, Szabo, & Griffiths, 2004; Thompson & Pasman, 1991). It yielded excellent internal consistency in the current sample (Cronbach’s alpha = .92).

2.2.4 Supplement and substance use. Use of appearance and performance enhancing supplements was measured utilizing the 18-item Appearance and Performance Enhancing Supplement/Substance Use Scale (APES; Cooper, Hicks, & Griffiths,
The APES comprises an 8-item Supplement subscale and an 11-item Substance subscale requiring participants to indicate usage on a 4-point Likert scale “how much time you have used each of these supplements in the past 3 years” (0=never, 1= a little (0-1 month), 2=moderate (1-6 months), 3=a lot (6 months or more). The APES has demonstrated strong internal consistency, medium to strong concurrent validity, and good to excellent test-retest reliability in the current sample (Cooper et al., submitted).

2.2.5 Body dissatisfaction. The Male Body Attitudes Scale (MBAS; Tylka, Bergeron, & Schwatrz, 2005) was utilized to evaluate body dissatisfaction in the sample. The 24-item measure employs a 6-point Likert scale ranging from 1 representing ‘Never’ to 6 representing ‘Always.’ Constructed to evaluate aspects of body image preoccupation and dissatisfaction, the MBAS comprises three subscales: Muscularity, Body Fat and Height. The MBAS has demonstrated high internal consistency as well as strong convergent validity with other measures of body esteem and attitudes towards muscularity (Tylka et al., 2005). The MBAS demonstrated excellent internal consistency in the current sample (Cronbach’s alpha = .93).

2.2.5 Body image perception. The Body Image Assessment for Obesity (BIA-O; Williamson et al., 2000) was utilized to evaluate body image perception in the sample. The BIA-O comprises a series of silhouettes that vary in body shape, the participant being requested to select the silhouette that they believe most resembles their own shape. The tool is a valid measure of current body shape and has demonstrated good temporal stability (Williamson et al.). Although the original BIA-O was originally intended to be administered on individual cards, we presented the silhouettes online for the purpose of this study.

NB: This study is currently under review. Please refer to the final published material when citing.
Procedure

Interested participants accessed the Information Sheet and survey through Google Docs. The Information Sheet provided participants with an overview of the research purpose and procedure as well as any risks or benefits associated with participation. The total battery was estimated to take participants a total of 20 minutes to complete.

Statistical Analyses

Data from participants were exported as a CSV file directly from the Google Forms website. Analyses were then performed using IBM SPSS Statistics version 25 (IBM Corporation, 2017). Body image distortion was calculated as discrepancy scores in a two-stage process: 1) the BMI and the BIAO scores for every individual were each converted into z-scores using population normative data (ABS, 2015; Williamson et al., 2000 respectively); 2) for each individual, the BMI z-score was subtracted from BIAO z-score. Larger discrepancy values, regardless of direction, indicated a greater degree of body image distortion relative to the community norm. Independent t-tests were conducted to explore gender differences. Pearson’s r was used to explore bivariate correlations between dependent and independent variables. Significant demographic predictors were then included in a hierarchical regression to examine the individual contribution of muscle dissatisfaction and body image distortion to each dependent variable. A multivariate multiple linear regression was conducted to examine the model of muscularity-enhancing attitudes and behaviors (DMS-Att, OEQ, EMS, and APES).

3. Results

3.1 Descriptive results

The descriptive findings for each of the measured variables are presented in Table 1. Notably, compared with females, males showed significantly higher

NB: This study is currently under review. Please refer to the final published material when citing.
muscularity dissatisfaction, attitudes related to the drive for muscularity, and reported use of appearance- and performance-enhancing supplements/substances. Accordingly, gender was included in the multivariate multiple regression. There were no gender differences in muscularity-oriented disordered eating or obligatory exercise.

<INSERT TABLE 1 HERE>

3.2 Bivariate correlations

Bivariate associations between outcome measures and between age and each outcome variable are shown in Table 2. Age was not related to any outcome variable, while higher BMI was associated with lower muscularity-oriented eating and heightened obligatory exercise. Accordingly, BMI was included in the following multivariate multiple regression. As can be seen small, positive associations were found between dissatisfaction with muscularity and each of the dependent outcome variables (ie. exercise, APED usage, attitudes for the drive for muscularity and muscularity-oriented disordered eating). Moderate significant relationships were also found between each of the dependent variables. Notably, body image distortion was not associated with any of the dependent variables.

<INSERT TABLE 2 HERE>

3.3 Individual hierarchical regression analyses

Individual hierarchical regressions were performed for each individual dependent variable controlling for demographic variables (BMI and gender) to determine the final model. The first analysis controlled for BMI and gender, finding that both gender and muscle dissatisfaction significantly contributed to muscularity-oriented eating. The second analysis examined APED use, finding that once gender and BMI were controlled for, only muscle dissatisfaction made a significant contribution. Results from the third analysis revealed that once controlling for gender
and BMI, both BMI and muscle dissatisfaction predicted obligatory exercise. The final analyses revealed that by controlling for gender and BMI, only muscle dissatisfaction predicted attitudes related to the drive for muscularity. After controlling for BMI and gender, body image distortion did not significantly contribute to any of the dependent variables.

3.4 Multivariate multiple regression analyses

The overall multivariate model, including BMI, gender and muscle dissatisfaction, predicted 52% of the variance in the MD variables (Wilks’ Lambda = .48, F(4, 141) = 38.78, p<.001). Post-hoc analyses examining the contribution of muscle dissatisfaction to each of the variables can be seen in Table 3. Muscle dissatisfaction uniquely predicted 79% of the variation in MD symptomatology (Wilks’ Lambda = .21, F(4,140)= 133.54, p<.001) and, as can be seen, had the strongest relationship with attitudes related to the drive for muscularity.

