# saxinstitute

**Evidence Check** 

Survey instruments for monitoring selfesteem in children

An **Evidence Check** rapid review brokered by the Sax Institute for the NSW Ministry of Health. March 2018.

An **Evidence Check** rapid review brokered by the Sax Institute for the NSW Office of Preventative Health. March 2018.

This report was prepared by:

Maxine Bonham, Leah Brennan, Julie Brimblecombe, Claire Bristow, Kaitlin Day, Aimee Dordevic, Terry Haines, Margaret Murray, Laura Tirlea, Helen Truby.

Monash University

March 2018 © Sax Institute 2018

This work is copyright. It may be reproduced in whole or in part for study training purposes subject to the inclusions of an acknowledgement of the source. It may not be reproduced for commercial usage or sale. Reproduction for purposes other than those indicated above requires written permission from the copyright owners.

Enquiries regarding this report may be directed to the:

Principal Analyst Knowledge Exchange Program Sax Institute www.saxinstitute.org.au knowledge.exchange@saxinstitute.org.au Phone: +61 2 91889500

Suggested Citation:

Bonham M, Brennan L, Brimblecombe J, Bristow C, Day K, et al. Survey instruments for monitoring selfesteem in children: an Evidence Check rapid review brokered by the Sax Institute (<u>www.saxinstitute.org.au</u>) for the NSW Ministry of Health, 2018.

doi:10.57022/vjnq5286

# Disclaimer:

This **Evidence Check Review** was produced using the Evidence Check methodology in response to specific questions from the commissioning agency.

It is not necessarily a comprehensive review of all literature relating to the topic area. It was current at the time of production (but not necessarily at the time of publication). It is reproduced for general information and third parties rely upon it at their own risk.

# Survey instruments for monitoring self-esteem in children

An **Evidence Check** rapid review brokered by the Sax Institute for the NSW Ministry of Health. March 2018.

This report was prepared by Maxine Bonham, Leah Brennan, Julie Brimblecombe, Claire Bristow, Kaitlin Day, Aimee Dordevic, Terry Haines, Margaret Murray, Laura Tirlea, Helen Truby.

# Contents

Glossary of terms	6
Executive summary	7
Background / Purpose of the review	7
Review questions	7
Summary of methods	7
Evidence grading	7
Key findings	7
Gaps in the evidence	9
Discussion of key findings	9
Conclusion	10
Background	13
Purpose and audience	13
Self-esteem defined	14
How is self-esteem measured?	14
Methods	15
Peer review literature	15
Evidence grading	16
Grey literature	18
Findings	19
Question 1: What short-form questionnaires for monitoring self-esteem in children aged 7-13 years hav been validated for use in healthcare intervention settings?	'e 19
Question 2: What validated short-form self-esteem questionnaires have been evaluated for their usabili and acceptability in vulnerable populations?	ty 32
Gaps in the evidence	34
Discussion/synthesis of findings	35
Conclusion / Recommendations	37
Objective recommendations	38
Expert opinion	38
References	40
Appendices	44
Appendix 1: Example database search, strategy 1	44
Appendix 2: Example database search, strategy 2	45
Appendix 3: NHMRC levels of evidence	46

Appendix 4: Intervention studies measuring self-esteem as identified in search 1	47
Appendix 5: Summary of systematic reviews that examine self-esteem as a component of weight loss interventions	82
Appendix 6: Validation studies in populations of interest of identified tools from search 1	. 83
Appendix 7: Overview of self-esteem measurement tools	. 88
Appendix 8: Multicomponent weight management interventions that assessed self-esteem in overweigl and obese adolescents from vulnerable population groups	ht 91
Appendix 9: Summary of NHMRC Guidelines for working with Aboriginal and Torres Strait Islander populations	93
Appendix 10: Recommendations for validating a self-esteem tool	. 94
Appendix 11: Harter's Self Perception Profile Children (SPPC)	95

# **Glossary of terms**

**Construct validity:** Construct validity is perhaps the most complex index of whether an instrument assesses what it purports to measure. Basically, it refers to the demonstration that if a given measure of a particular construct is inserted into a matrix of theoretical predictions or a model where specific predictions are advanced, and the predictions that involve the construct are supported, then one indirectly concludes that the measure of the construct is valid.

**Convergent validity:** Convergent validity typically refers to the fact that scores on one index of a given construct "converge" with parallel indices of the same constructs on different instruments.

Face validity: Refers to the fact that "on its face", the content of items on a given instrument look like credible markers of the construct in question, that is, they are relatively transparent (thus, also the term content validity).

**Factorial validity**: Determines the strength of clusters (factors) within a questionnaire and the interrelations of the individual items that make up the cluster.

Global self-esteem: A person's overall appraisal of their worth.

Healthcare intervention settings: Programs designed to target health outcomes specifically run by/from healthcare professionals/facilities i.e. General Practitioners, dietitians, psychology, physiotherapists.

**Multicomponent intervention**: An intervention with more than one of the following aspects; physical activity/exercise program, diet prescription, nutrition education and overall behaviour change.

**Multidimensional self-esteem tools:** Tools implementing numerous subscales of self-esteem (i.e. academic, sports performance), which form the overall total measure of global self-esteem.

**Reliability:** Cronbach's alpha is a measure of internal consistency: that is, how closely related a set of items is as a group. It is considered to be a measure of scale reliability.

**Self-concept:** The idea a person holds of oneself, constructed from their own beliefs and responses from others.

**Self-esteem:** A person's beliefs of their own self-worth including how much they like, respect and accept themselves.

Validity: Validity refers to the fact that a measure assesses what it was intended to measure.

Vulnerable populations: Persons or groups at high risk of developing poor health outcomes i.e. Indigenous, elderly and low SES populations.

**Weight-management program**: Any program specifically designed to target weight loss and weight maintenance in the group in which the program is implemented.

# **Executive summary**

# **Background / Purpose of the review**

This rapid review was commissioned by the NSW Office of Preventive Health to provide recommendations on the most suitable survey instruments to monitor self-esteem for use in the Go4Fun, Aboriginal Go4Fun and Go4Fun Online Programs. Go4Fun, Aboriginal Go4Fun and Go4Fun Online are evidence-based 10-week healthy lifestyle programs that promote healthy eating, physical activity and confidence building in children aged seven to 13 years who are above a healthy weight.

Building self-esteem is an important component of interventions that focus on achieving a healthy weight.<sup>1</sup> The outcomes from this review will inform changes to the current pre- and post-program self-esteem questionnaire used in the Go4Fun programs. The recommendations will take into consideration user burden; validity and reliability; ability to detect change in self-esteem; and usability and acceptability in culturally and linguistically diverse (CALD), lower socioeconomic status (SES) and low-literacy populations.

#### **Review questions**

This review aimed to address the following questions:

# **Question 1**:

What short-form questionnaires for monitoring self-esteem in children aged 7-13 years have been validated for use in healthcare intervention settings?

#### **Question 2:**

What validated short-form self-esteem questionnaires have been evaluated for their usability and acceptability in vulnerable populations?

# **Summary of methods**

In order to provide as detailed a review as possible, a two-tiered strategy to the review was undertaken. Firstly, a systematic search was undertaken to identify multicomponent weight-management interventions (similar to Go4Fun) in children and adolescents (overweight or obese) from 2007 onwards that had used a measure of self-esteem. This enabled us to identify short-form self-esteem tools currently in use in behavioural weight-management interventions. The second search strategy identified studies on children and adolescents that had specifically reported on subsequent psychometric testing of any of the tools identified in search strategy 1. This strategy assisted in making recommendations regarding the usability and acceptability of self-esteem tools in vulnerable populations (as unlikely to be explicitly reported) and their applicability in healthcare settings. A summary of the findings is outlined below.

# **Evidence grading**

Sixteen randomised controlled trials (NHMRC Level II evidence), two non-randomised experimental trials (NHMRC Level III-2 evidence), one interrupted time-series controlled study (NHMRC Level III-3 evidence) and 17 pre-test and post-test trials (NHMRC Level IV evidence) were identified in search strategy 1; an adequate number of high-level evidence studies to draw conclusions from.

# **Key findings**

# **Question 1:**

In this review we identified seven validated tools that measure self-esteem, with a number of tools having a number of subscales. All seven identified tools have demonstrated acceptable validity and reliability,

however; not necessarily in all types of validity testing. Poor study design, the population group originally validated in, and year of validation (many more than 15 years ago) makes recommending an appropriate tool for this current setting and population group challenging.

Overall, the review highlights a lack of empirical evidence regarding the psychometric properties of selfesteem tools currently used in paediatric healthcare settings in Australia. Where a large number of validated self-esteem tools do exist, no short-form questionnaires for monitoring self-esteem in healthy children aged 7-13 years have been validated specifically for use in such settings. Validation studies of well-regarded selfesteem tools have typically been cross-sectional, school-based and primarily undertaken in the US. Subsequently, these tools have been used in a range of interventions, age ranges and ethnicities without consideration of the limited psychometric testing, and/or the validity and reliability of the tools in that specific setting. The only short-form questionnaire that has been validated via an intervention to improve self-esteem (US children, Grades 4 and 7) is the Piers-Harris Self Concept scale, and this study was conducted in a school setting.

These limitations have not deterred researchers from using self-esteem tools in healthy lifestyle programs for children. In our search (2007 onwards) of multicomponent weight-loss interventions in children and adolescents (overweight and obese) we retrieved 37 papers reporting on 15 randomised controlled trials (RCTs), 17 pre-test and post-test trials, two non-randomised experimental trials and one interrupted time-series controlled study using self-esteem measures. Across the studies, seven different validated self-esteem measurement tools were identified:

- The Harter's Self-Perception Profiles for children (SPPC) and for adolescents (SPPA)
- The Rosenberg Self-Esteem Scale (RSES)
- Piers-Harris Self Concept Scale, second edition (Piers-Harris-2)
- Marsh's Self-Description Questionnaire-I (SDQ-I)
- Children and Youth Physical Self-Perception Profile (CY-PSPP)
- The Beck Youth Inventory II (BYI-II).

The included tools either measure global self-esteem, self-esteem as a series of constructs or include a measure of self-esteem as one of the subscales of a measure of broader psychosocial constructs. All tools are validated, show acceptable reliability (> 0.7 Cronbach's alpha) and validity (face, construct and convergent). Eight of the intervention studies reported their own reliability data, with Cronbach's alpha ranging from 0.75 to 0.95, suggesting good to excellent internal consistency. All tools except for the RSES are multidimensional, meaning they implement numerous subscales of self-esteem (i.e. academic, sports performance) that form the overall total measure of global self-esteem. However only the SPPC, SPPA, SDQ-I and CY-PSPP were used to report subscales of self-esteem among the identified intervention studies; the RSES, Piers-Harris-2 and BYI-II were used to report global self-esteem only.

# **Question 2:**

Of the multicomponent weight management interventions that assessed self-esteem in overweight and obese children and adolescents, six were conducted in populations considered vulnerable or of low SES.<sup>2-7</sup> An improvement in self-esteem was identified in five of these six studies, with one showing no change. Two different tools were used in these studies: the RSES and the CY-PSPP. These tools are child completed or in conjunction with a parent/carer.

No study commented directly on the acceptability and usability of the identified tools. A second search sought to identify whether the tools identified in search strategy 1 have been subsequently validated in 'vulnerable' 7-13-year olds. Studies that used a translated version of the questionnaire in another language, or were not undertaken in Australia, New Zealand, Canada, Europe or the US were excluded. Only participants in one of the nine identified psychometric testing studies were considered a vulnerable

population. This study used Harter's SPPC in African-American girls from low SES areas (Stewart et al 2010). Moderate to poor validity was reported, with the authors stating that 'subscales of the SPPC should be cautiously interpreted when used with different ethnic and racial groups'.

No tools administered online (English versions only included) have been validated, making recommendations regarding their usability and acceptability challenging. None of the included self-esteem tools have been specifically validated in Aboriginal children and adolescents. Through search strategy 2, however, the Racial Identity and Self-Esteem of Children (IRISE\_C scale)<sup>8</sup> was identified. This tool measures racial identity and self-esteem in Aboriginal children and appears to show good validity, but has not been used in pre- or post-test conditions so its ability to detect change is currently unknown.

# Gaps in the evidence

- 1. There is a lack of intervention studies with self-esteem as primary outcomes. As such, the reporting of self-esteem data is limited
- 2. Validation studies of self-esteem tools are lacking in Australian children and adolescents aged between seven and 13 years
- 3. Of the validation studies found only one was conducted within the past 10 years and this study focused on the validation of the SPPC in African-American adolescent girls. None of the tools have been validated in populations that would reflect modern-day children in the Go4Fun Program
- 4. The self-esteem tool we propose for use in Go4Fun, the SPPC, has not been the primary focus of a validation study in Australia. However, Marsh & Holmes<sup>9</sup> demonstrated good validity when comparing the SPPC to the SDQ-I in Australian children. It has been translated into a number of languages (for example Spanish, Dutch and Italian) and validated in vulnerable populations outside of Australia
- 5. There is a paucity of validation studies of self-esteem tools in an online setting. The Piers-Harris-2 has been validated online but only in Taiwan
- 6. This review highlights the need for psychometric testing of self-esteem measures in Australian children and adolescents both in a face-to-face and online setting. The Go4Fun program offers an intervention that can be used constructively to achieve this aim
- 7. Guidance on whether self-esteem tools can be child administered or in association with a parent/carer (especially in the younger age group) also needs to be considered

# **Discussion of key findings**

# **Question 1**:

Our findings summarise the use of validated self-esteem tools in multicomponent weight management interventions for children and adolescents. All self-esteem tools identified in search strategy 1 have been validated pre-2007 and are collectively suitable for children and adolescents between the age ranges of seven and 18. Typically, validation studies of self-esteem tools have been undertaken in the school setting.

Of the identified intervention studies, 23 identified an improvement in self-esteem over the course of the intervention, 10 studies found no change, one saw a reduction in self-esteem, and in one study it was unclear whether there was a change in self-esteem. Physical self-esteem was the most commonly reported subscale of self-esteem to show improvement following intervention. Of the 16 studies that reported on subscales of self-esteem, 12 identified an improvement in physical self-esteem. The RSES (13 studies) and the Harter (22 studies) tools (SPPC and SPPA) were the most frequently used tools.

In our expert opinion tools that do not focus primarily on self-esteem (BYI-II) or have infrequently been used in behavioural weight-management programs since 2007 (SDQ-I, Piers-Harris-2 and the CY-PSPP) should not form part of the recommendations, especially as their psychometric properties do not perform better than other tools. In addition, the CY-PSPP primarily focuses on physical self-esteem.

The only self-esteem tool identified through the search strategy that has been validated in Australia is the SDQ-I. This Australian-developed tool has been validated only in adolescents. As such it cannot be recommended for the 7-10 age group without validation. Only one self-esteem tool has been validated in Aboriginal children and adolescents. The IRISE-C is a tool that evaluates racial identity and self-esteem. Although its validity and reliability are fair, it has not been tested in pre- or post-intervention settings and therefore its ability to measure change in self-esteem is currently unsubstantiated. Further we would question the relevance of this measure to programs for weight reduction unless the program also aimed at improving racial identity.

A search of the grey literature points to the use of multidimensional self-esteem tools in intervention setting such as Go4Fun. Overall, small changes in self-esteem are observed after weight-management interventions. The SPPC and RSES were reported as the most commonly used tools. The 'Young Minds Matter' program reported validation of a new short form tool (13 items) in Australian adolescents (Adolescent Self-Esteem Questionnaire) in 2017. However, the tool has not been validated in younger children (7-10 years), nor in an intervention setting and as such cannot be recommended for immediate use.

# **Question 2:**

Evidence regarding the usability/acceptability of the included validated self-esteem tools in vulnerable populations is scarce in both the published literature and grey literature. No included intervention studies that measured self-esteem commented directly on the acceptability or usability of the identified tools. In order to assess this, data were extracted to determine, of the participants who completed the interventions, the percentage of participants who also completed the self-esteem measures. Completion rates for the RSES were highest, with rates of or close to 100%, where these figures could be identified. The CY-PSPP had the poorest completion rate of the three tools used in vulnerable populations, being 45.8%. All tools were completed by the child or adolescent participants with assistance, if required in order to complete the RSES. This is not to say that assistance should not be offered, particularly to vulnerable or CALD populations where English is not the first language, if administering the other tools, it just wasn't offered in the identified studies.

None of the tools (English version) have been validated in an online environment, making commentary regarding their usability and acceptability challenging.

# Conclusion

Low self-esteem is a complication of paediatric overweightness and obesity and is associated with adverse consequences. Although fluctuations in self-esteem are a feature of growing up, children who are overweight are reported to be at risk for greater declines in self-esteem as they enter early adolescence when compared to their healthy-weight counterparts.<sup>10</sup> Multicomponent weight-management programs that target self-esteem can induce weight loss as well as positively impacting on self-esteem.<sup>1</sup> Based on evidence and expert opinion we are making the following recommendations to the MOH for the inclusion of a validated measure of self-esteem in the Go4Fun program:

# **Objective recommendations**

- 1. All identified tools have been validated in children and/or adolescents and show good reliability and validity
- 2. Based on ease of use, and the understanding that RSES is already used in Go4Fun, keeping the RSES is a consideration, but this review identified that there are more suitable, multidimensional tools
- 3. A thorough review of the literature points to adoption of a multidimensional self-esteem tool (which excludes RSES). In the absence of changes in global self-esteem, multidimensional tools can be helpful in examining sub-components of self-esteem
- 4. After consideration of the available evidence, we propose adopting the Harter SPPC questionnaire into the Go4Fun program. This tool appears to be appropriate for both the younger and older ages that enrol in the program, eliminating the challenges associated with two different tools.
- 5. Whichever tool is used, it is recommended that its reliability is tested in a Go4Fun cohort
- 6. For the younger age group (7-10 years) we recommend parental/carer support in completion of the SPPC. This would be the case for the majority of validated self-esteem tools selected and not specific to the SPPC
- 7. As no tools have been validated in an online environment we would recommend adoption of the SPCC for online use and also validation in an online setting
- 8. Aboriginal children require the presence of an Aboriginal staff member to support completion of the survey. This staff member should preferably be from the children's community and preferably occupy the lead position in tool administration and be an integral member of the team
- 9. We would not recommend the IRISE-C. Although its validity and reliability are fair, it has not been tested in pre- or post-intervention settings and therefore its ability to measure change in self-esteem is currently unsubstantiated. Furthermore, we question the relevance of this measure to programs for weight reduction unless the program also aimed to improve racial identity.

# **Expert opinion**

- Prior to adoption of the Harter SPPC (or any tool) we recommend a validation of its use in Australian children and adolescents. The Go4Fun program provides the ideal opportunity. For details on how to achieve this please see Appendix 10
- 2. We see no issues with adopting the SPPC in Aboriginal Go4Fun, however it would need to be translated into the appropriate language, in consultation with the specific communities. However, there are aspects of identifying as an Aboriginal person that would be integral to self-esteem that are not captured in this tool. This would require a separate study, as this review has not identified any available self-esteem tools that are specific to Aboriginal people
- 3. The use of one tool for the age range 7-13 years is also beneficial in terms of pre-post change as the children age. Having to change tools as children move from one bracket (e.g. 7-10 years) to another (e.g. 11-13 years) has implications for evaluation
- 4. The SPPC has been translated into a number of languages (for example Spanish, Dutch and Italian) and validated in vulnerable populations outside of Australia, indicating its potential for use in Australia
- 5. While the SPPC is the best available tool for this purpose, we realise that the length of the SPPC may be a barrier to its use. To reduce the time associated with completion of the SPPC, selecting

subscales previously shown to be particularly relevant to obesity in children and adolescents may be considered as an alternative. The following subscales most frequently reported to change in children and adolescents after participation in behavioural lifestyle interventions are physical appearance and athletic competence

- 6. When validating self-esteem measures in Aboriginal and Torres Strait Islander populations we recommend that the NHMRC guidelines for ethical conduct in Aboriginal and Torres Strait Islander Health Research be considered, Appendix 9. Referral to the 'Cultural adaptation of the Go4Fun program for delivery with Aboriginal communities in NSW (pages 9 and 10) is also recommended
- 7. Relying on a single form of self-esteem measurement may not be sufficient. Clinically, multiple methods to gauge one's self-esteem are used including: interviews, behavioural observation and ratings by others (e.g., teachers, counsellors, medical professionals, parents)

# Background

Go4Fun is an evidence-based 10-week healthy lifestyle program that promotes healthy eating, physical activity and building confidence amongst children, aged 7 to 13 years who are above a healthy weight. Go4Fun is based on the UK MEND (Mind, Exercise, Nutrition ... Do it!) program<sup>11, 12</sup> and has subsequently been adapted for community settings in Australia.<sup>13, 14</sup> This multidisciplinary program has been designed to achieve long-term behaviour change for the whole family by incorporating education on nutrition, physical activity and behaviour change.<sup>15, 16</sup> The program started out as a face-to-face program led by trained health professionals but has undergone various iterations to improve the reach and accessibility of the program to priority population groups. These adaptations include Aboriginal Go4Fun and Go4Fun online. Aboriginal Go4Fun is a culturally adapted version of the program, which is aimed at Aboriginal children and their families, whereas Go4Fun online enables broader access of the program to remote and rural areas.

The program is a key initiative of the NSW Office of Preventive Health (OPH), who commissioned a number of reports to improve the program. These reviews have included a cultural review<sup>17</sup> and a routine review of the survey instruments used to assess physical activity and diet.<sup>18</sup> However, other measures included in the program such as self-esteem were not included in this latter review.

In 2017 the NSW OPH commissioned a new review to evaluate survey instruments used to measure selfesteem in similar programs in order to provide guidance and recommendations on the appropriateness and acceptability of the current self-esteem scale (Rosenberg Self-Esteem Scale) used in Go4Fun. This tool was adopted into the program in 2008 and has not been psychometrically tested for use in this population. Feedback from program leaders has indicated the current self-esteem measure (a modified version of the Rosenberg Self-Esteem Scale) has some limitations. This has become more significant since the Aboriginal adaptation of the program in which the scale is not well received.

# **Purpose and audience**

The purpose of this review was to inform the NSW OPH of the best available evidence for:

1. Existing short-form questionnaires for monitoring self-esteem in children aged 7-13 years that have been validated for use in healthcare interventions

Scope to include questionnaires that are able to detect change over time and have been specifically designed, tested or validated in healthy lifestyle programs for children

2. Questionnaires that have been evaluated for their usability and acceptability in vulnerable populations with a focus on whether questionnaires are suitable for self-completion by the child alone or together with a parent/carer

Scope to include use of the questionnaires in online settings and whether it is suitable for use in Aboriginal populations

This review will inform changes to the current pre- and post-program self-esteem questionnaire for children 7-13 years old with recommendations/expert opinion provided for the most suitable instrument for use in the Go4Fun, Aboriginal Go4Fun and Go4Fun Online Programs.

# Self-esteem defined

Self-esteem has been variously defined in the literature but is taken to mean a person's overall appraisal of their worth, including how much they like, respect and accept themselves ("How do I feel about who I am?").<sup>19</sup> There are two distinctive aspects to the self-esteem definition. Self-esteem involves an evaluation along a continuum ranging from positive to negative with an optimum level being somewhere in the middle of this continuum. Secondly, the evaluation includes a cognitive component (i.e., how one thinks about oneself) and an affective component (i.e., how one feels about oneself).<sup>20, 21</sup>

# How is self-esteem measured?

Self-esteem has traditionally been measured via self-report questionnaires. Although there are limits to what clinicians/researchers can expect when using self-esteem instruments (i.e. self-report bias), the use of pen and paper, self-report instruments remain the primary and most reliable means of ascertaining self-esteem specifically and psychosocial variables more generally.

Either global or domain specific aspects are captured by self-esteem measures. Global self-esteem can be distinguished from domain-specific evaluations of the self and is defined as "the totality of the individual's thoughts and feelings having reference to himself [sic] as an object".<sup>22 p. 7</sup> Domain-specific evaluations (i.e., academic competence, physical appearance, social competence) are usually included in the overall, global self-esteem measure. Depending on the research target either global or specific domains are used to assess the whole or parts of self-esteem. Whereas specific self-concepts are found to be more highly related to content-specific outcomes (e.g., academic self-concept as a predictor of academic achievement), global self-esteem was found to be more strongly related to more global outcomes such as mental health and wellbeing.<sup>23</sup>

# Methods

# Peer review literature

A rapid review of both peer-reviewed and grey literature was conducted to provide an evidence check for the review questions. The research team was commissioned in October 2017 to complete the rapid review, and the search strategy was discussed and agreed. The review team had expertise in: systematic reviews, multicomponent lifestyle interventions that measured self-esteem as an outcome; working with Indigenous communities and psychometric testing studies.

The first draft of the report was submitted for feedback in November and the final report submitted in December 2017.

A two-tiered approach was undertaken to fulfil the requirements of the Evidence Check as outlined below. Firstly, a systematic search strategy was undertaken to identify self-esteem tools used in multicomponent weight-management interventions in overweight/obese children and adolescents, similar to the Go4Fun program. On identification of the tools a second search strategy was undertaken to evaluate the validity of these tools in 7-13-year olds under similar intervention conditions as Go4Fun.

The literature predominantly comprised peer-reviewed journals, however grey literature searching identified systematic reviews and reports on public health interventions that met inclusion criteria and were included in the review. Selected reports from the NSW OPH were also provided to the research team.

The following database searches were used to provide the published peer-reviewed literature for this review:

# Search strategy 1

Search strategy 1 was run in Ovid Medline (2007 – current); CINAHL Plus (2007-current); Scopus (2007 – current) and PsycINFO (2007 – current). In accordance with the original commissioned request from the Sax Institute, a search strategy was devised by the research investigators in consultation with the commissioning group to identify relevant studies published between 2007 and current day in the English language (Table 1). A representative list of free-text keywords was generated and entered into electronic databases for mapping to subject headings.

A final database search was conducted on the 19th October 2017. The full list of the search terms is included in Table 1 with an example search strategy in Appendix 1:

Table 1 Search strategy for identification of self-esteem tools used in behavioural interventions in children and adolescents

Field 1	Field 2	Field 3	Field 4
adolescen* OR youth* OR child* OR p?ediatric OR teenage* OR boy* OR girl* OR student* OR pupil* OR school*	obes* OR overweight	(Nutrition* OR Diet* OR Health* OR Behavio* OR Lifestyle OR Community OR "Physical activity" OR Exercise OR Fitness OR Weight OR Psycho* OR Multi- component OR Multicomponent	"self-esteem" OR "self- concept" OR "self- perception" OR psychosocial OR "self- worth" OR "self- evaluation" OR "self- confidence" OR "self-

	OR Family* OR Obes* OR School*)	love" OR "self- compassion"
	WITHIN 3 WORDS	
	(Educat*, Program*, Intervention, Care, Therap*, Treatment, Promotion)	
	OR	
	"group work" OR intervention OR prevention	

Combine fields with AND

Limit from 2007 to current

# Search strategy 2

Ovid Medline, CINAHL Plus and PsycINFO were searched for the second strategy. The second search strategy was developed to capture peer-reviewed literature regarding studies evaluating the validity of the self-esteem measurement tools identified in search strategy 1, for use with Australian children aged 7-13 years. Two versions of the search were run. One search using fields 1-3 (see Table 2) and a second search using fields 1, 3 and 4 that specifically searched for validation studies in Aboriginal populations. Population groups were limited to Australia, New Zealand, UK, US and Europe to ensure relevance to the Australian target audience and the Go4Fun program. Self-esteem measurement tools were limited to versions in the English language. Studies that included Aboriginal and Torres Strait Islander children and vulnerable populations such as low SES were searched for as part of the inclusion criteria, but only Aboriginal was used as a specific search term. There was no date limit on this search. A final database search was conducted on the 3rd November 2017. The full list of the search terms is included in Table 2 with an example search strategy in Appendix 2:

Field 1	Field 2	Field 3	Field 4
child* OR teen* OR adolescen*	hild* OR teen* OR dolescen* Harter* self-perception" OR Rosenberg OR Marsh adj self- description OR "Beck Youth Inventory" OR "Piers-Harris" OR "Children and Youth Physical Self- Perception Profile"		Aborigin*
Combine fields with AND			

#### Table 2 Search strategy for self-esteem validation studies

# **Evidence grading**

Fifteen randomised controlled trials (NHMRC Level II evidence), two non-randomised experimental trials (NHMRC Level III-2 evidence), one interrupted time-series controlled study (NHMRC Level III-3 evidence) and 17 pre-test post-test trials (NMHRC Level IV evidence) were included in the review from search strategy

1. Overall there were a large number of high-level evidence studies to draw conclusions from (See Appendix 3).

# **Included studies**

# Search strategy 1

All study eligibility criteria were devised in accordance with the scoping document provided by the Sax Institute and subsequent discussions with the commissioning group. It was agreed that key focus areas for this review include: RCTs and pre-test and post-test studies, children and adolescents aged 7-13 years (those who fall outside this age group may be acceptable) who are overweight or obese with a focus on multicomponent weight-management interventions in line with Go4Fun. The list of agreed inclusion and exclusion criteria is outlined in Table 3.

A total of 3669 articles were identified. Titles and abstracts were screened by three investigators with 111 full texts assessed for eligibility. Thirty-seven papers (representing 35 studies) were deemed to meet the inclusion criteria for this review. Included papers were: 1) in overweight and obese children and adolescents, 2) investigating a multicomponent weight-management intervention, and 3) reporting pre- and post-intervention self-esteem as an outcome. In agreement with NSW OPH we extracted relevant information from these studies including intervention details, length of intervention, population (children and adolescents), self-esteem tools used and the impact of the intervention on markers of self-esteem (see Appendix 4). Traditional quality assessment was not conducted, as the quality of the study (typically referring to evaluation of the RCT design) was not reflective of the validity of the tool used.