4. Discussion

This study aimed to explore the relative contributions of body image dissatisfaction and body image distortion to engagement in muscle dysmorphic behaviors. This is the first study, that we are aware of, that has compared the roles of these two key components of body image disturbance in predicting MD symptomatology. The results suggest that body image dissatisfaction, and not body image distortion, are associated with increased engagement with risky muscle enhancing strategies.

The association between muscularity dissatisfaction and MD symptomatology is supported by a significant body of research (Diehl & Baghurst, 2016; Olivardia, Pope, & Hudson, 2000). In contrast, the role of body image distortion in MD symptomology...
has been presumed based upon case studies and anecdotal reports (Pope et al., 1993; Pope et al., 1997). BDD, and MD as a consequence, has been defined as a disorder characterized by a distortion in one’s perception of self in one or more areas of concern (American Psychiatric Association, 2013). As such our results are surprising, raise questions about the DSM definition of MD and have the potential to contribute to the future understanding of the condition.

To date, meta-analyses have found insufficient evidence for MD as a clinically distinct diagnosis from other existing diagnoses including BDD, anorexia nervosa, bulimia nervosa and obsessive-compulsive disorder (Cooper et al., in preparation; dos Santos Filho et al., 2016; Nieuwoudt et al., 2016). The current findings suggest that body image distortion may prove an important factor in differentiating MD from similar psychiatric diagnoses. Our findings also further contribute to evidence that MD may be more appropriately classified as a body image disorder (Murray et al., 2013). As there are no current evidence-based interventions for MD, such classification may support the trialing of treatments for MD that have been modified from those developed for anorexia and bulimia nervosa. In a recent case report by Murray and Griffiths (2015), a 15-year old male with MD was successfully treated with an adapted family-based approach for anorexia (Lock & Le Grange, 2013). Further studies are required to examine the treatment efficacy of this approach on a larger scale as well as to trial the effectiveness for MD of other effective eating disorder treatments such as enhanced cognitive therapy (Fairburn, 2008).

Despite the strengths of this study there were limitations. Firstly, our study found smaller correlations between the BIA-O and BMI than in previous studies using silhouette measures (Conti et al., 2013; Ralph-Nearman & Filik, 2018). Whilst this may indicate higher levels of overall distortion in our sample, it is possible that the

NB: This study is currently under review. Please refer to the final published material when citing.
use of the online method of presentation reduced the validity of the measure. Future studies are required that utilize a measure validated for online delivery. Furthermore, as with similar epidemiological studies of body image discrepancies (Conti et al., 2013; Gardner, Jappe, & Gardner, 2009), our use of the BIA-O limited the choice of current body image into larger or smaller physique. Due to this, we were unable to differentiate increasing size related to higher muscularity from increasing size related to higher adiposity. Future studies would benefit from the inclusion of a silhouette measure that allows for both changes in muscularity as well as adiposity as separate variables. Furthermore, as can be seen from the gender differences on most of the measures utilized, females are typically less likely to exhibit the drive for muscularity and muscularity-enhancing behaviors. As such, future studies would be recommended to include a larger proportion of female participants.

5. Conclusions

To our knowledge, this research is the first to explore the relative contributions of two components of body image disturbance to engagement with MD behaviors. Significantly, our results contradicted the presumption of perceptual disturbances in MD symptomology, a finding which, if replicated, has substantial implications for the classification and understanding of MD. Additional research is required to further elucidate similarities and differences between thinness-oriented eating disorders (e.g., anorexia and bulimia nervosa), BDD and MD. It is hoped that this research will allow for further development of MD and similar diagnoses in clinically meaningful heterogeneous populations, to improve research, treatment and assessment outcomes.

NB: This study is currently under review. Please refer to the final published material when citing.
References


Cooper, M., Hicks, R., & Griffiths, K. M. (submitted). Preliminary psychometric characteristics of a measure for appearance and performance-enhancing supplement and substance usage.


*NB: This study is currently under review. Please refer to the final published material when citing.*
dysmorphia among personal trainers. *Cogent Psychology, 3*, 1-20. doi:
10.1080/23311908.2016.1243194

dos Santos Filho, C. A., Passarelli Tirico, P., Stefano, S. C., Touyz, W. W., &
category muscle dysmorphia. *Australian and New Zealand Journal of
Psychiatry, 50*, 323-33. doi:10.1177/0004867415614106

York: Guilford Press.

Foster, A. C., Shorter, G. W., & Griffiths, M. D. (2013). Muscle dysmorphia: Could it
be classified as an addiction to body image? *Journal of Behavioral
Addictions, 4*, 1-5. doi:10.1556/JBA.3.2014.001

new figural drawing scale of body-image assessment: The BIAS-BD. *Journal
of Clinical Psychology, 65*, 113-122. doi:10.1002/jclp.20526

Detection of Muscle Dysmorphia symptoms in male weightlifters. *Anales de
Psycologia, 33*, 204-210. doi: 10.6018/analesps.33.1.233311

IBM Corp. (2011). *IBM SPSS Statistics for Mac* (Version 20.0) [Computer Software].
Armonk, NY: IBM Corp.


McCreary, D.R. (2007). The Drive for Muscularity Scale: Description,


*NB: This study is currently under review. Please refer to the final published material when citing.*


*NB: This study is currently under review. Please refer to the final published material when citing.*


*NB: This study is currently under review. Please refer to the final published material when citing.*
TABLES AND FIGURES

1. The person has a preoccupation with the idea that one’s body is not sufficiently lean and muscular. Characteristic associated behaviors include long hours of lifting weights and excessive attention to diet.
2. The preoccupation causes clinically significant distress or impairment in social, occupational, or other important areas of functioning, as demonstrated by at least two of the four following criteria:
   i. The individual frequently gives up important social, occupational, or recreational activities because of a compulsive need to maintain his or her workout and diet schedule;
   j. The individual avoids situations where his or her body is exposed to others, or endures such situations only with marked distress or intense anxiety;
   k. The preoccupation about the inadequacy of body size or musculature causes significant distress or impairment in social, occupational, or other important areas of functioning; and
   l. The individual continues to work out, diet or use ergogenic substances despite knowledge of adverse physical or psychological consequences.
3. The primary focus of the preoccupation and behaviors is on being too small or inadequately muscular, as distinguished from fear of being fat, as in anorexia nervosa, or a primary preoccupation only with other aspects of appearance, as in other forms of BDD.

Box 1. Pope and colleagues’ criteria for MD. From Pope, Gruber, Choi, Olivardia, & Phillips, 1997, p.556

1. The person has a preoccupation with one or more perceived defects in their physical appearance.
2. The preoccupation causes clinically significant distress or impairment in social, occupational, or other important areas of functioning.
3. Preoccupations trigger excessive and repetitive behaviors.
4. The primary focus is not related to body fat or weight as in an eating disorder
MD specifier: The preoccupation is focused on physique being too small or insufficiently muscular.


Table 1
Mean (standard deviations) for each measure as a function of gender

<table>
<thead>
<tr>
<th>Gender</th>
<th>MBAS-MD</th>
<th>BDistort</th>
<th>DMS-Att</th>
<th>APES</th>
<th>EMS</th>
<th>OEQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female (n=176)</td>
<td>24.60 (10.44)</td>
<td>0.24 (0.81)</td>
<td>2.21 (1.01)</td>
<td>3.95 (5.12)</td>
<td>1.42 (1.42)</td>
<td>44.58 (12.31)</td>
</tr>
<tr>
<td>Male (n=83)</td>
<td>28.79 (11.33)</td>
<td>1.78 (0.64)</td>
<td>3.03 (1.30)</td>
<td>5.86 (5.79)</td>
<td>0.99 (0.74)</td>
<td>44.18 (11.02)</td>
</tr>
<tr>
<td>Total (n=261)</td>
<td>25.89** (10.83)</td>
<td>0.75** (1.04)</td>
<td>2.46** (1.17)</td>
<td>4.53** (5.37)</td>
<td>1.27 (1.24)</td>
<td>44.47 (11.76)</td>
</tr>
</tbody>
</table>

Table 2
**Bivariate correlations between measures**

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>BMI</th>
<th>BDistort</th>
<th>MBAS-MD</th>
<th>APES</th>
<th>DMS-Att</th>
<th>EMS</th>
<th>OEQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1</td>
<td>.17*</td>
<td>-.06</td>
<td>-.07</td>
<td>-.02</td>
<td>-.10</td>
<td>-.03</td>
<td>-.07</td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td>1</td>
<td>.06</td>
<td>.09</td>
<td>.13*</td>
<td>.06</td>
<td>-.03</td>
<td>-.16*</td>
</tr>
<tr>
<td>BDistort</td>
<td>1</td>
<td></td>
<td>.14</td>
<td>.12</td>
<td>.12</td>
<td>-.01</td>
<td>-.08</td>
<td></td>
</tr>
<tr>
<td>MBAS-MD</td>
<td></td>
<td></td>
<td>1</td>
<td>.47**</td>
<td>.86**</td>
<td>.42**</td>
<td>.40**</td>
<td></td>
</tr>
<tr>
<td>APES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>.42**</td>
<td>.46**</td>
<td>.52**</td>
</tr>
<tr>
<td>DMS-Att</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Table 3
**Results of post-hoc analyses for model of muscle dissatisfaction**

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>SE B</th>
<th>t</th>
<th>p</th>
<th>Partial Eta squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>APES</td>
<td>0.27</td>
<td>0.04</td>
<td>6.54</td>
<td>&lt;.001</td>
<td>.23</td>
</tr>
<tr>
<td>DMS-Att</td>
<td>0.10</td>
<td>0.00</td>
<td>22.18</td>
<td>&lt;.001</td>
<td>.78</td>
</tr>
<tr>
<td>EMS</td>
<td>0.06</td>
<td>0.01</td>
<td>5.81</td>
<td>&lt;.001</td>
<td>.19</td>
</tr>
<tr>
<td>OEQ</td>
<td>0.57</td>
<td>0.10</td>
<td>5.75</td>
<td>&lt;.001</td>
<td>.19</td>
</tr>
</tbody>
</table>

**p<.001 APES = Appearance and Performance Enhancing Supplement and Substance Use Scale. BDistort – Body image distortion, DMS-Att – Drive for Muscularity Scale –attitudes, EMS = Eating for Muscularity Scale, MBAS – Male Body Attitudes Scale (Muscle Dissatisfaction), and OEQ – Obligatory Exercise Questionnaire.