# Table 3: Search strategy 1 study eligibility criteria

Inclusion	Exclusion
<ul> <li>Overweight and obese participants</li> <li>Any participants aged 7 to 13 years</li> <li>Multicomponent lifestyle interventions for weight management</li> <li>Studies that report self-esteem pre- and post- intervention as an outcome</li> </ul>	<ul> <li>Papers or measurement tools in languages other than English</li> <li>Reviews and other published works not based on primary research</li> <li>Participants with comorbidities where the focus of the intervention was not on weight loss or lifestyle change.</li> <li>Studies conducted outside of Australia, UK, Canada, USA, New Zealand or Europe</li> </ul>

# Search strategy 2

A second search was undertaken to specifically focus on the validity of the self-esteem tools identified in the first search. Study eligibility criteria were devised in accordance with the scoping document provided by the Sax Institute and similar to the search strategy from the Evidence Check on validated short-form survey instruments for children's diet and physical activity commissioned in 2016.<sup>18</sup> The final list of agreed inclusion and exclusion criteria is outlined in Table 4.

A total of 628 articles were identified. Titles and abstracts were screened by two independent investigators, following which, 98 full texts were assessed for eligibility. Nine papers were deemed to meet the inclusion criteria for this review. Included papers were: 1) validation studies 2) in children and adolescents, 3) reporting self-esteem as an outcome. Extracted data included study population, self-esteem tool and reported validity of tool (see Appendix 4). When 'Aborigin\* was included in the search strategy only one paper was retrieved. This paper<sup>11</sup> had also been identified through grey literature searching.

# Table 4: Search strategy 2 study eligibility criteria

Inc	lusion	Ex	clusion
•	Validation studies	•	Studies that validated questionnaires in
•	Age group to include children and adolescents 7-13		languages other than English
	years of age	•	Reviews and other published works not
•	Self-esteem tools selected from search strategy 1 or		based on primary research
	alternate self-esteem measurement tools deemed of		
	relevance		

# **Grey literature**

Grey literature was also sourced from web-based searches including; Google, Google Scholar, PAIS Index, National Health and Medical Research Council (NHMRC), Aboriginal Health and Medical Research Council (AHMRC), Australian Department of Health, Nutrition.gov (US) and the Food and Nutrition Information Center (US). Sources were searched using a combination of similar free text keywords such as "self-esteem" (and) "child\*" (and) "intervention" (or) "program" (and) "activity" (or) "exercise" (or) "nutrition" (or) "diet" (and) "indigenous" (or) "aboriginal". These searches were completed in October 2017.

# Findings

# Question 1: What short-form questionnaires for monitoring self-esteem in children aged 7-13 years have been validated for use in healthcare intervention settings?

In order to answer this question, the findings have been segregated as follows. First an overall summary of the included studies (search strategy 1) is presented followed by a breakdown of the included studies by self-esteem tool. This enables two objectives to be met 1) the effectiveness of programs to impact self-esteem and 2) an indication of current trends in usage of specific self-esteem tools. Details on the included self-esteem tools are then presented with a brief commentary on their overall validity in relation to the age range of interest (7-13 years) (search strategy 2).

# Summary of the intervention studies that measured self-esteem (search strategy 1)

All included studies were multicomponent lifestyle weight-management interventions in overweight and obese children and adolescents that reported self-esteem (or self-concept, self-worth etc.) as an outcome. Thirty-seven papers were identified, reporting on 15 RCTs, 17 pre-test post-test trials, two non-randomised experimental trials and one interrupted time-series controlled study, conducted across Australia (seven studies), UK (eight studies), US (15 studies), Canada (two studies), Norway (one study) and Portugal (one study). The identified studies covered commercial settings (one study), community settings (25 studies), primary care (five studies) and research settings (four studies). Six of the studies reported including participants of a low SES or from a vulnerable population group (e.g. low-income communities or participants with learning disabilities). Studies included between 17 and 13,998 overweight and obese children and adolescents, with ages ranging from four to 19 years. The age ranges studied were largely inconsistent across the studies, with the most commonly identified age range being 13 to 16 years, for adolescents. Among children there was no commonly studied age range. Interventions varied in duration from two weeks up to 12 months, with 10 weeks and 12 weeks being the most common durations.

Across the studies, seven different self-esteem measurement tools were identified representing seven scales: the Harter's Self-Perception Profiles for children (SPPC) and adolescents (SPPA), the Rosenberg Self-Esteem Scale (RSES), Piers-Harris Self-Concept Scale (Piers-Harris-2), Marsh's Self-Description Questionnaire-I (SDQ-I), Children and Youth Physical Self-Perception Profile (CY-PSPP) and Beck Youth Inventory II (BYI-II).

Nineteen studies reported on global self-esteem only, and the most commonly used tool to report global self-esteem was the Rosenberg Self-esteem Scale (13 studies). Sixteen studies reported on global self-esteem in combination with other domains of self-esteem, such as physical appearance and social acceptance. The Harter's Self-Perception Profile (children and adolescents) was the most commonly used tool to report multiple domains of self-esteem (13 studies).

Overall, 23 studies identified an improvement in self-esteem over the course of the intervention, 10 studies found no change, one saw a reduction in self-esteem, and in one study it was unclear whether there was a change in self-esteem. Physical self-esteem was the most commonly reported subscale of self-esteem to show improvement following intervention. Of the 16 studies that reported on subscales of self-esteem, 12 identified an improvement in physical self-esteem. Whether or not a change was detected is more likely to be due to the intervention program, as opposed to the tool selected to measure self-esteem.

To assess the broad acceptability and usability of the identified tools, data was collected to determine, of the participants who completed the interventions, the percentage of participants who also completed the self-esteem measures. Overall the SPPC, SPPA, RSES, SDQ-I and BYI-II tools had relatively good completion rates, particularly the SPPA, which is possibly due to the older age range in which it is administered. The

Piers-Harris-2 and CY-PSPP tools were each only used in one study, so the interpretation of usability is limited, but these two tools had lower completion rates of 19% and 45.8%, respectively.

# Harter's Self-Perception Profile for Adolescents (SPPA)

There was one pre-test post-test trial and four RCTs identified that used the SPPA to measure self-esteem. Of these studies, two were conducted in Australia, two in the US and one did not specify a location. This tool was used across the community, and research-based settings in populations ranging from 22 up to 151 participants. This tool was used exclusively in a 13 to 16-year-old age group. Studies that used this tool ranged in duration from seven weeks up to five months, over which time three studies saw an improvement in global self-esteem, four saw improvements in other self-esteem subscales. Across all studies this tool was completed by the adolescent participant. Of the five studies that used this tool, two reported on the internal consistency of the tool within their study populations. Internal consistency across the five subscales of the tool was reported as Cronbach's alpha 0.75 to 0.88<sup>24</sup>, and as 0.79 to 0.90<sup>25</sup>, which was rated as good to excellent internal consistency. One study cited a previously reported Cronbach's alpha value of 0.74 to 0.92.<sup>26</sup>

# Harter's Self-Perception Profile for Children (SPPC)

The SPPC was the second-most commonly used tool, used by 11 studies, including six RCTs and five pretest post-test studies. This tool has been used in multicomponent weight-management interventions across Australia, Norway, the UK and US, in community, research and primary care settings. Studies that used this tool included from 21 up to 258 participants aged from four up to 17 years, with interventions ranging from two weeks up to six months in duration. Overall, five studies identified an improvement in global selfesteem following intervention, and five identified an improvement in subscales of self-esteem. One study identified a reduction in global self-esteem, but an improvement in physical self-esteem. This questionnaire was also completed by the child or adolescent participant, with assistance for younger children where required. Two studies using this tool reported on internal consistency within their own study populations, which were Cronbach's alpha 0.77 to 0.95<sup>27</sup>, and 0.91.<sup>28</sup> Four studies stated previously reported values for internal consistency, which ranged from Cronbach's alpha 0.71 to 0.86.<sup>29-32</sup> Similar to the SPPA, the SPPC has good to excellent internal consistency in the included studies (Appendix 4).

#### Rosenberg Self-Esteem Scale (RSES)

The RSES was the most commonly used tool among multicomponent weight-management interventions for overweight and obese adolescents. This tool was used in eight pre-test post-test studies and five RCTS across Australia, Portugal, the UK and US. It was primarily used in a community-based setting, but has also been used in primary care and commercial settings. This tool has been used in studies ranging from 19 up to 13,998 participants, with ages ranging from five to 19 years. Studies that used the RSES had interventions ranging from two weeks up to 48 weeks in duration. Ten of the 13 studies found an improvement in self-esteem following intervention; three saw no change. The RSES only measures global self-esteem, so no changes were identified in subscales of self-esteem. One study that used this tool reported a Cronbach's alpha for internal consistency of 0.84 for within their study population.<sup>33</sup> One other study cited a previously reported value for internal consistency of the RSES, of 0.80.<sup>6</sup>

# Piers-Harris Self-Concept Scale (Piers-Harris-2)

Only one of the identified studies used the Piers-Harris-2. This was a pre-test post-test study conducted in Canada in a community-based setting. It consisted of 345 participants aged six to 12 years of age, who participated in an 8 to 12-week multicomponent weight-management intervention. They observed an improvement in global self-esteem over time; subscales were not assessed. This questionnaire was completed by the child participants. There was no mention of any reliability or validity testing of the tool within the study population.

# Marsh's Self-Description Questionnaire (SDQ-I)

Two studies were identified that used the SDQ-I. Two non-randomised experimental trials were conducted in the US in community-based settings. They included between 231 and 269 participants aged nine to 12 years of age. The interventions lasted 12 weeks in duration, after which they observed an improvement in self-esteem and self-esteem subscales. This questionnaire was completed by the child participants. The two studies reported internal consistency of the tool within their study populations, as Cronbach's alpha 0.81 and 0.83.<sup>2, 3</sup>

#### Children and Youth Physical Self-Perception Profile (CY-PSPP)

Again, only one identified study used this tool; a community-based pre-test post-test study in the US. Eighty-four participants, aged eight to 16 years participated in a 10-week multicomponent weightmanagement intervention, which resulted in an improvement in their global and physical self-esteem. This questionnaire was completed by the child/adolescent participants. There was no mention of any reliability or validity testing of the tool within the study population.

# Beck Youth Inventory II (BYI-II)

Two studies used the BYI-II to measure self-esteem (a subscale), one pre-test post-test trial and one interrupted time series controlled study. The latter was a community-based 10-week intervention program, with 144 participants aged six to 17 years, held in Canada. The other was a seven to 13-week intervention that included 17 participants with a mean age of 10.7 years, conducted in a primary care setting in the U.S. The Canadian study observed an improvement in global self-esteem following intervention, whereas the U.S. study saw no change. Subscales of self-esteem were not assessed. In both studies the questionnaire was completed by the child/adolescent participants. One study reported on the internal consistency of this tool, stating that Cronbach's alpha typically exceeds 0.86, and in their study population was 0.94, suggesting excellent internal consistency.<sup>34</sup>

#### Grey literature: question 1

#### Summary of SLRs that have examined measurement of self-esteem in weight-management programs.

The inclusion of systematic literature reviews in this project was to identify parallel literature in order to complement that identified through the systematic search. Systematic literature reviews (SLRs) were identified from the two search strategies outlined above and a brief search from the first three pages of Google Scholar (search terms: children, adolescents, systematic review, self-esteem). Only reviews from 2005 onwards were included. A summary of the most relevant content from four SLRs is included below. A general summary is included in Appendix 5.

Lowry et al.<sup>35</sup> summarised the effects of weight-management programs that reported on self-esteem. Twenty-one studies were included in the review. The age range of the participants were separated into three categories 7-12 years, 11-18 years and a mixed range that included children and adolescents from 7-18 years. Tools used in the included studies were typically the ones being evaluated in this commissioned review. The most widely used tools were the Piers-Harris (original and version 2; 25% of all studies) and Harter (SPPC or SPPA; 50% of all studies included in the review). The majority of the studies used multidimensional assessment, with 10 of the 16 included studies reporting improvements in global selfesteem. Of the remaining studies that did not show an impact on global self-esteem, improvements in components of self-esteem were reported in four studies. Studies using only global self-esteem scales, of which there were two, both reported improvements and three studies did not report the tool used. The authors of the review conclude that overall small positive effects of weight loss on self-esteem in paediatric populations were observed after intervention. However, certain components related to self-esteem such as body image may be impacted first before any changes to global self-esteem are observed. Failure to observe changes in global self-esteem therefore may be due to length of the intervention, rather than ability of the designated tools to measure change. Longer term follow up of intervention studies may be required to fully evaluate the longer-term impacts of short-term weight loss on global self-esteem. Such studies are currently limited.

Griffiths et al.<sup>36</sup> identified 17 papers that measured self-esteem in interventional, cross-sectional or longitudinal studies dating back to 1994. The tools identified in the review were predominantly multidimensional enabling a more hierarchically organised overview of several areas of competence to be assessed. This review included many of the studies identified in other reviews in this section (1, 35) and confirms the Harter's tool (SPPC or SPPA) as being the most frequently used. Six out of nine cross-sectional studies, the only prospective study included and six out the seven interventions all reported use of the SPPC or SPPA. The cross-sectional studies were informative in that they generally reported a lower self-esteem in obese children compared to their normal weight counterparts (six out of nine studies) indicative of the tools' ability to differentiate self-esteem status between these groups. Generally, weight loss did result in increases in global self-esteem in the majority of studies. However, three studies that reported no change in weight status, did report improvements in self-worth, athletic competence, social acceptance and feelings/emotions.<sup>37-39</sup> These findings suggest a benefit of participation in interventions even in the absence of weight loss. This review identified that the domains most likely to improve as a result of intervention were appearance, social function and competence. No differences in the effect of obesity on self-esteem between children and adolescents were observed and the authors conclude evidence relating to ethnicity and also gender were lacking.

A more current systematic review and meta-analyses have been conducted by Murray et al.<sup>1</sup> In this review 13 studies were included representing 1157 overweight or obese children. The focus of this review was adolescents aged between 10 and 19 years who were enrolled into a multicomponent weight-management program where self-esteem had been measured. The studies included were similar to Go4Fun in that they had content relating to nutrition, physical activity and improving psychosocial outcomes. Of interest in this review are four studies that intervened in children in the 10-13 age range as they were identified in search strategy one and therefore applicable to Go4Fun.<sup>12, 28, 40, 41</sup> The self-esteem tool used was the SPPC supporting Lowry et al's<sup>35</sup> recommendation that multidimensional tools be used.

In a 2005 systematic review conducted by Butler et al.<sup>42</sup> the authors examined articles that reported on measures of 'self' in children and adolescents. Fourteen scales were included for further assessment. The most frequently reported measure was the Piers-Harris Children's self-concept scale (both versions). The authors report that studies that include the RSES tend to focus on adolescents rather than children and as such did not make the final list of included 14 scales – a factor to consider in this review. The authors concluded that currently there appears, both theoretically and psychometrically, to be an acceptance of multidimensionality with respect to the self, with the latest scales designed around such a notion. These findings are supported by the previously mentioned reviews by Lowry et al., Griffith et al. and Murray et al. where a focus on summarising studies of multiple competencies, particularly in relation to childhood obesity, appears to be the norm.

# Grey literature (other than systematic reviews)

Only one study meeting the inclusion criteria was found that was not identified from the original search strategy.<sup>43</sup> This multicomponent intervention was conducted in children aged 11-15 years, using the Rosenberg Self-Esteem Scale (RSES) for measuring pre-test and post-test self-esteem respectively. The study did not find any significant changes in self-esteem levels upon conclusion of the intervention. Another study of interest is the Children's Health and Activity Modification Program (C.H.A.M.P.)<sup>44</sup>, a fourweek day camp for obese children aged 8-14 years. Though a specific tool for measuring self-esteem was not implemented, the researchers conducted qualitative interviews with the children at the end of the camp to explore whether they felt increased levels of self-esteem, of which they reported improvements.