**NB:** Assumption testing was undertaken for all analyses. As well as those already included in text, all assumptions were met or parametric testing was still deemed viable due to large sample sizes. An a priori power analysis was conducted determining that a minimum sample of 119 participants would be required for this study.
CHAPTER 6

Overview
This chapter provides a review of the papers included in this thesis. It summarises and integrates the findings across the papers, considers the research and clinical implications of the results, analyses the strengths and limitations of the studies, and concludes with a discussion of directions for future research.

6.1 Summary of overall project findings
This project undertook a systematic review of the extant research literature and conducted three original investigations to advance our understanding and assessment of MD. The studies sought to provide a clear picture of MD and to develop psychometrically sound assessment instruments to better inform researchers and clinicians in the field. More specifically, as indicated in Chapter 1, the project aimed to address the following research questions:

   RQ1: Is there currently sufficient evidence to differentiate MD from healthy individuals?

   RQ 2: Is there currently sufficient evidence to differentiate MD from existing psychiatric diagnoses?

   RQ 3: Can we accurately measure symptoms of MD and, where required, develop psychometrically sound instruments to measure these symptoms?

   RQ 4: Can we identify further factors to differentiate MD from similar psychiatric diagnoses?
The study findings for each of these questions are briefly summarised.

**RQ1: Is there currently sufficient evidence to differentiate MD from healthy individuals?**

The review undertaken in Study 1 indicated that individuals with MD are likely to constitute a clinically relevant population. Our findings suggested that individuals with MD exhibit greater APED use, body dissatisfaction, disordered eating, exercise pathology, obsessive-compulsive symptomology, impairment, and overall mental ill-health than healthy controls. Additionally, individuals with MD were found to report lower subjective wellbeing than their healthy counterparts. Despite these promising results, all findings must be interpreted with caution due to high degree of heterogeneity between studies and the small number of papers identified.

**RQ 2: Is there currently sufficient evidence to differentiate MD from existing psychiatric diagnoses?**

This project was unable to find sufficient evidence that MD can be delineated from existing psychiatric diagnoses. We found evidence that individuals with MD had lower eating pathology than individuals with an eating disorder and higher muscularity-oriented body change behaviours than individuals with BDD. Despite this, there were no significant differences between MD and eating disorders on measures of exercise pathology and muscle dissatisfaction; furthermore, there were no differences between MD and BDD on measures of disordered eating and obsessive-compulsive pathology. However, the body of evidence comprised few studies comparing MD with other clinical diagnoses and there was a high degree of heterogeneity between studies.
RQ3: Can we accurately measure symptoms of MD and, where required, develop psychometrically sound instruments to measure these symptoms?

Among the factors highlighted in Study 1 to delineate MD and similar disorders, both muscularity-oriented disordered eating and APED use were identified as lacking psychometrically sound instruments. Therefore, this project developed and assessed the validity and reliability of two novel measures: the Appearance and Performance-Enhancing Supplement and Substance Use Scale (APES; Study 2) and the Eating for Muscularity Scale (EMS; Study 3). These measures were constructed using a comprehensive development process and were each validated in two community samples. Our findings demonstrated that both measures possess strong internal consistency, very good test-retest reliability, and good concurrent validity with associated measures.

RQ 4: Can we identify further factors to differentiate MD from similar psychiatric diagnoses?

A final aspect highlighted in Chapter 1, was the absence of evidence from quantitative empirical studies supporting body image distortion as a core component of MD. An important outcome of this project was the finding that MD-related attitudes and engagement in MD behaviours (including obligatory exercise, muscularity-oriented disordered eating, and APED use) were associated with muscle dissatisfaction and not body image distortion. This preliminary research indicates that body image distortion may represent a factor that differentiates MD from similar existing psychiatric diagnoses.
6.2 Clinical and research implications of findings
The following section will provide an overview of the results of this project in the context of their clinical and research implications.

6.2.1 MD as a clinically relevant population

As a relatively new psychiatric diagnosis, a key stage in validating MD is determining whether it is truly a clinically condition. This stage appears to be particularly relevant in the case of MD, where both the scientific and general communities have questioned whether this diagnosis is a popular media fad (Vandereycken, 2011). Our findings from Study 1 provided cautious support for MD as a clinically relevant condition, individuals with MD demonstrating significantly higher psychopathology and lower protective factors when compared with healthy controls and their bodybuilding peers. These findings have important implications for: increasing community awareness of masculinity-oriented body image concerns; training and education of healthcare providers; and future research into body image concerns.

Body image-related disorders have historically been viewed as affecting young, affluent, Caucasian women motivated by the pursuit of thinness (Smolak & Striegel-Moore, 2001). Perhaps consequently, males who experience anorexia and bulimia nervosa are less likely than their female counterparts to seek treatment, despite similar likelihood of mortality to females following hospital discharge (Griffiths et al., 2015; Hoang, Goldacre & James, 2014; Lewisohn, Seeley, Moerk, & Stiegel-Moore, 2002). As males with both traditional eating disorders and MD commonly report masculinity concerns and low self-worth, the social stigma related to male body image concerns is likely to serve as an additional barrier to help-seeking behaviours and may increase social isolation
(Griffiths et al., 2015; Mussap, 2008). In addition, the conceptualisation of body image-related disorders through the lens of the drive for thinness also stigmatises individuals who exhibit dysfunctional investment in body ideals unrelated to the thin ideal (such as the mesomorphic ideal).

Consequently, we anticipate that the wider dissemination of our findings supporting MD as a clinical condition will increase social awareness of the nuances in maladaptive body image concerns. As awareness of these diverse body image ideals increases, there is potential for society to change the heuristics perceived of those at risk of experiencing a body image-related disorder (i.e. young, Caucasian females). Arguably, these changes are inextricable from efforts to decrease the public stigma associated with body image concerns and diverse maladaptive body change behaviours in non-cisgendered females and other diverse groups. Ultimately, this decreased stigma and increased societal awareness may lead to increased treatment seeking behaviours by individuals with MD.