A recent report (2017) on the Australian 'Young Minds Matter' program details the validation of a new tool for measuring adolescent self-esteem (The Adolescent Self-Esteem Questionnaire (ASQ), a 13-item measure of global self-esteem.<sup>45</sup> The tool was used on 2964 11 to 17-year-olds and 304 school-aged students (14-17-year olds) and was compared against the RSES. The report found the tool was able to detect changes in self-esteem over two-time points and had similar validity and reliability to the RSES, with high internal consistency, construct validity and test-retest reliability. It did not perform significantly better than the RSES. The researchers advise however that the cross-sectional nature of the data means that the ASQ has not yet been validated for measuring changes in self-esteem outcomes as a result of interventions/programs.

# Summary of the validation studies that measured self-esteem (search strategy 2)

In the second search, nine validation studies were retrieved (English versions of the tools only). Studies were conducted in the UK (one study), US (five studies), and Australia (three studies) and validated tools used included RSES (one study), SDQ (two studies), Piers-Harris-22 (one study), and SPPC (one study). Alternative scales that came out of this search included Coopersmith Self Esteem (one study), State Self Esteem Scale (one study), and Self Perception Score (one study). Most validation studies were performed as a single testing occasion in a school or classroom setting (see Appendix 6).

Of the three validation studies performed in Australia, one was the alternative tool Self Perception Score (O'Dea 2009) and two were the SDQ-I<sup>9</sup>. The Self Perception Score was validated in males and females aged 11-14.5 years, from two schools in the same district in Sydney, Australia. Ethnicity and other demographics were not reported. Students completed the scoring independently during class time under the supervision of the researcher. Correlation scores were reported with SPPC, Beck Junior Depression, Speilberger State and Trait Anxiety and Eating Disorder Inventory. The Marsh reports were two papers based on the same study that assessed male and female children aged 10 years from five suburban public schools in Sydney, Australia. Participants were described as varied SES from working class to upper middle class (percentages and ethnicity were not reported). The authors reported on the convergent validity and factor analysis of the SDQ-I. The questionnaire was delivered in a classroom setting with researchers' guidance. The SDQ-I was administered to the children alongside the Pier-Harris Self-Concept Scale for Children (original version) and an older version of the SPPC called the Perceived Competence Scale for Children (PCSC). When comparing the three tools factor analysis and construct validity was tested. The PCSC and the SDQ-I performed better in factor analysis than the Piers-Harris, suggesting that these two scales more appropriately measure their intended factors. With regards to construct validity all three tools were closely correlated with other measures of self-esteem but the SDQ-I has stronger divergent validity in that it had a weaker correlation with measures of a separate construct from self-esteem.

The RSES was validated in the UK in children 12 to 19 years old (Bagley & Mallick 2001). The questionnaires were completed independently by students in a classroom setting across four schools in two working to middle class English counties (Hampshire and Yorkshire). Ethnicity data were not collected. The authors reported on internal reliability (Cronbach's alpha values) and construct validity for 12-13-year-old males and females.

The studies that were performed in the US included validation of the Piers-Harris<sup>46</sup>, Harter's SPPC<sup>47</sup>, CY-PSPP<sup>48</sup>, and two alternative tools, the Coopersmith Self Esteem Scale<sup>49</sup>, and the State Self Esteem Scale.<sup>50</sup> The Piers-Harris-1 was validated in students in grades 4 and 7, ages were not reported. Five schools participated from South-western US region and students were reported as 21% Hispanic, 75% Anglo-Saxon, and 4% other ethnic classification. The tool was administered in a self-report format. The author reports on convergent validity, discriminant validity, and internal consistency. Sensitivity to change was also reported on a separate cohort from the same five schools and were 60% Hispanic, 22% Anglo-Saxon and 17% Native American. The Harter's SPPC was validated in African-American females aged 12 years as a self-report in a group setting on the first day of a summer camp for at risk teens. The SPPC was compared with the RSES to establish stability of the construct assessment in an African-American population. The participants were generally from the Midwest US and were from lower SES, presented with multiple risky behaviours, and were from at-risk environments with high exposure to violence. Factor analysis, internal reliability, and convergent validity were reported. The CY-PSPP was validated as a self-report form during a physical education class in a school setting in 8 to 12-year olds. Schools were from predominantly middle-class neighbourhoods and the majority of the children were white (70%). The Coopersmith was a self-report tool that was validated in a school setting in the US state of Kentucky. The questionnaire was administered by teachers who were trained in the delivery of the questionnaire. Randomly selected children were assessed twice, two weeks apart. Ethnicity was not reported. The authors measured convergent validity, internal consistency and sensitivity to change. The State Self Esteem Scale was validated in 11 to 13-year olds from a Southeast Texas middle-school. Participants were reported as 42% white non-Hispanic, 57% African American, <1% Asian. The questionnaire was a self-report that was completed weekly during a six-week mandatory 'Success Program' class at the school which was designed to promote self-esteem and school success. The questionnaire was compared and correlated with RSES.

# Important characteristics of included self-esteem tools

All self-esteem tools reported in this review were designed for use in children/adolescents and have undergone some psychometric testing, just not necessarily in healthcare (intervention) settings. The selfesteem tools vary considerably in their length but are considered short form in that they can be completed in one sitting. Although the majority report a completion time of between 10 and 15 minutes this will vary based on levels of literacy, understanding and support required to complete the tool. Self-esteem tools specifically targeted at younger children may require assistance to complete. Appendix 7 summarises the self-esteem tools reported in this review and the number of items and subscales assessed by each tool. The RSES remains the shortest tool available to assess self-esteem. However, it is unidimensional and although it has been validated in children as young as eight, it is primarily used in adolescents. In terms of popularity Harter's SPP (children and adolescent versions) were used in 16 of the 36 included studies (search strategy 1). These findings are supported by the systematic review and meta-analysis of Murray et al. who reported Harter's SPP (children and adolescent versions) to be the most commonly used tool, as did the review by Lowry et al. The review by Butler et al.<sup>42</sup> reported the Piers-Harris self-concept scales to be the most frequently used tools, however this review is older and these findings not supported by the more recent reviews and the current search. Of the multidimensional tools used, scales by Harter benefit from being the shortest in length, are multidimensional in terms of outcomes measures and validated children and adolescent versions exist.

All tools identified in the search strategies are able to detect changes in self-esteem over time. However, there are some caveats to this statement. Changes in certain domains of self-esteem (such as body image) may occur before overall changes in self-esteem are observed. Therefore, self-esteem tools that assess various domains (multidimensional) provide a more useful assessment strategy.

All self-esteem tools included in this review have undergone some psychometric testing in children and/or adolescents. Many of the tools have been used in Australian settings, however not many of the tools have been validated in Australian children. Where possible we report on reliability, face validity, construct validity and content validity (a definition of these terms is included in the Glossary) of the tools. However, no one tool (long form) has been validated against a short-form version. In the section below are summaries of the tools used to measure self-esteem including commentary on their reliability and validity. This information is also summarised in Appendix 6.

#### Summary of self-esteem tools

#### Harter's Self-Perception Profile for Children (SPPC)

*Profile:* The SPPC instrument (8 – 13 years) includes five specific self-concept subscales: scholastic competence, athletic competence, social competence, physical appearance, and behavioural conduct. In addition, a separate, sixth subscale, covers global self-worth (or self-esteem). There are a total of 36 items, six for each subscale. Global self-worth is rated by its own set of items and scored separately. All subscales employ "structured alternative response format" i.e. yes/no.

*Cross-cultural considerations:* Specifically designed for American children and adolescents, in terms of the content of the domains, the structure of the instruments, and the question format, and therefore the SPPC is not necessarily applicable in other countries and cultures without validation.

Availability: Available online free of charge.

*Administrative burden*: The scale can be administered in groups (either small groups or classrooms) or individually. It is advisable to read all the items of the questionnaire to younger children (i.e. 3rd and 4th grades) or to children who have difficulty with English language comprehension.

**Online use:** This tool has not been validated for online use in an English-speaking population. Online surveys are not recommended to be completed by young children by themselves.

#### **Reliability and validity**

*Reliability*: Internal reliability for the SPPC has been assessed across four samples of children from the general population. Cronbach's alpha values reached acceptable levels on average, ranging from .74 to .83 for the SPPC (51).

Face validity: The SPPC meets this criterion because items directly ask about the concepts in question.

*Content validity*: According to the creator, the goal was to develop an instrument where the transparency of the content was so obvious that anyone could understand the intent.

*Construct validity*: The authors report that through the development of self-concept and global self-esteem subscales, embedded within a model of the determinants, correlates, and consequences of global self-esteem, the construct of the SPPC has met with empirical support<sup>23, 52</sup> and therefore demonstrated construct validity of the instrument.

*Convergent validity:* A comparison with Marsh et al.<sup>53, 54</sup> between subscales of similar content has shown evidence for convergent validity. The subscale of global self-worth correlated (0.56) with Marsh's general self-concept subscale of the SDQ-I questionnaire.

#### Rosenberg self-esteem scale (RSES)

**Profile:** The scale was introduced in 1965 in Rosenberg's study of a large sample of high school students in New York State in the US. The appeal of the Rosenberg scale derives from both its theoretical and its practical attributes. Furthermore, the conceptualisation of self-esteem underlying the development and construction of the RSES is close to that presented in psychological theories about the self and is also consistent with the layman's view of self-esteem.

*Administrative burden:* The popularity of the questionnaire may in part be related to its short length and ease of administration.

*Cross-cultural considerations:* Specifically designed for use with American teenagers, the scale has been widely translated and used in many countries, but not necessarily validated, e.g. Australia.

Availability: Freely available.

**Score interpretation**: The RSES, a 10-item scale, was originally presented as a Guttman scale, but it is typically administered using a four-point Likert-type response format ranging from strongly disagree to strongly agree. Higher scores indicate higher self-esteem.

# **Reliability and validity**

**Reliability:** The RSES presented high ratings in reliability areas in initial testing; internal consistency was 0.77, minimum coefficient of reproducibility was at least 0.90.<sup>55</sup> A number of independent studies using such samples as parents, men over 60, high school students, and civil servants, showed Cronbach's alpha coefficients ranging from 0.72 to 0.87 (all fairly high). Test-retest reliability for the two-week interval was calculated at 0.85, the seven-month interval was calculated at 0.63.<sup>56</sup>

*Validity*: Demonstrates face validity, construct validity and concurrent and predictive validity using known groups.

*Face validity:* Rosenberg<sup>22</sup> argues, for example, that self-concept is "not a collection but an organisation of parts, and components," which are hierarchically organised and interrelated. Therefore, a measure of the global evaluation of self-concept, an overall picture of positive or negative attitudes toward the self, is appropriate.

**Construct validity:** The RSES was investigated using item response theory. Factor analysis identified a single common factor, contrary to some previous studies that extracted separate self-confidence and self-depreciation factors. A unidimensional model for graded item responses was fit to the data. A model that constrained the 10 items to equal discrimination was contrasted with a model allowing the discriminations to be estimated freely. The test of significance indicated that the unconstrained model better fit the data – that is, the 10 items of the RSES are not equally discriminating and are differentially related to self-esteem. The pattern of functioning of the items was examined with respect to their content, and observations are offered with implications for validating and developing future personality instruments.

*Convergent validity:* The RSES correlates significantly with other measures of self-esteem, including the Coopersmith Self-Esteem Inventory. In addition, the RSES correlates in the predicted direction with measures of depression and anxiety.

Online application: RSES has not been specifically validated in an online setting.

# Piers-Harris children's self-concept scale 2 (Piers-Harris-2)

*Profile:* This original scale (unidimensional) was designed in the early 1960s to measure self-concept, a term that has been used interchangeably with self-esteem and self-regard. An updated version (multidimensional) of the scale exists (Piers-Harris-2) that contains 60 items and has been validated in children aged 7-18. In this document we use the term Piers-Harris when referring to the original unidimensional scale or Piers-Harris-2 when referring to the multidimensional version.

*Cross-cultural considerations:* Whereas the original Piers-Harris was evaluated in a homogenous sample of American public-school children, Piers-Harris-2 has been validated in 1387 ethnically representative districts.

Availability: At a cost and can only be administered by Allied Health or other Education Professionals.

*Administrative burden:* Test items are simple descriptive statements, written at a Year 2 reading level (i.e. '*I am a happy person*') with a yes/no response format. Although the scale has 60 items, the length of time to complete the questionnaire is usually between 10 and 15 minutes.

*Score interpretation:* The scale comprises 60 items, which cover six subscales. Two validity scales identify biased responding and the tendency to answer randomly. The updated version of the tool (Piers-Harris-2) provides a total score that reflects overall self-concept, plus subscale scores that permit more detailed interpretation.

# **Reliability and validity**

**Reliability:** Original authors reported good internal consistency ranging from Cronbach's alpha .74 to .81 for the six subscales and .91 for the total score of the Piers-Harris-2 across six age groups (7-8, 9-10, 11-12, 13-14, 15-16, 17-18 years old). Test-retest reliability for the original and revised Piers-Harris-2 in both normal and special populations, completed in the 1960s, 1970s, 1980s, reported reliability coefficients ranging from r=.71 to r=.88.

Face validity: Has demonstrated acceptable face validity.

*Content validity*: The items were written to maximise content validity and to test qualities that children reported liking or disliking about themselves.

**Construct validity**: Examination of the tool's structural characteristics was conducted by the authors. Both the original and the revised Piers-Harris-2 scale showed the domains are interrelated and measure separate aspects of overall self-concept.

*Convergent validity*: Concurrent data was collected to establish convergent validity and the extent to which the Piers-Harris-2 scale correlated with measures of similar psychological constructs.

**Online application:** This tool has not been validated in an online version.

# Marsh's Global self-esteem (Self-Description Questionnaire- SDQ-I)

*Profile:* Designed to measure multiple dimensions of self-concept for pre-adolescents. In particular, the scale measures self-perceptions relative to four non-academic areas (Physical Ability, Physical Appearance, Peer Relations, and Parent Relations) and three academic areas (Reading, Mathematics, and school in general), as well as a global perception of self.<sup>57</sup> This scale is the only widely used scale that was developed in Australia. The SDQ-I was originally constructed in order to empirically test the self-concept model proposed by Shavelson, Hubner and Stanton<sup>58</sup> with pre-adolescent children.

Cross-cultural considerations: Developed and validated for use in Australian classrooms for 8-12-year olds.

*Availability*: This scale is freely available. However, it is a condition of use that any data on this instrument that is collected will be made available in electronic form to the SELF Research Centre.

Administrative burden: On average this scale can take from 15-25 minutes to complete.

*Score interpretation:* Each question is answered through a 5-point true/false scale with 12 of the 76 questions worded negatively. However; as it appears that valid answers are not produced in pre-adolescent responses to negatively worded questions, these 12 items do not form part of the score.