As well as their potential to change community perceptions of those at risk of experiencing body image disturbance, the current research has potential relevance to healthcare and medical professionals. Primary health care providers are typically the first to identify mental health conditions in the community (Swartz, King & Rider, 2011). However, screening for mental health conditions in a primary health care setting is shockingly low, even with more prevalent mental health concerns (Diamond et al., 2011; Garnick, Horgan, Merrick, & Hoyt, 2007), and providers commonly report inadequate training in identifying and treating mental health concerns (McCue Horwitz et al., 2015). The use of formal screening tools has been demonstrated to significantly increase provider
effectiveness in identifying mental health issues (Brown & Wissow, 2010; Beger-Jenkins, McCord, Gallgher, & Olfson, 2011). Consequently, a key factor in supporting individuals with MD is to improve education and adequate tools for primary health care providers. Furthermore, as athletes and frequent gym-users may be at heightened risk for MD, further education and screening tools may also be beneficial for physical trainers and sports coaches. We consider that the dissemination of the findings of this research, along with previous MD research, will equip providers and adjunct professionals with knowledge that supports their informed clinical responses to MD and encourages the use of screening tools with individuals presenting with MD-related symptomatology such as APED use, anxious or mood-related symptomatology, and frequent exercise injuries.

Finally, it is also hoped that the support of MD as a clinically relevant condition will influence researchers within the body image field to broaden study investigations to include muscularity-oriented and other diverse body image concerns. Approximately 1% of all research examining anorexia nervosa has been conducted in males (Murray, Griffiths, & Nagata, 2018) and our initial study found that 70% of the MD research undertaken thus far has used male bodybuilding and weightlifting samples. It could be argued that conducting the majority of research in these selected samples (such as young females for AN and weightlifting males for MD) facilitates an understanding of individuals most at risk of developing these disorders. Indeed, it is important to undertake research that investigates specific gender, ethnic, and cultural groups, especially when the targeted group is likely to demonstrate unique characteristics or risk factors, as may be the case for male bodybuilders. However, an over-reliance on limited
samples also leads to biases and limitations in the development of valid clinical pictures of diagnoses that are applicable in diverse populations.

An illustration of the limitations associated with using narrowly defined samples to determine diagnostic criteria can be seen in the domain of eating disorders (ED). Prior to the release of the DSM-5 (American Psychiatric Association, 2013), no ED diagnostic criteria permitted the diagnosis of binge eating without compensatory behaviours, or the diagnosis of purging behaviours without binge eating behaviours, and individuals with anorexia nervosa were required to achieve a BMI of less than 16. Consequently, these criteria led to between 70-80% of individuals diagnosed with an ED being classified under the category of eating disorder not otherwise specified, a cluster of symptoms in which individuals met some but not all symptoms of the “typical” ED such as anorexia nervosa and bulimia nervosa (Le Grange, Swanson, Crow, & Merikangas, 2012; Machado, Goncalves, & Hoe, 2013). Although the DSM-5 has introduced the new diagnoses of binge eating disorder, and purging disorder (under the category of other specified feeding or eating disorder), and removed the specific weight criterion from anorexia nervosa, these diagnostic categories still remain overly homogeneous, a factor which is likely contribute to the ongoing level of public stigma highlighted earlier.

Consequently, it is hoped that the findings from Study 1 supporting MD as a clinically relevant disorder will stimulate future research to examine more diverse body image concerns. This may include incorporating samples with greater gender, size, and ethnic diversity, as well as examining a greater range of areas of body image focus. These designs have the potential to contribute to
more culturally sensitive, body-diverse, and gender-neutral methods of understanding psychopathology.

6.2.2 Delineation of MD from existing diagnoses

The development of new psychiatric conditions requires that these diagnoses are clinically distinct from existing disorders (Blashfield, Sprock, & Fuller, 1990). As there has been disagreement regarding the nosological classification of MD as well as its delineation from similar disorders, our project aimed to examine the demarcation of MD from other disorders. Consistent with previous reviews, we were unable to find sufficient evidence to support the conclusion that MD is clinically distinct from existing diagnoses. Despite this, Studies 1 and 4 both highlighted potential areas that may delineate MD in future research including APED usage, eating pathology, muscle-enhancing behaviours, and body image distortion. These findings have important implications for informing clinical-decision making and the development of MD treatment protocols as well as future MD research.

With respect to the delineation of MD, this project had two primary purposes: first, to determine whether there is sufficient evidence for MD as a clinically distinct diagnosis from other existing diagnoses and secondly to identify similarities with existing disorders to improve the categorization of MD. An over-focus on categorization and classification may be judged an academic exercise, presenting minimal benefit to those struggling with mental health conditions. Nevertheless, these determinations are critical to our understanding of newly identified disorders such as MD. The development of efficacious interventions for new disorders is both a costly and time-consuming procedure,
with estimates suggesting that it can take at least 17 years for interventions to be transition from conceptualization to clinical practice (Institute of Medicine, 2001). Consequently, during this period practitioners who prescribe interventions must do so without the benefit of an evidence-base to guide their decisions.