#### **Reliability and validity**

*Factor Analysis:* The determination of how many factors (variables) a scale measures. Analysis has shown that 97% of the target factors had a correlation coefficient of above 0.3 and only 1% of non-target factors had a correlation coefficient of above 0.3(9). This demonstrates that the SDQ-I measures the domains it is stated to measure and does not measure alternative domains of self-esteem.

**Reliability:** The SDQ-I has shown good internal consistency across the domains, total academic, total nonacademic and general self-concept. However; general self-concept had lower reliability scores and larger standard deviations than the other two domains and therefore it remains important to consider each sample population's reliability especially with regards to general self-concept. When multiple sample populations were pooled together, SDQ-I offered good reliability with general self-concept yielding slightly lower reliability than total academic and total non-academic domains.

*Face validity:* The SDQ-I was specifically designed around the Shavelson model of self-concept and as such has been used to test the model. The items of the instrument appear to have good face validity.

*Construct validity:* Construct validity can be demonstrated through a combination of convergent and divergent validity. Marsh has shown that the SDQ-I has high convergent validity with other self-esteem tools (Piers Harris and The Perceived Competence Scale for Children) and low correlation with unrelated measures through multi-trait multi-method analysis.<sup>9</sup> Together this demonstrates that the SDQ-I does, in fact, measure multidimensional self-esteem.

**Convergent validity.:** When compared to Piers-Harris and PCS, the SDQ-I showed high convergent validity (r= 0.87) and low divergent validity (r= 0.38) demonstrating that the SDQ-I has good correlation with other measures of self-concept and poor correlation with measures that do not represent self-concept. When the academic self-concept scale was compared with academic achievement there was good correlation (r=0.36) this further supports good convergent validity for the subscales of the SDQ-I.

**Online Application:** Online format is possible via importing the instrument in e-form but has not been validated in an online setting.

#### Child and youth physical self-perception profile (CY-PSPP)

*Profile:* The CY-PSPP consists of 36 items about children's physical self-perception. Six domains are represented: global self-esteem, physical self-worth, sport competence, body attractiveness, physical strength and physical condition. The instrument is an adaptation of the adolescent PSPP.<sup>59</sup> Modifications to the PSPP to improve its utility in younger children resulted in the revised CY-PSPP<sup>60</sup> which was validated in younger children in 2005.<sup>48</sup>

*Cross-cultural considerations:* Developed in the US the scale has also been found to have good cross-cultural validity as similar findings have been observed in other countries including the UK and Canada.

Availability: Scale could not be retrieved so availability unclear.

*Administrative burden:* Original validation manual unavailable but the original validation study reported that the survey was completed in physical education classes which indicates administration is within 45 minutes.<sup>48</sup>

*Score interpretation:* Each item consists of two statements and a four-point structured alternative format to reduce the tendencies for socially desirable responses and approximately half of the items are reverse coded.

### **Reliability and validity**

The CY-PSPP was originally developed as the Physical Self-Perception Profile (PSPP) and the factorial validity and psychometric properties of the PSPP were originally demonstrated with college students.<sup>59</sup>

*Reliability*: Previous work<sup>60</sup> revealed high alpha reliability for the CY-PSPP scales (range: 0.77–0.91) in a high school sample. In another study, which sought validation of the scales with a younger cohort (elementary school) employed confirmatory factor analyses to evaluate the a priori dimensions for the evidence of convergent and discriminant validity of the CY-PSPP scales.<sup>48</sup> Factor loadings in the measurement model revealed no substantial areas of concern. Median loadings for the total sample, boys subsample and girls subsample were 0.69 (range 0.41–0.82), 0.67 (range 0.32–0.85), and 0.70 (range 0.43–0.83). Most latent factor intercorrelations were reasonably moderate although some were notably large. In summary, these findings suggest an adequate fit for the CY-PSPP measurement model to these data and reasonable psychometric attributes.

*Face validity:* The instrument seems to have good face validity. For instance, in order to provide a more complete and powerful subdomain coverage, product, process and perceived confidence items were included by the authors. Examples of these in the sports competence subscale were "some people are good at most sports" (product), "some people seem to learn sports skills very slowly" (process), and "some people feel very confident when it comes to playing sports' (perceived confidence).

*Construct validity:* The psychometric properties and factor structure of the CY-PSPP were evaluated to provide support for the reliability and discriminant validity of the subscales and the total scale. Statistics provided by the confirmatory factor analysis clearly indicate that the observed data are consistent with the constraints specified by the model. The scores on the subscale items therefore are adequately described by four independent but correlated latent factors.

*Convergent validity:* The convergent validity of the scale is generally favourable and supported by relationships found by comparing the CY-PSPP domains with other measures of physical activity and physical fitness (i.e. Harter).

**Online availability:** Online surveys are not recommended to be administered to young children by themselves. A parent or group leader should assist with completing surveys online with the child present.

#### Beck Youth Inventory II

**Profile:** The new Beck Youth Inventories<sup>™</sup> - Second Edition for Children and Adolescents are designed for children and adolescents aged seven through 18 years. Five self-report inventories (20 questions each) can be used separately or in combination to assess symptoms of depression, anxiety, anger, disruptive behaviour and self-concept. Self-concept is one of the separate inventories. Children and adolescents describe how frequently the statement has been true for them during the past two weeks, including today. The self-concept inventory measures cognitions of competence, potency, and positive self-worth.

*Cross-cultural considerations:* The psychometric properties of the Beck Youth Inventories were studied in a nationally drawn, standardisation sample of 800 children<sup>61</sup>, therefore, representative of the US population for age, gender, ethnicity and socioeconomic status.

Availability: Not freely available.

Administrative burden: 5 minutes per inventory.

*Score interpretation:* The BYI-II scoring is available on Pearson's web-based platform for scoring and reporting.

#### **Reliability and validity**

*Reliability*: The internal consistencies of the five inventories were all high (coefficients >.84), and the test-retest reliability correlations for the five inventories were all large.

*Face validity*: The BYI-II appears to have good validity with all items referring to one aspect of emotional and social impairment (i.e. depression, anger, self-concept).

**Convergent validity:** The Conners-Wells Adolescent Self-report Scale<sup>62</sup> was used to investigate the convergent validity of the disruptive behaviour inventory in a subsample of 108 children. The scores on the Conner-Wells Conduct Problems scale were well correlated with the scores on the disruptive behaviour inventory (r = .70, p < .001). Furthermore, the scores on the anger inventory were also well correlated with the scores on the Conners-Wells ADHD Index (r = .73, p < .001).

**Construct validity:** The construct validity of the five inventories has been assessed against a variety of different scales in subsamples of the standardisation sample. For example, the scores on the Beck Youth Depression Inventory were highly correlated (r = .72, p < .001) with the scores on the Children's Depression Inventory<sup>63</sup> in a subsample of 128 children.

Online availability: Possible at a cost.

# Racial Identity and Self-Esteem of Children (IRISE-C)

*Profile:* The IRISE\_C was specifically developed to examine the sense of self among Indigenous Australian children, including self-concept, self-esteem and racial identity. Four subscales were constructed; knowledge of Aboriginal culture, knowledge of racial identity, salience of Aboriginal culture, salience of racial identity.

The IRISE-C is currently reported as a static measure of identity and self-esteem therefore does not take into account that self-esteem can be influenced by significant others. Therefore, the IRISE\_C provides a snapshot of identity and self-esteem at one-time point. It was developed by the first author, an Aboriginal researcher.

*Cross-cultural considerations:* The IRISE-C has been developed for use in Australian Aboriginal children and adolescents. However, the authors do caution that results may not be generalised to Aboriginal children living in remote areas as these children were not included in the study. Additionally, the sample used in the validation study were from regional and rural sites that may not have been large enough to adequately describe the full diversity of identity and self-esteem among Aboriginal children. The children included in the sample were from Western Australia, where there are approximately 69,665 Aboriginal people, and more than 250 Aboriginal communities (not including Aboriginal residents in metropolitan areas).<sup>64</sup> Within these communities are distinct Aboriginal groups with a complex and rich system of language. The authors caution that the items contained in the IRISE\_C may be different for other Aboriginal groups not yet researched. Another challenge reported by the authors is that some children may identify as being from more than one Aboriginal group, or may identify with both Aboriginal and non-Aboriginal heritage.

#### Availability: Unclear.

*Administrative burden:* Paper survey with verbal administration ideally by an Aboriginal research assistant/relevant person from the Aboriginal community. The scoring key is A4 in size and contains the Likert scale with images of "smiley faces." The pre-pilot and subsequent pilot of the scoring key revealed that school children, who are the target group were very familiar with the use of smiley faces on the IRISE\_C scoring chart as it was an acceptable, recognisable symbol.

The authors emphasis the following elements to be vital in order to achieve authentic and reliable data when using the IRISE-C:

- Recruiting Aboriginal community research assistants from the local area of the study site and training them in interviewing and survey techniques
- Surveys being administered in a one-on-one basis with each participant
- Verbally asking each individual survey item of participants and recording the response accordingly on the paper survey
- Using a visual (image) scorecard.

#### Score interpretation: TBC.

#### Reliability and validity

The questionnaire is validated in children aged 6-13 years. The 71-item IRISE\_C was originally piloted in 35 urban Aboriginal children (8-12 years). The current form of the IRISE\_C contains 40 items and was evaluated in 229 children from rural, remote and urban regions of Western Australia. Statistically, the IRISE\_C is a valid and reliable instrument that captures identity and self-esteem for Australian Aboriginal children 8–12 years of age.

*Reliability:* Factor 1 (Aboriginal culture) had a Cronbach's alpha of 0.835; Factor 2 (racial identity) had a Cronbach's alpha of 0.800 (Table 2). Correlation between factors 1 and 2 (knowledge): 0.628

*Face validity*: A series of consultations, negotiations and reviews from other Aboriginal community members, Aboriginal teachers, and professionals ensured that the concepts contained in the IRISE\_C were culturally sound and acceptable. In following such a protocol, a high level of Aboriginal ownership has been

encouraged and, in doing so, the development of the IRISE\_C inventory with culturally safe and secure procedures has ensured authentic and valid results. Further, the concepts captured by the instrument have been deemed of value and acceptability by the Aboriginal carers who provided a high response rate.

Convergent validity: Not yet assessed.

*Construct validity:* The confirmatory factor analysis has shown that the 4 subscales: 1. knowledge of identity, 2. salience of identity, 3. knowledge of culture and 4. Salience of culture represent "good" and "acceptable" fitting models. This demonstrates that each subscale effectively captures a single, consistent underlying factor or concept.

Online application: Possible to have an e-form survey administered by an Aboriginal Research Assistant.

# Considerations of self-esteem tool properties (search strategy 1)

All identified tools show acceptable reliability and validity (Appendix 7) although the IRISE-C has not been validated in an intervention setting. However, tools specifically validated in the age range associated with Go4Fun (7-13 years) include the SPPC and the SDQ-I. Tools that have not been validated specifically in 7-10-year age group include the RSES. The other tools covered age ranges between 7 and 18.

All of the commonly used, validated tools have demonstrated acceptable validity and reliability, however; not necessarily in all types of validity testing. The problem arises in the poor study design of these validation studies and their age, therefore making recommendations, based on the psychometric properties, on an appropriate tool for this setting and population group challenging. All of the identified tools were created and validated more than 15 years ago, raising the question of whether the results are applicable to the populations in which they are used today. The reliability of the tool, for one, is population specific and as such it is strongly recommended that the Go4Fun team test the reliability of their chosen tool in their population group.

With regards to other considerations when selecting a tool, there is some differentiation between the tools. Both the RSES and SPPC are both freely available and have comparatively low administrative burden (being 10 and 36 items respectively) compared to BYI-II, which is 100 items. None of the seven identified tools have been validated online. Although, it is increasingly easy to translate tools from paper to online form this does not mean that the validity is as transferrable. An online tool is easy to administer and can be sent directly to participants to complete at a time convenient to them. However, it is recommended that the validity and reliability of the tool be tested in an online format first.

Cultural considerations are imperative; with a vast range of differences in language use and communication style between cultures, it is easy to see why tools that have been designed in one population will not necessarily transfer well to another. Common issues include differences in reading ability, cultural norms dictating responses (some cultures will only respond in the affirmative so as to not disagree with the interviewer) and difference in ethnicity between interviewer and respondent.<sup>65</sup> When choosing a tool, it is important to consider and seek expert advice on whether the wording and administration of the tool would need to be adapted to suit the needs of children from diverse backgrounds. Once the tool is adapted it would need to be thoroughly tested in the group of interest.

# Question 2: What validated short-form self-esteem questionnaires have been evaluated for their usability and acceptability in vulnerable populations?

#### Summary of included studies in low SES or vulnerable populations

Of the multicomponent weight-management interventions that assessed self-esteem in overweight and obese children and adolescents identified in search strategy one, six were conducted in populations considered vulnerable or low SES<sup>2-6</sup> (Appendix 8). There were two RCTs, two pre-test post-test studies and two non-randomised experimental trials identified, conducted across Australia (one study), the UK (one study) and US (four studies), all in community-based settings. Studies included between 21 and 347 overweight and obese children and adolescents, with ages ranging from eight to 17 years. Interventions varied in duration from 10 weeks up to six months, with 10 weeks and 12 weeks being the most common durations.

Across the studies, two different self-esteem measurement tools were identified: the RSES and CY-PSPP.

The CY-PSPP was used in populations experiencing or exposed to poverty. The RSES was used in participants with a diagnosed learning disability, culturally diverse populations of high social disadvantage, as well as low-income populations. No identified studies assessed self-esteem following multicomponent weight-management interventions in overweight or obese Aboriginal or Torres Strait Islander children.

Three studies reported on global self-esteem only, and the most commonly used tool to report global selfesteem was the RSES. Three studies reported on global self-esteem in combination with other domains of self-esteem, such as physical self-concept. Overall, five studies identified an improvement in self-esteem over the course of the intervention, and one saw no change. Improvements were seen in both global selfesteem and physical self-esteem.

There was no tool that stood out above the rest as being more likely to identify a change in self-esteem. Whether or not a change was detected is more likely to be due to the intervention program, as opposed to the tool selected to measure self-esteem. As no study commented directly on acceptability or usability of the self-esteem tool used, data was collected to assess the percentage of completers of the intervention who also completed the self-esteem tool. Completion rates for the Rosenberg Self-Esteem Scale (RSES) were at or close to 100%, where it could be identified. The CY-PSPP had the poorest completion rate of 45.8%. Both tools were completed by the child or adolescent participants, with assistance if required for completing the RSES. This is not to say that assistance shouldn't be offered, particularly to vulnerable or CALD populations, if administering the other tools, it just wasn't offered in the identified studies.

# Summary of included validity studies in low SES or vulnerable populations

Only one of the included validation studies reported on the use of Harter's SPPC in a vulnerable population.<sup>47</sup> The participants were 12-year-old African-American girls reportedly from lower SES, presented with multiple risky behaviours, and were from at-risk environments with high exposure to violence. The study measured factor analysis, internal reliability, and convergent validity of the Harter's SPPC in this cohort. Moderate to poor validity was reported, with the authors stating that "subscales of the SPPC should be cautiously interpreted when used with different ethnic and racial groups".

The Australian validation studies do not adequately report on ethnicity or SES of participants. The validation studies from the US have included more detailed information on their sample populations. Generally, the populations in the US studies are a mixture of ethnicities including Hispanic, non-Hispanic white and African-American children. One study that assessed the sensitivity to change of the Piers-Harris scale<sup>46</sup> included 17% children from Native American background, however did not report their results separately. Some questionnaires have been validated in vulnerable populations with specific disease-states but were not considered relevant for this review.

Search strategy two identified no tools administered online that have been validated. Therefore, although self-esteem tools are used in the online environment there has been no attempt to validate for the online setting. Our recommendation would be to undertake an online validation. The two papers that did report on the validation of a self-esteem tool in an online environment were excluded because of the age of the participants (13-17 years) or location of validation study (Taiwan).

None of the self-esteem tools identified through search strategy one have been specifically validated in Aboriginal children and adolescents. Through search strategy two the IRISE\_C scale was identified. This tool appears to show good validity but has not been used in pre-test or post-test conditions so its ability to show change is currently unknown.

# Grey literature (findings): question 2

Of the grey literature retrieved for question 2, no specifically relevant tools surfaced, however, two studies are worth mentioning. The first<sup>66</sup> reports on a cross-cultural analysis of the SDQ-II (for adolescents in grades 7-11). The study compares Indigenous and non-Indigenous student samples. Overall, the report found that the SDQ-II results for Indigenous and non-Indigenous participants were highly correlated and therefore is appropriate to administer in both populations. Of note, this refers to SDQ-II (adolescents) and is therefore not generalised to the SDQ-I (child version).

The second study of interest is the Australian Government program '*Kids Matter*', a nationwide initiative to improve mental health outcomes in pre- and primary-school-aged children. The program does not use a specific self-esteem measurement tool, however they have developed a specific, comprehensive guide to conducting health interventions such as *Kids Matter* in Aboriginal and Torres Strait Islander children and communities.<sup>67</sup> Some of the main recommendations are:

- Not positioning every Aboriginal and Torres Strait Islander person into one or even two groups. While understanding the similarities are important, it must be remembered that every person (and community) is unique
- Welcome or acknowledgement of country preferably at the beginning of a session
- Cultural competence embracing local cultural knowledge
- Building community capacity particularly in remote communities, involving local elders and other leaders is crucial
- Acknowledging lack of knowledge and asking questions. It is however important to consider cultural sensitivity in the way questions are asked
- Taking time to unpack meanings of mental health and wellbeing in these communities and contexts, especially considering language, meaning and translation (i.e. 'mental' health can have negative connotations, 'well spirit' may be better received)
- Participants may need a considerable amount of time to process and think about their responses, particularly if English isn't their first language
- Parent involvement has been shown to improve student achievement, and it is important to recognise the role of the extended family and the community in the upbringing of children in these communities, which can vary from typical Western family-like structures.

For further guidelines from the NHMRC for working with Aboriginal and Torres Strait Islander populations refer to Appendix 9.

# Gaps in the evidence

- 1. There is a lack of intervention studies with self-esteem as primary outcomes. As such, the reporting of self-esteem data is limited
- 2. Validation studies of self-esteem tools are lacking in Australian children and adolescents aged between seven and 13 years
- 3. Of the validation studies found, only one was conducted within the past 10 years and this study focused on the validation of the SPPC in African-American adolescent girls. None of the tools have been validated in populations that would reflect modern day children in the Go4Fun Program.
- 4. The self-esteem tool we propose for use in Go4Fun, the SPPC, has not been the primary focus of a validation study in Australia. However, Marsh & Holmes<sup>9</sup> demonstrated good validity when comparing the SPPC to the SDQ-I in Australian children. It has been translated into a number of languages (for example Spanish, Dutch and Italian) and validated in vulnerable populations outside of Australia
- 5. There is a paucity of validation studies of self-esteem tools in an online setting. The Piers-Harris-2 has been validated online, but only in Taiwan
- 6. This review highlights the need for psychometric testing of self-esteem measures in Australian children and adolescents both in a face-to-face and online setting. The Go4Fun program offers an intervention that can be used constructively to achieve this aim
- 7. Guidance on whether self-esteem tools can be child administered or in association with a parent/carer (especially in the younger age group) also need to be considered.

# Discussion/synthesis of findings

Overall, 23 intervention studies identified an improvement in self-esteem over the course of the intervention, 11 studies found no change, one saw a reduction in self-esteem, and in one study it was unclear whether there was a change in self-esteem. Referral to the original manual/publications for these tools (where possible) provided details of reliability and validity – some in more detail than others. Where relevant data was available good reliability and validity was reported for all included self-esteem tools.

Typically, these tools have remained relatively unchanged since their development with researchers tending to validate the tools in different languages, different cohorts of children and ages rather than develop alterations to the actual measurement construct of the tools. Changes of note included the RSES, which has a shortened version (five items compared to 10) but its validity has not been explicitly reported. The RSES has also been adapted so that the scaled response has been replaced with typographical emoticons<sup>68</sup> which may enhance its application in younger children, however the validity and reliability of this revised version do not appear to have been tested. The Piers-Harris questionnaire was refined in 2001, modified to a 60item guestionnaire from 80 items and validated in an ethnically diverse and mixed SES group to become the Piers-Harris-2. However, it is not free to use and it is recommended that interpretation is by a psychologist, therefore it is a more clinically focused instrument. The CY-PSPP tool was revised version from PSPP, however its focus is on physical aspects of self-esteem and as such has limited use in terms of assessment of global self-esteem. The Harter's SPPC instrument is part of an age-graded, developmental battery only appropriate for US grades three through six (typically ages seven through to 12), although it has been used successfully with middle school students (grades seven and eight, typically ages 12 to 14). The tool has not been updated intentionally as the developers suggest that shortening the tool will impact the subscales and their usefulness. Of note, if other age ranges are to be considered in the Go4Fun program, a Pictorial Scale of Perceived Competence and Social Acceptance for Younger Children (ages four to seven) has been developed<sup>69</sup> and an adolescent version (ages 13-18) is also available and free to use. The only included tool to be developed and validated in Australia (excluding grey literature and therefore the ASQ) is Marsh's SDQ-I. We have also highlighted the IRISE\_C as a tool that has been developed and validated in Aboriginal children to measure racial identity and self-esteem. Although its validity and reliability are fair, it has not been tested in pre- or post-intervention settings and therefore its ability to detect change in self-esteem is currently unknown. All the tools included in the review (including the IRISE\_C) are considered short-form and can be completed by the child/adolescent on their own or with a parent/carer in one sitting. Tools requiring parental/carer assistance include the Harter's SPPC and recommendations on using the IRISE C include involving an Aboriginal care worker.

Four of the tools (search strategy one) accounted for only seven of the included 36 studies. The studies that used the SDQ-I both reported on the Youth Fit for Life program, which had a focus on physical activity. Only two scales of the SDQ-I were used (general scale and physical survey scale). It is not clear if validation of these subscales has been undertaken. The BYI-II was used twice. One study was reported as an efficacy trial only had a sample size of 17 (15 completed) and the second study recruited participants with at least one co-morbidity. Both studies only used three subscales: self-concept inventory, depression inventory and anxiety inventory. The Piers-Harris-2 and the CY-PSPP were only used once. Marsh's SDQ-I was only used in three studies but is an Australian tool and as such has been validated in Australian children. The RSES (13 studies) and the Harter (22 studies) tools (SPPC and SPPA) were the most frequently used tools.

There appears to be an increased utilisation of self-esteem tools that are multi-dimensional rather than unidimensional. Multidimensional tools are made up of subscales that measure constructs of self-esteem

that contribute to global self-esteem. It is reasonably common for a weight-management intervention to not impact global self-esteem but to demonstrate beneficial improvements in subscales. Constructs of selfesteem such as physical appearance and social competence have been reported to change before global self-esteem and therefore are increasingly important to consider in the adoption of a self-esteem tool. All tools considered in this review are classified as multidimensional except for RSES and the original Piers-Harris tool.

Search strategy two identified studies that have subsequently (post original validation studies) validated the included self-esteem tools identified in search one in children and adolescents. No date restriction was placed on this search but included studies were limited to Australia, New Zealand, US, Canada and Europe and children who were overweight or obese. Of the nine validation studies retrieved, the majority were performed in a single testing occasion in a school or classroom setting.

Evidence regarding the usability/acceptability of the included self-esteem tools is scarce outside of the initial validation studies. Of the 36 studies (search strategy one) that reported measuring self-esteem, none commented on their usability. We also examined completion rates of questionnaires as an indication of burden. Seventeen of the 36 included studies did not report completion rates of self-esteem measures, likely a result of self-esteem being a secondary outcome measure. Of those studies that reported an overall analysis of completion for the intervention, completion rates for measures of self-esteem were generally lower.
# Conclusion / Recommendations

Low self-esteem is a complication of paediatric overweight and obesity and is associated with adverse consequences. Although fluctuations in self-esteem are a feature of growing up, children who are overweight are reported to be at risk for greater declines in self-esteem as they enter early adolescence when compared to their healthy weight counterparts.<sup>10</sup> Multicomponent weight management programs that target self-esteem can induce weight loss as well as positively impacting on self-esteem.<sup>1</sup>

In making recommendations for identification of a short-form questionnaire for use in Go4Fun a number of considerations were taken into account. These included the reliability and validity of the tool, use in Australian children and adolescents, ability to measure changes in self-esteem and usability and acceptability in vulnerable populations. If preferable to use a tool validated in Australian children and adolescent Self-Esteem Questionnaire (ASQ) and Marsh's Self-Description Questionnaire (SDQ-I) should be considered. Limitations, however, include lack of validation in the 7-10-year age range, in an online setting and also in Aboriginal children. Furthermore, both tools have not been highly used in intervention studies. Although both tools show good validity and reliability, they do not outperform the Rosenberg Self-Esteem Scale (RSES) and the ASQ has not been validated for measuring change in self-esteem.

In considering validation of a self-esteem measure in Aboriginal and Torres Strait Islander populations, we recommend that the NHMRC guidelines for ethical conduct in Aboriginal and Torres Strait Islander Health Research<sup>70</sup> be considered. These guidelines provide a process for validation that respectfully engages the Aboriginal and Torres Strait Islander population in the research and provides capacities, outcomes or opportunities that are of benefit and of value to them. This involves consideration of:

- Consultation with the community so that participating communities have had equal and respectful
  input into the development of the project and how it will be implemented and have understood and
  expressed satisfaction with the project, its potential benefits and how study findings will be
  disseminated
- Values and cultures being respectfully considered in the development of the self-esteem tool in addition to consideration of differences in values, norms and aspirations
- Attention being given to involving the community in decisions regarding the project and that local structures are acknowledged and used in these processes
- Recognition of different people's input into the project
- Information provided as a result of the project being understood and usable in decision-making by the participating communities.

Based on its popularity, validity and reliability, use in behavioural intervention studies and age-specific nature the SPPC appears to be suitable for Go4Fun. Shorter than most tools at 36 items, the tool is freely available, has scales for younger and older age ranges and also parent and teacher proxy versions. Restrictions are that it has not been the direct focus of a validation study for use in Australian populations or in an online environment – however no identified tool meets all requirements.

Based on evidence and expert opinion we are making the following recommendations to the MOH for the inclusion of a validated measure of self-esteem in the Go4Fun program:

#### **Objective recommendations**

- 1. All identified tools have been validated in children and/or adolescents and show good reliability and validity
- 2. Based on ease of use, and the understanding that RSES is already used in Go4Fun, keeping the RSES is a consideration, but this review identified that there are more suitable, multidimensional tools
- 3. A thorough review of the literature points to adoption of a multidimensional self-esteem tool (which excludes RSES). In the absence of changes in global self-esteem, multidimensional tools can be helpful in examining sub-components of self-esteem
- 4. After consideration of the available evidence we propose adopting the Harter SPPC questionnaire into the Go4Fun program. This tool appears to be appropriate for both the younger and older ages that enrol into the program eliminating the challenges associated with two different tools
- 5. Whichever tool is used it is recommended that its reliability be tested in a Go4Fun cohort
- For the younger age group (7-10 years) we recommend parental/carer support in completion of the SPPC. This would be the case for the majority of validated self-esteem tools selected and not specific to the SPPC
- 7. As no tools have been validated in an online environment we would recommend adoption of the SPCC for online use and also validation in an online setting
- 8. Aboriginal children require the presence of an Aboriginal staff member to support completion of the survey. This staff member should preferably be from the children's community and preferably occupy the lead position in tool administration and be an integral member of the team
- 9. We would not recommend the IRISE-C. Although its validity and reliability are fair, it has not been tested in pre- or post-intervention settings and therefore its ability to measure change in self-esteem is currently unsubstantiated. Furthermore, we question the relevance of this measure to programs for weight reduction unless the program also aimed to improve racial identity.

#### **Expert opinion**

- Prior to adoption of the Harter SPPC (or any tool) we recommend a validation of its use in Australian children and adolescents. The Go4Fun program provides the ideal opportunity. For details on how to achieve this please see Appendix 10
- 2. We see no issues with adopting the SPPC in Aboriginal Go4Fun, however it would need to be translated into the appropriate language, in consultation with the specific communities. However, there are aspects of identifying as an Aboriginal person that would be integral to self-esteem that are not captured in this tool. This would require a study in itself, as the review has not identified any available self-esteem tools that are specific to Aboriginal people
- 3. The use of one tool for the age range 7-13 years is also beneficial in terms of pre- or post-change as the children age. Having to change tools as children move from one bracket (e.g. 7-10 years) to another (e.g. 11-13 years) has implications for evaluation
- 4. The SPPC has been translated into a number of languages (for example Spanish, Dutch and Italian) and validated in vulnerable populations outside of Australia thus indicating its potential for use in Australia
- 5. While the SPPC is the best available tool for this purpose, we realise that the length of the SPPC may be a barrier to its use. To reduce the time associated with completion of the SPPC, selecting

subscales previously shown to be particularly relevant to obesity in children and adolescents may be considered as an alternative. The following subscales most frequently reported to change in children and adolescents after participation in behavioural lifestyle interventions are physical appearance and athletic competence

- 6. When validating self-esteem measures in Aboriginal and Torres Strait Islander populations we recommend that the NHMRC guidelines for ethical conduct in Aboriginal and Torres Strait Islander Health Research be considered (Appendix 9). Referral to the 'Cultural adaptation of the Go4Fun program for delivery with Aboriginal communities in NSW' (pages 9 and 10) is also recommended
- 7. Relying on a single form of self-esteem measurement may not be sufficient. Clinically, multiple methods to gauge one's self-esteem are used including: interviews, behavioural observation, and ratings by others (e.g., teachers, counsellors, medical professionals, parents).