Since there are currently no evidence-based interventions for MD, there is a clear need for research to guide clinicians working in this field. Although the results from Study 1 were unable to delineate MD from traditional eating disorders or BDD, our findings did contribute to the existing literature by providing evidence that MD may be more accurately classified as a body image disorder (Studies 1 and 4). The diagnosis of both anorexia and bulimia nervosa presently include the criterion of individuals exhibiting either an overevaluation of weight/shape or undue influence of weight/shape on an individual’s self-worth (American Psychiatric Association, 2013). Consequently, these individuals typically engage in maladaptive behaviours and are dissatisfied with their weight/shape in comparison to their desired thin ideal (American Psychiatric Association). Similarly, our findings from Study 4 indicate that individuals engaging in maladaptive muscle-enhancing behaviours are dissatisfied with their shape in comparison to a muscular ideal. Classification of MD as a body image disorder with diagnostic similarities to anorexia and bulimia nervosa would support the extrapolation of research into these diagnoses to MD populations.

In fact, there has been considerable support for reclassifying MD as a body image disorder among the scientific community (Murray et al., 2012; Phillipou, Blomley, & Castel, 2016). This position has been bolstered by evidence suggesting that MD has similar aetiological mechanisms, course, and treatment
response to ED (Lamanna, Grieve, Pitt Derryberry, Hakmann & McClure, 2010; Murray et al., 2012). Study 1 identified a number of cases studies that provided support for the utility of ED protocols for the treatment of MD including the use of pharmacological interventions (Ung, Fones & Ang, 2000), cognitive therapy (Phillips, O’Sullivan, & Pope, 1997; Ung et al., 2000), and adapted family therapy for anorexia nervosa (Murray & Griffiths, 2015). Each of these are indicated interventions for individuals with anorexia and bulimia nervosa. In the absence of more direct, higher quality evidence, the current analysis may indicate that a clinician seeking to prescribe a treatment for MD might best select one of these interventions. Other evidence-based ED protocols such as enhanced cognitive therapy (Fairburn, 2008) and specialist supportive clinical management (McIntosh et al., 2006) may also have potential for the treatment of individuals with MD. However, there is a clear need to undertake appropriate trials of these interventions to develop an evidence-base on which to base authoritative clinical practice guidelines.

6.2.3. Assessment of MD

Accurate assessment measures are required to facilitate the empirical investigation of new psychiatric diagnoses, their delineation from existing conditions, and the evaluation of their treatment efficacy. Consequently, a core aim of this project was to improve the assessment of MD-related symptomatology. The APES and EMS, developed in Study 2 and 3 are novel psychometric tools designed to facilitate the accurate assessment of two symptoms of MD, namely APED use and muscularity-oriented disordered eating.
These instruments have key implications for the assessment of MD in clinical and research settings.

Standard clinical practice recommends that practitioners assess the pre- and post-treatment symptomology of patients to evaluate treatment outcomes. Furthermore, the use of questionnaires can be a vital component in the early stages of the clinical information gathering process. As there are currently no psychometrically sound survey tools for the evaluation of APED and muscularity-oriented disordered eating, the measures developed in this project may support more accurate assessment of symptomology in individuals experiencing MD or those presenting with APED usage. These measures aim to contribute to more competent information gathering procedures in treatment as well as improved evaluation of treatment outcomes in clinical settings. This contribution will be of further benefit with future research to identify clinical norms and cut-off values for these measures.

These assessment tools provide similar benefits for the field of research. Previous research into both MD and APED use has been hindered by inconsistent measurement and the use of inappropriate assessment tools. Valid and reliable assessment tools not only allow for further delineation of MD from existing diagnoses, as noted earlier, but can also support the future examination of aetiological mechanisms of MD, treatment outcomes, and the prevalence of these behaviours at a population level. Additionally, standardised measurement across studies can permit more effective comparison of findings. We anticipate that this will lead to greater understanding of MD within the scientific community and, consequently, dissemination of these findings to the general population.
6.2 Methodological strengths and limitations of this project

6.2.1 Strengths

This project overall had multiple strengths. Firstly, it was the first to examine the validity and reliability of MD as a construct, as opposed to focusing on testing specific diagnostic criteria (either those suggested by Pope and colleagues, 1997, or the American Psychiatric Association, 2013). As highlighted in Chapter 1, there are key concerns with not only the reliability and validity of both existing MD criteria but also with the lack of a theoretical framework, adequate sampling, and sufficient evidence used to inform their development. Consequently, in our view, a focus on examining these criteria is premature. Study 1, therefore, utilised the framework of Robins and Guze’s criteria (1970) for psychiatric validity to develop a clear clinical picture of MD. We acknowledge that there is a need for further examination and exploration of the clinical pictures outlined in Study 1. However, in our view they provide a promising starting point for developing clinically valid diagnostic criteria for MD.

In contrast to our examination of MD using Robins and Guze’s criteria (1970) for determining psychiatric validity, previous meta-analyses (dos Santos Filho, Passarelli Tirico, Stefano, Touyz, & Mederios Claudino, 2016; Nieuwoudt, Zhou, Coutts, & Booker, 2012) in the field have employed Blashfield and colleagues’ (1990) criteria for proposing new diagnostic categories. The latter criteria focus on whether there is sufficient empirical research for a new diagnostic category and whether there is a set of reliable and clinically unique diagnostic criteria that typically present as a co-occurring syndrome. While the Blashfield criteria might achieve their aim of preventing the premature inclusion of under-researched diagnostic criteria in diagnostic systems, they do not
include as a criterion the differentiation of diagnostic groups from healthy populations. Thus, despite our challenges in delineating MD from existing diagnoses, we were able to present the novel finding supporting MD as a clinically relevant condition when compared with healthy controls.