# References

- 1. Murray M, Dordevic AL, Bonham MP. The Impact of Multicomponent Weight Management Interventions on Self-Esteem in Overweight and Obese Adolescents: A Systematic Review and Meta-Analysis. Journal of Pediatric Psychology. 2016:1-16.
- 2. Annesi JJ, Faigenbaum AD, Westcott WL, Smith AE. Relations of self-appraisal and mood changes with voluntary physical activity changes in African American preadolescents in an after-school care intervention. Journal of Sports Science and Medicine. 2008;7(2):260-8.
- 3. Annesi JJ, Faigenbaum AD, Westcott WL, Smith AE, Unruh JL, Hamilton FG. Effects of the Youth Fit For Life protocol on physiological, mood, self-appraisal, and voluntary physical activity changes in African American preadolescents: Contrasting after-school care and physical education formats. International Journal of Clinical and Health Psychology. 2007;7(3):641-59.
- 4. Christison AL, Evans TA, Bleess BB, Wang H, Aldag JC, Binns HJ. Exergaming for health: A randomized study of community-based exergaming curriculum in pediatric weight management. Games for Health. 2016;5(6):413-21.
- 5. Nobles J, Radley D, Dimitri P, Sharman K. Psychosocial Interventions in the Treatment of Severe Adolescent Obesity: The SHINE Program. Journal of Adolescent Health. 2016;59:523-9.
- 6. Tirlea L, Truby H, Haines TP. Pragmatic, Randomized Controlled Trials of the Girls on the Go! Program to Improve Self-Esteem in Girls. American Journal of Health Promotion. 2016;30(4):231-41.
- 7. Weintraub DL, Tirumalai EC, Haydel F, Fujimoto M, Fulton JE, Robinson TN. Team Sports for Overweight Children The Stanford Sports to Prevent Obesity Randomized Trial (SPORT). Archives of Pediatric and Adolescent Medicine. 2008;162(3):232-7.
- 8. Kickett-Tucker CS, Christensen D, Lawrence D, Zubrick SR, Johnson DJ, Stanley F. Development and validation of the Australian Aboriginal racial identity and self-esteem survey for 8-12 year old children (IRISE\_C). Int J Equity Health. 2015;14:103.
- 9. Marsh HW, Holmes IW. Multidimensional self-concepts: Construct validation of responses by children. American Educational Research Journal. 1990;27(1):89-117.
- 10. Strauss RS. Childhood obesity and self-esteem. Pediatrics. 2000;105(1):e15.
- 11. Kolotourou M, Radley D, Gammon C, Smith L, Chadwick P, Sacher PM. Long-Term Outcomes following the MEND 7–13 Child Weight Management Program. Childhood Obesity. 2015;11(3):325-30.
- 12. Sacher PM, Kolotourou M, Chadwick PM, Cole TJ, Lawson MS, Lucas A, et al. Randomized Controlled Trial of the MEND Program: A Family-based Community Intervention for Childhood Obesity. Childhood Obesity. 2010;18(1):S62-S8.
- 13. Hardy LL, Mihrshahi S, Gale J, Nguyen B, Baur LA, O'Hara BJ. Translational research: are community-based child obesity treatment programs scalable? BMC Public Health. 2015;15:652.
- 14. Welsby D, Nguyen B, O'Hara BJ, Innes-Hughes C, Bauman A, Hardy LL. Process evaluation of an up-scaled community based child obesity treatment program: NSW Go4Fun(R). BMC Public Health. 2014;14:140.
- 15. Khanal S, Welsby D, Lloyd B, Innes-Hughes C, Lukeis S, Rissel C. Effectiveness of a once per week delivery of a family-based childhood obesity intervention: a cluster randomised controlled trial. Pediatric Obesity. 2016;11(6):475-83.
- 16. NSW Office of Preventive Health. Go4Fun. Sydney2016.
- 17. Innes-Hughes C HL, Radvan D, , C R. Cultural adaptation of the Go4Fun® program for delivery with Aboriginal communities in NSW. Sydney; 2016.
- 18. Flood V GJ, Gifford J, Tuner N, Hardy L, Evidence on existing, validated short-form survey instruments for children's diet, physical activity, and sedentary behaviour: an Evidence Check review brokered by the Sax Institute www.saxinstitute.org.au; 2016.

- 19. Mann M, Hosman, C.M.H., Schaalma, H.P., de Vries, N.K., . Self-esteem in a broad-spectrum approach for mental health promotion. Health Education Research 2004(19):357-72.
- 20. Rodgers RF, Melioli, T., Laconi, S., Bui, E., Chabrol, H. Internet addiction symptoms, disordered eating, and body image avoidance. Cyberpsychology, Behavior, and Social Networking. 2013(16):56-60.
- 21. Silvia C. Assessing body image from a parental and child's perspective. International Section A: Humanities and Social Sciences: Dissertation Abstracts. 2014;74.
- 22. Rosenberg M. Conceiving the self. New York: Basic Books; 1979.
- 23. Harter S. The Construction of the Self. New York: Guildford Press; 1999.
- 24. Jelalian E, Sato A, Hart CN. The Effect of Group-Based Weight-Control Intervention on Adolescent Psychosocial Outcomes: Perceived Peer Rejection, Social Anxiety, and Self-Concept. Children's Health Care. 2011;40(3):197-211.
- 25. Iloyd-Richardson EE, Jelalian E, Sato AF, Hart CN, Mehlenbeck RS, Wing RR. Two-Year Follow-up of an Adolescent Behavioral Weight Control Intervention. Pediatrics. 2012;130(2):e281-e8.
- 26. Mehlenbeck RS, Jelalian E, Lloyd-Richardson E, Hart CN. Effects of behavioral weight control intervention on binge eating symptoms among overweight adolescents. Psychology in the Schools. 2009;46(8):776-86.
- 27. Murdoch M, Payne N, Samani-Radia D, Rosen-Webb J, Walker L, Howe M, et al. Family-based behavioural management of childhood obesity; service evaluation of a group programme run in a community setting in the United Kingdom. European Journal of Clinical Nutrition. 2011;65:764-7.
- 28. Danielsen YS, Nordhus IH, Juliusson PB, Maehle M, Pallesen S. Effect of a family-based cognitive behavioural intervention on body mass index, self-esteem and symptoms of depression in children with obesity (aged 7-13): a randomised waiting list controlled trial. Obesity Research & Clinical Practice. 2013;7(2):e116-e28.
- 29. Christison A, Khan HA. Exergaming for health: a community-based pediatric weight management program using active video gaming. Clinical Pediatrics. 2012;51(4):382-8.
- 30. Incledon E, Gerner B, Hay M, Brennan L, Wake M. Psychosocial predictors of 4-year BMI change in overweight and obese children in primary care. Obesity (19307381). 2013;21(3):E262-70.
- Watson PM, Dugdill L, Pickering K, Owen S, Hargreaves J, Staniford LJ, et al. Service evaluation of the GOALS family-based childhood obesity treatment intervention during the first 3 years of implementation. BMJ Open. 2015;5:e006519.
- 32. Wong WW, Abrams SH, Mikhail C, Terrazas NL, Wilson TA, Arceo D, et al. An Innovative Summer Camp Program Improves Weight and Self-esteem in Obese Children. Journal of Pediatric Gastroenterology and Nutrition. 2009;49:493-7.
- 33. Quinlan NP, Kolotkin RL, Fuemmeler BF, Costanzo PR. Psychosocial outcomes in a weight loss camp for overweight youth. International Journal of Pediatric Obesity. 2009;4:134-42.
- 34. Jacobson D, Mazurek Melnyk B. A Primary Care Healthy Choices Intervention Program for Overweight and Obese School-Age Children and Their Parents. Journal of Pediatric Health Care. 2012;26:126-38.
- 35. Lowry KW, Sallinen BJ, Janicke DM. The Effects of Weight Management Programs on Self-Esteem in Pediatric Overweight Populations. Journal of Pediatric Psychology. 2007;32(10):1179-95.
- 36. Giffiths LJ, Parsons TJ, Hill AJ. Self-esteem and quality of life in obese children and adolescents: A systematic review. International Journal of Pediatric Obesity. 2010;5:282-304.
- 37. Brehm BJ, Rourke KM, Cassell C, Sethuraman G. Psychosocial outcomes of a pilot multidisciplinary program for weight management. American Journal of Health Behavior. 2001;27(4):348-54.
- 38. Kotler LA, Etu SF, Davies M, Devlin MJ, Attia E, Walsh BT. An open trial of an intensive summer day treatment program for severely overweight adolescents. Eating and weight disorders. 2006;11(4):e119-22.
- 39. Sahota P, Rudolf MCJ, Dixey R, Hill AJ, Barth JH, Cade J. Randomised controlled trial of primary school based intervention to reduce risk factors for obesity. British Medical Journal. 2001;323:1029.

- 40. Croker H, Viner R, Nicholls D, Haroun D, Chadwick P, Edwards C, et al. Family-based behavioural treatment of childhood obesity in a UK national health service setting: Randomized controlled trial. International Journal of Obesity. 2012;36(1):16-26.
- 41. Lochrie AS, Wysocki T, Hossain J, Milkes A, Antal H, Buckloh L, et al. The Effects of a Family-Based Intervention (FBI) for Overweight/ Obese Children on Health and Psychological Functioning. Clinical Practice in Pediatric Psychology. 2013;1(2):159-70.
- 42. Butler RJ, Gasson SL. Self Esteem/Self Concept Scales for Children and Adolescents: A Review. Child and Adolescent Mental Health. 2005;10(4):190-201.
- 43. Huang JS, Norman GJ, Zabinski MF, Calfas K, Patrick K. Body Image and Self-Esteem among Adolescents Undergoing an Intervention Targeting Dietary and Physical Activity Behaviors. Journal of Adolescent Health. 2007;40:245-51.
- 44. Pearson ES, Irwin JD, Burke SM. The Children's Health and Activity Modification Program (C.H.A.M.P.): Participants' perspectives of a four-week lifestyle intervention for children with obesity. Journal of Child Health Care. 2012;16(4):382-94.
- 45. Hafekost K, Boterhoven de Haan K, Lawrence D, Sawyer M, Zubrick S. Validation of the Adolescent Self-Esteem Questionnaire: Technical Report: Telethon Kids Institute and the Graduate School of Education. The University of Western Australia, Perth, Australia; 2017.
- 46. Franklin MR, Duley SM, Rousseau EW, Sabers DL. Construct validation of the Piers-Harris Children's Self Concept Scale. Educational and Psychological Measurement. 1981;41(2):439-43.
- 47. Stewart PK, Roberts MC, Kim KL. The psychometric properties of the Harter Self-Perception Profile for Children with at-risk African American females. Journal of Child & Family Studies. 2010;19(3):326-33.
- 48. Welk GJ, Eklund B. Validation of the children and youth physical self perceptions profile for young children. Psychology of Sport and Exercise. 2005;6(1):51-65.
- 49. Johnson BW, Redfield DL, Miller RL, Simpson RE. The Coopersmith self-esteem inventory: A construct validation study. Educational and Psychological Measurement. 1983;43:907-13.
- 50. Linton KE, Marriott RG. Self-esteem in Adolescents: Validation of the State Self-Esteem Scale. Personality and Individual Differences. 1996;21(1):85-90.
- 51. Rudasill K, Callahan C. Psychometric Characteristics of the Harter Self-Perception Profiles for Adolescents and Children for Use With Gifted Populations. Gifted Child Quarterly. 2008;52(70-86).
- 52. Harter S. The construction of the self: Developmental and socio-cultural foundations. New York: Guildford Press; 2012.
- 53. Marsh HW. Self-Description Questionnaire-I. San Antonio, TX: Psychological Corporation; 1988.
- 54. Marsh HW. Self-Description Questionnaire-III. San Antonio, TX: Psychological Corporation; 1991.
- 55. Rosenberg M. Society and the adoelscent self-image. Princeton, NJ: Princeton University Press; 1965.
- 56. Silber E, Tippet J. Clinical assessment and measurement validation. Psychological Reports. 1965;16:1017-71.
- 57. Marsh HW. Self Description Questionnaire (SDQ) I: A theoretical and empirical basis for the measurement of multiple dimensions of preadolescent self-concept. An interim test manual and research monograph. MacArthur, New South Wales: University of Western Sydney, Faculty of Education.; 1992.
- 58. Shavelson R, Hubner JJ, Stanton JC. Self-concept: Validation of Construct Interpretations. Review of Educational Research. 1976;46:407-41.
- 59. Fox K, B. Corbin C. The Physical Self-Perception Profile: Devlopment and Preliminary Validation Journal of Sport and Exercise Psychology. 1989;11:408-30.
- 60. Eklund RC, Whitehead JR, Welk GJ. Validity of the Children and Youth Physical Self-Perception Profile: A Confirmatory Factor Analysis. Research Quarterly for Exercise and Sport. 1997;68(3):249-56.
- 61. Steer RA KG, Beck JS, Beck AT. Evidence for the construct validities of the Beck Youth Inventories with child psychiatric outpatients. Psychological Reports. 2001;89(3):559-65.

- 62. Connors C. Conners' Rating Scales–Revised: Technical manual. Toronto, Ontario, Canada: Multi-Health Systems; 1997.
- 63. Kovacs M. Children's Depression Inventory (CDI) manual. New York: Multi Health Systems; 1992.
- 64. Kickett-Tucker C, Christensen D, Lawrence D, Zubrick S, Johnson D, Stanley F. Development and validation of the Australian Aboriginal racial identity and self-esteem survey for 8–12 year old children (IRISE\_C). International Journal for Equity in Health. 2015;14(103).
- 65. Purdie N MA. Measurement of self-concept among Indigenous and non-Indigenous Australian students. Australian Journal of Psychology. 2004;56(1):50-62.
- 66. Bodkin-Andrews GH, Ha MT, Craven RG, Yeung AS. Factorial Invariance Testing and Latent Mean Differences for the Self-Description Questionnaire II (Short Version) with Indigenous and Non-Indigenous Australian Secondary School Students. International Journal of Testing. 2010;10(1):47-79.
- 67. Australian Government Department of Health. Engaging with Aboriginal and Torres Strait Islander communities: Supplement to the KidsMatter Primary Action Team Handbook. 2014.
- 68. Fraser C, Lewis K, Manby M. Steps in the Right Direction, Against the Odds, An Evaluation of a Community-Based Programme Aiming to Reduce Inactivity and Improve Health and Morale in Overweight and Obese School-Age Children. Children & Society. 2012;26(2):124-37.
- 69. Harter S, Pike R. The pictorial scale of perceived competence and social acceptance for young children. Child Development. 1984(55):1969-82.
- 70. NHMRC. Values and Ethics: Guidelines for ethical conduct in Aboriginal and Torres Strait Islander health research. Australia; 2003.