A further strength of the current project was that we purposefully selected psychometric instruments that focused on muscularity-oriented concerns and, where no psychometrically sound measures were available, we developed novel measures of muscularity-related symptomology. Both Chapter 1 and 2 (Study 1) identified that the current clinical understanding of MD symptomology has been hindered by the use of body image and behavioural assessments that examine the drive for thinness. It is likely that these measures have contributed to the ongoing underestimation of body image concerns and eating pathology in individuals exhibiting the drive for muscularity (Murray, Griffiths, & Nagata, 2018), particularly in males who typically score lower on these measures than female comparators (Jennings & Phillips, 2017; Smith et al., 2017). Consequently, our novel measures in this project were specifically developed in samples comprising individuals from both indicated and universal populations with an, overall, equal gender representation. This sampling method was chosen to optimise the validity and generalisability of our instruments by reducing the limitations associated with the biased and limited sampling that have characterised previous MD research.

In addition to the strengths associated with the use of a rigorous psychometric development and validation processes, the measures themselves have strengths. For example, in contrast with much of the MD and APED literature, the APES incorporates items assessing both licit and illicit APED
usage. Licit APED use has previously been recognised as a potential risk factor for future illicit APED usage (Hildebrandt et al., 2012) with supplement users indicating more positive attitudes towards doping and increased endorsement of illicit APED use as safe (Barkoukis, Lazarus, & Tsorbatzoudis, 2015; Carvey, Farina, & Lieberman, 2012; Hartmann & Siegrist, 2016). Additionally, licit APED use, whilst legal, is often undertaken without the supervision of medical professionals, which has led to increased presentations to emergency departments and preventable doping bans in due to mislabeled ingredients (Abbate et al., 2014; Kohler, Thomas, Petrou, Schanzer, & Thevis, 2009; Nordt et al., 2012). Thus, unsupervised usage of licit APED is a key public health concern along with illicit APED usage. Addressing this concern and undertaking research that can better inform future policy and decision-making in this area requires accurate assessment of both licit and illicit APED usage with a psychometrically sound measure such as the APES.

Finally, this research was the first to explore the relative contributions of body image disturbance, and two of its subcomponents, to MD symptomology. The association between muscularity dissatisfaction and MD symptomology has been supported by a significant body of research (Diehl & Baghurst, 2016; Olivardia, Pope, & Hudson, 2000). However, the assumption that body image distortion is inherent to MD has been based predominantly on findings from case reports (Hernandez-Martinez, Gonzalez-Marti & Jordan, 2017; Pope et al., 1993; Pope et al., 1997). In isolation, our results do not refute the presence of body image distortion in MD presentations. However, they do strongly highlight the need for further exploration of the accuracy of the BDD conceptualization of MD.
6.2.2 Limitations

We have reviewed the individual limitations of each project within their relevant papers. To reduce redundancy, this section will review the limitations of the project overall.

Firstly, a particular limitation of our project was the use of self-report instruments. Self-report measures rely on individuals to accurately report their symptomology. Although there is always expected error in self-report instruments due to factors such as impression management and response bias, a key concern in psychopathology research is the impact of the individual introspective ability of study participants (Bartholome, Peterson, Raatz, & Raymond, 2013; Schaub, Brune, Bierfoh, & Juckel, 2012). While, at present, there is not specific research on the introspective ability of individuals with MD symptomology, preliminary findings by Pope and colleagues’ (2005) did find significant delusionality in their sample of 14 males with co-occurring MD and BDD. Furthermore, individuals with MD possess a similar ego syntonic attitude towards their symptomology as individuals with AN. As individuals diagnosed with AN and other eating disorders often engage in denial of their symptomology, it has been suggested that this may result in low introspective awareness of their disorder and, subsequently, their self-reported symptomology (Bartholome et al., 2013; Garner & Garfinkel, 1979; Swenne, Belfrage, Thurfjell, & Engstrom, 2005). Further research is required to assess the presence and impact of symptom denial on results in clinical MD groups. This may include using convergent measures of symptomatology that do not rely on self-report measures such as structured clinical interviews and clinician-rated measures.
A further limitation of our research was the sampling bias associated with sourcing participants from Facebook. A key aim of this project was to include a sample representative of the general population, including some at-risk individuals. We anticipated that this methodology would address the biases that have previously limited the generalisability of MD research undertaken in male weightlifting and bodybuilding populations. However, as noted in Studies 2-4, the Facebook sample included approximately 88% Caucasian and 90% heterosexual participants. Similar overrepresentation of Caucasian participants is a common critique of social media recruiting methods (Gu et al., 2016; Stern, Bilgen, McClain, & Hunscher, 2016). Overreliance on heterosexual, Caucasian participants at the exclusion of ethically diverse individuals has previously contributed to poor cultural validity and generalisability of findings, perpetuating underestimations of body image concerns in minority populations (Kashubeck-West et al., 2013; Smolak & Stiegel-Moore, 2001). As both the APES and EMS were validated in the Facebook sample, further research must be undertaken in culturally diverse samples to ensure the cross-cultural suitability of these assessments. It should be noted that, despite a similarly high proportion of heterosexual individuals, our sample sourced from Amazon Mechanical Turk did comprise a more ethnically diverse sample. This is consistent with previous research that has suggested that participants recruited from Amazon Mechanical Turk exhibit greater ethnic, geographic, and economic diversity than other convenience samples (such as undergraduate samples) (Casler, Bickel, & Hackett, 2013).

A final limitation was our selection of community samples for this project. The decision to focus on non-clinical settings was based on a number of factors.
Individuals sourced from clinical groups may not be typical of individuals with MD given that few seek professional help for the condition (Leone, Sedory, & Gray, 2005). Targeting individuals from mixed-gender community settings also provides an opportunity to reduce the level of gender bias in sampling. Further, community samples can provide sufficient participants to support meaningful statistical analyses. However, we acknowledge that our lack of a clinical sample does limit the clinical inferences that are possible from the results. Nevertheless, the data and findings can inform future research on clinical populations with an MD diagnosis and treatment history. The EMS and APES in particular may prove useful in the assessment and differential diagnosis of MD-related concerns in clinical populations.

6.3 Future research directions

This project identified that: the clinical classification of MD is not resolved; there are challenges in delineating MD from similar diagnoses; there is an absence of follow-up studies into MD; new measures APES and EMS demonstrate promising psychometric characteristics in community and at-risk samples; and the evidence did not support the widely presumed role of body image distortion in MD behaviours.

The evidence that the nosology of MD is yet to be resolved points to the need to develop valid and reliable diagnostic criteria for MD. Study 1 presented two potential clinical pictures of MD, as a purging disorder and a body image disorder. These clinical pictures, we hope, will guide future researchers to develop valid and reliable diagnostic criteria for MD. As these findings were based upon the clustering of individual case studies, they need to be further examined in larger samples.
Study 1 also identified the need for further research to better understand factors that delineate MD from existing diagnoses to support its clinical utility. Both criteria for MD (American Psychiatric Association, 2013; Pope et al., 1997) have suggested that the direction of weight/shape concerns represent a differentiating criterion for MD and anorexia nervosa. However, with Study 1 indicating that exercise symptomology and muscle dissatisfaction is comparable between eating disorders and MD, we question whether this differentiation is meaningful. To address this concern, future research should utilise the EMS to more accurately assess MDE as part of a suite of assessments that examines whether there is a need to differentiate these disorders.

As Study 1 and Study 4 also identified that APED use and body image distortion may differentiate MD from BDD, future research should use the APES and measures of body image distortion amongst a battery of assessments to delineate MD from BDD. The utilisation of our novel instruments could be combined with an Ockham’s razor approach to address the demarcation of MD. Delineation of psychiatric diagnoses should be meaningful in their contribution to our understanding, assessment and treatment of the condition. Although this study, and prior research, have highlighted the direction of body image dissatisfaction, use of APED, quality of life, and delusionality, it must be questioned whether these differences are meaningful to patients with MD and their care. The Research Domain of Criteria (RDoC; Insel, Cuthbert, Garvey, Heinssen, & Pine, 2010) represents a paradigm shift in psychiatry for redesigning psychiatric dimensions with adequate heterogeneity. Similar to Robins and Guze’s (1970) second criteria of laboratory studies, the RDoC recommends exploring diagnostic commonalities and differences based on
factors such as positive and negative valence systems, cognitive systems, social processing systems, and arousal systems (Morris & Cuthbert, 2012). We propose that the development of evidence-based clusters of body image symptomology can improve the clinical validity of these diagnoses and reduce redundancy that results from meaningless delineation of disorders. This can, thus, facilitate the development of clearer nosology of disorders, whereby researchers and funding sources can undertake or support projects which address larger systemic group challenges rather than focusing on unnecessarily specific sub-categories.

Study 1 also identified an absence of research into the course, prognosis, and heritability of MD. Future research must examine these areas to continue improving our understanding of MD. Increased understanding of MD and its aetiological mechanisms may be utilised for delineation studies, but as mentioned previously, there is also a need for increased social awareness of masculinity-oriented concerns to reduce public stigma. Awareness raising initiatives need to be informed by research into the effectiveness of interventions (e.g. educational programs) to increase MD literacy, reduce MD stigma, and increase MD help-seeking. Additionally, there is a clear need for research that will guide the future development of efficacious prevention and treatment protocols evaluate the effectiveness of these interventions in the MD population using high quality randomised controlled trials.

Finally, whilst Study 2 and 3 provided promising psychometric examinations of the APES and the EMS, there is a need for future research to examine these measures in clinical samples. We have critiqued previous instruments for their absence of support from clinical groupings and while we were able to demonstrate the known-group validity of the EMS with individuals
exhibiting heightened MD symptomatology, there is a need for further replication of these findings. In these future studies, we aim to continue to examine the APES and EMS to develop clinical norms and cut-offs as well as examine known-group validity in individuals with clinically presenting individuals with MD and APED abuse/dependence. These findings will lead to improvements in the accurate assessment of MD symptomology in both clinical and research settings as well as facilitate future empirical research into populations with APED use and musculature-oriented disordered eating.

6.4 Conclusion

This project aimed to investigate the clinical validity and assessment of MD and its relevant symptomology. We established the validity of MD as a clinically relevant condition; although further research is required to establish a clear clinical picture of MD and to determine whether it is clinically distinct from existing psychiatric diagnoses. Our inability to demarcate MD from existing diagnoses resulted, to some extent, from a paucity of research. However, it was also hindered by the absence of psychometrically valid and appropriate assessment measures. Our two novel and valid instruments, the APES and the EMS, are tools that can be used in future studies examining the clinical validity and delineation of MD as a psychiatric diagnosis. The unexpected finding that distorted body image does not predict MD in community samples may prove a critical finding in the understanding of the MD if it is replicated in clinical MD populations. In summary, the current research has the potential to stimulate future research that will enhance our understanding of MD and ultimately lead to the development of effective prevention and treatment protocols.


20, 324-330. doi:10.1177/1359104514521639