# Appendices

### Appendix 1: Example database search, strategy 1

#### **Ovid Medline**

946 documents retrieved

1	(adolescen* or youth* or child* or p?ediatric or teenage* or boy* or girl* or student* or pupil*).mp. or school* Pediatrics/ or Adolescent/ or Child/ [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	3468349
2	limit 1 to (english language and yr="2007 -Current")	1128066
3	(overweight or obes*).mp. or overweight/ or obesity/ [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	287527
4	limit 3 to (english language and yr="2007 -Current")	164756
5	2 and 4	44436
6	((nutrition* or diet* or health* or behavio* or lifestyle or community or physical activity or exercise or fitness or weight or psycho* or multi-component or multicomponent or family or obes* or school*) adj3 (educat* or program* or intervention or care or therap* or treatment or promotion)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	1163803
7	limit 6 to (english language and yr="2007 -Current")	489048
8	weight reduction programs/ or health education/ or weight loss/ or intervention.mp. or prevention.mp. or "group work".mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	959466
9	limit 8 to (english language and yr="2007 -Current")	469719
10	7 or 9	840137

11	("self-esteem" or "self-concept" or "self-perception" or "self-worth" or "self-evaluation" or "self-confidence" or "self-love" or "self-compassion" or psychosocial).mp. or self- concept/ [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	141869
12	limit 11 to (english language and yr="2007 -Current")	63547
13	5 and 10 and 12	946

## Appendix 2: Example database search, strategy 2

### Ovid Medline

62 documents retrieved

1	(adolescen* or teen* or child*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	326769 1
2	validat*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	404980
3	("Harter* self-perception" or rosenberg or (marsh adj3 self- description) or "Children and Youth Physical Self-Perception Profile" or "beck youth inventory" or "Piers-Harris").mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	1711
4	1 and 2 and 3	62

#### **Appendix 3: NHMRC levels of evidence**

Level of Evidence	Study Design	Number of included studies
I	A systematic review of Level II studies.	0
II	A randomised controlled trial.	15
III-1	A pseudo-randomised controlled trial (i.e., alternate allocation or some other method).	0
III-2	A comparative study with concurrent controls (i.e., non- randomised experimental trials, cohort studies, case-control studies, interrupted time series studies with a control group).	2
III-3	A comparative study without concurrent controls (i.e., historical control study, two or more single arm studies, interrupted time series studies without a parallel control group).	1
IV	Case series with either post-test or pre-test/post-test outcomes.	17

The majority of multicomponent weight management interventions for overweight and obese children and adolescents, reporting measures of self-esteem, comprise level II (RCTs) and level IV (pre-test post-test trials) evidence.

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
Annesi et al 2007	Non- randomised, experimental trial (III-2)	US - Communi ty	SDQ-I general self 8-item scale, physical appearance 8 item scale	16	Child complete d	NS	Aged 9 to 12 years, M&F, After school group BMI M: 84th percentile F: 86th percentile, PE group BMI M:92nd percentile F: 91st percentile, AfricanAme rican, no exclusion criteria	African America n Lower to lower- middle SES strata	Youth Fit for Life Program: 12 weeks, 3 x 45- minute sessions a week comprising of cardiovascular and resistance exercise, nutrition and health information and behavioural skills training in school gymnasium, group sessions. Control group: unstructured physical activity sessions	Ν	General Self- Esteem improved Physical appearanc e improved Physical self- concept improved

# Appendix 4: Intervention studies measuring self-esteem as identified in search 1

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
Annesi et al 2008	Non- randomised, experimental trial (III-2)	US - Communi ty	SDQ-I: general self 8 item scale, physical appearance 8 item scale	16	Child complete d	100	Average age 10.6 years, M&F, Mean BMI M: 82nd percentile, F: 79th percentile, African American, no exclusion criteria given	African America n Lower to lower- middle SES strata	Youth Fit for Life Program: 12 weeks, 3 x 45- minute sessions a week comprising of cardiovascular and resistance exercise, nutrition and health information and behavioural skills training in an after-school care setting, group sessions. Control group: unstructured physical activity sessions	Ν	General Self- Esteem improved Physical appearanc e improved Physical self- concept improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
Archuleta et al 2016	Pre-test Post- test (IV)	US - Communi ty	SPPC 6 subscales, each with a 6-item scale	36	Child complete d	NS	Aged 8 to 17 years, M&F, BMI ≥85th percentile, mainly Hispanic, needed physician referral	Majority Hispanic populati on, SES not stated	The Fit Families Program: 7 weeks, 2-hour session a week in school and YCMA facilities. The sessions aim to maintain or decrease BMI by increasing physical activity, decreasing sedentary activity and improve self-esteem and knowledge about healthy lifestyle behaviours. No control group	Ν	Global self-worth no change Physical appearanc e improved Athletic competen ce improved No other changes

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
Bonham et al 2017	RCT (II)	Australia - Commerci al	Rosenberg 5- item scale	5	Adolescen t complete d	91	Aged 13 to 17 years, M&F, BMI z score ≥1.282, otherwise healthy	Majority high or highest SES	The JenMe Program: 12 weeks, 12 weekly sessions with commercial trained consultant comprising of dietary and behavioural education with progress reviews, slowly transitioned from pre-prepared meals to own meal planning. One-on-one sessions. Control group: waitlist control received standard healthy eating guidelines	Ν	Global self- esteem improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
Brennan et al 2012	RCT (II)	Australia - Communi ty	Rosenberg 10 item scale	10	Adolescen t complete d	NS	Aged 11 to 19 years, M&F, BMI internationa Ily defined as overweight or obese, capable of fully participatin g in program and at least one parent willing to actively participate	Not stated	Choose Health Program: 16 weeks, 12 60- minute sessions (10 weekly, 2 fortnightly) and one phone call. Program assisted adolescents in making behavioural changes aimed at improving dietary choices and increasing physical activity. Sessions were one-on-one in a clinic setting, followed by a maintenance phase. Control	Ν	Global self- esteem no change

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									group: wait list control		
Christins on et al 2012	Pre-test Post- test (IV)	USA - Communi ty	SPPC- 6 subscales	NS	Child complete d	55	Aged 8 to 16 years, M&F, BMI≥85th percentile, mixed ethnicities, otherwise healthy	Majority white SES not stated	Exergaming for Health Program; 10 weeks, 10 x 2- hour session weekly comprising of exergaming activities, nutrition education and behavioural management in groups. No control group	N	Global self-worth improved Behaviour al conduct improved no other changes

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
Christins on et al 2016	Pre-test Post- test (IV)	US - Communi ty	CY-PSPP 36- item scale with 6 domains including global self- worth	36	Child complete d	45.8	Aged 8 to 16 years, M&F, BMI≥85th percentile, mixed ethnicities, otherwise healthy	Located near neighbo urhoods with poverty	Exergaming for Health Program; 10 weeks, 10 x 2- hour session weekly comprising of exergaming activities, nutrition education and behavioural management in groups. No control group	Includes semi- online compone nt (exergami ng) but this was done in group setting in person	Global self-worth improved Physical self-worth improved
Croker et al 2012	RCT (II)	UK - Primary Care	SPPC (no. of items NS)	NS	Parent and child complete d together	56	Aged 8 to 12 years, M&F, mixed ethnicities, BMI internationa Ily defined as	NS	Family-based behavioural treatment: 6 months, 12 x 1.5- hour session over a 6-month period, treatment	Ν	Global self- esteem improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
							overweight or obese, otherwise healthy		focused on whole family lifestyle changes in group sessions. Control group: Wait-list control		
Danielse n et al 2013	RCT (II)	Norway - Research	SPPC 36 item scale with 6 subscales	36	Adolescen t complete d	NS	Aged 10 to 13 years, gender NS, BMI internationa Ily defined as obese, ethnicity NS, otherwise healthy,	NS	12 weeks, 12 x 45-minute weekly sessions to establish a healthy diet, increase physical activity and reduce sedentary activity by developing family and individual coping skills through modest long- term lifestyle modifications. One-on-one	N	Global self-worth improved Physical appearanc e improved Athletic competen ce improved Behaviour al conduct improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									sessions with clinical psychologist. Control group: wait-list control		
DeBar et al 2012	RCT (II)	US - Primary care	Rosenberg (no. of items NS)	NS	Adolescen t complete d	93.1	Aged 12 to 17 years, F only, BMI ≥90th percentile, mixed ethnicities, health plan members, not severely obese (BMI>45)	NS	5 months, 16 x 90 minute over 5 months encouraging changes in dietary intake and eating patterns, increasing physical activity and discussion of common issues for obese girls such as disordered and altered body image. Parent	N	Global self- esteem improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									support meeting were included. Sessions held in groups. Control group: Usual care		
Fagg et al 2014	Pre-test Post- test (IV)	UK - Communi ty	Rosenberg 10 item scale (modified wording for use by younger children)	10	Child complete d	37.7	Ages 7 to 13 years, M&F, BMI≥ 91st percentile, mixed ethnicities, no exclusion criteria	NS	The MEND program: 10 weeks, 20 sessions delivered over 10 weeks in schools and community health centres. Family-based intervention using a range of professionals to promote the adoption and to sustain a healthy lifestyle. No control group	Ν	Global self- esteem improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
Fonseca et al 2014	Pre-test Post- test (IV)	Portugal - Communi ty	Rosenberg 10 item scale	10	Adolescen t complete d	NS	Aged 12 years or older, M&F, BMI≥ 95th percentile, ethnicity NS, no behavioural problems	NS	2-week residential weight management camp. Including 2 x 45-minute walking sessions a day, 120 minutes of games-based vigorous activity. Meals were prepared by a nutritionist with participants encouraged to participate in meal prep and daily psycho- educational support to facilitate	Ν	Global self- esteem no change

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									behaviour modification. No control group		
Fraser et al 2012	Pre-test Post- test (IV)	UK - Communi ty	Rosenberg 10 item scale: written scaled responses replaced with typographica I emoticons (grinning smiley face, smiley face, sad face, sad face with tear)	10	Child complete d, likely with parent/car er	NS	Aged 5 to 16 years, M&F, BMI≥91st percentile, mixed ethnicities, no exclusion criteria	NS	The Newtown Kids Program: 48 weeks, 12-weekly review sessions involving nutritional advice. Attend 'Fusion' sessions or physical activity component of their choice. Fusion sessions team-based physical activity sessions with group healthy	N	Global self- esteem improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									eating education. No control group		
Incledon et al 2013	RCT (II)	Australia - Primary care	Modified (at baseline only) SPPC (no. of items NS)	NS	Complete d by both parent and child	83.1	Aged 5 to 12 years, M&F, BMI internationa Ily defined as overweight or obese, ethnicity NS, not currently receiving treatment for obesity,	NS	The LEAP2 Program: 12 weeks, 4 sessions with GP over 12 weeks aimed at reducing BMI through sustainable nutritional and physical activity behaviour choices using solution-focused therapy. Control group	N	Global self- esteem reduction Physical self- esteem improvem ent

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
							BMI z- score≥3.0				
Jacobson et al 2013	Pre-test Post- test (IV)	US - Primary care	BYI-II 20-item scale	20	Child complete d	100	Mean age 10.7 years, M&F, mean BMI z-score 1.91, mixed ethnicities, no exclusion criteria stated	NS	Healthy Choices Program: 7-13 weeks, 7 weekly sessions designed to use goal setting for improve nutrition and physical activity. No control group	N	Global self- concept no change

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
Jelalian et al 2011	RCT (II)	NS - Research	SPPA 25-item scale, 5 of a possible 8 domains	25	Adolescen t complete d	95.7	Aged 13 to 16 years old, M&F, 30-90% overweight defined by BMI, mixed ethnicities, at least one parent able to participate	NS	4 months, 16 x 1- hour weekly sessions, with a prescribed balanced deficit diet and a gradual increase in physical activity through 2 conditions. First Condition: Cognitive behavioural treatment (CBT) with aerobic exercise. Second Condition: CBT with peer- enhanced adventure therapy, which included group	Ν	Global self-worth improved Physical appearanc e improved Romantic appeal improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									based physical activities designed to improve peer relationships		
Joosse et al 2008	Pre-test Post- test (IV)	US - Communi ty	Rosenberg 10 item scale	10	Child complete d	NS	Aged 5 to 16 years, M&F, BMI ≥85th percentile, ethnicity NS, no exclusion criteria	NS	The Fit Kids/Fit Families Program: 12 weeks, 1 x 2- hour session a week in school and YCMA facilities. The sessions aim to maintain or decrease BMI by increasing physical activity, decreasing sedentary activity	Ν	Global self- esteem improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									and improve self-esteem and knowledge about healthy lifestyle behaviours. No control group		
Kolotour ou et al 2015	Pre-test Post- test (IV)	UK - Communi ty	Rosenberg self-esteem scale (no. of items NS)	NS	Child complete d	84.2	Aged 7 to 13 years, M&F, BMI ≥91st percentile, mixed ethnicities, no exclusion criteria	NS	The MEND program: 10 weeks, 20 sessions delivered over 10 weeks in schools and community health centres. Family-based intervention using a range of professionals to promote the adoption and to sustain a healthy	Ν	Global self- esteem improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									lifestyle. No control group		
Lloyd- Richards on et al 2012	RCT (II)	US - Research	SPPA 45-item scale, 5 of 8 potential domains included	45	Adolescen t complete d	NS	Aged 13 to 16 years old, M&F, 30-90% overweight defined by BMI, mixed ethnicities, at least one parent able to participate	NS	16 weeks, 16x 1- hour weekly sessions, with a prescribed balanced deficit diet and a gradual increase in physical activity through 2 conditions. First Condition: Cognitive behavioural treatment (CBT) with aerobic exercise. Second Condition: CBT with peer- enhanced adventure	Ν	Global self-worth improved Physical appearanc e improved Social acceptanc e improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									therapy, which included group based physical activities designed to improve peer relationships		
Lochrie et al 2013	RCT (II)	US - Communi ty	SPPC (No. of items NS)	NS	Child complete d	NS	Aged 8 to 11 years old, M&F, BMI ≥85th percentile, mixed ethnicities, otherwise healthy	NS	6 months, 14 x 60-90-minute sessions over 6 months. Group sessions led by psychologist and dietitian to cover nutrition, behaviour modification, psychosocial interventions, exercise topics and obesity related medical	N	Global self- esteem no change

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									issues. Control group: 1 x 60- minute session on general nutrition and physical activity recommendation s		
McCallu m et al 2007	RCT (II)	Australia - Primary care	SPPC: Modified scale	NS	Child complete d	100	Aged 5 and 9 years, M&F, BMI classified as overweight or obese, ethnicity NS, not with a BMI z- score ≥3	NS	The LEAP2 Program: 12 weeks, 4 sessions with GP over 12 weeks aimed at reducing BMI through sustainable nutritional and physical activity behaviour choices using solution-focused therapy. Control group	Ν	Global self-worth no change Physical appearanc e no change

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
Mehlenb eck et al 2009	RCT (II)	US - Research	SPPA (No. of items NS)	NS	Adolescen t complete d	NS	Both studies: Aged between 13 and 16, M&F, ethnicity NS, not already enrolled in weight loss program Study 1: 20%-80% overweight defined by BMI Study 2: 30%-90% overweight defined by BMI	NS	16 weeks, 16 x 1- hour weekly sessions, with a prescribed balanced deficit diet and a gradual increase in physical activity through 2 conditions. First Condition: Cognitive behavioural treatment (CBT) with aerobic exercise. Second Condition: CBT with peer- enhanced adventure therapy, which included group	Ν	Social Acceptanc e NS Athletic Competen ce NS and Physical Appearanc e NS

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									based physical activities designed to improve peer relationships		
Murdoch et al 2011	Pre-test Post- test (IV)	UK - Communi ty	SPPC (No. of items NS)	NS	Child complete d with assistance from facilitators if required	NS	Aged 7 to 14 years old, M&F, BMI ≥98th percentile, mixed ethnicities, exclusion criteria NS	NS	The Family Based Behavioural Management Program: 6 months 15 90- minute group sessions over 6 months. Aims to give advice to the whole family to encourage a healthy lifestyle. Using behavioural modification techniques to modify children's	N	Physical self- concept improved No change in other subscales

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									eating and exercise behaviours. No control group		
Nobles et al 2016	Pre-test Post- test (IV)	UK - Communi ty	Rosenberg Scale (no. of items NS)	NS	Adolescen t complete d	NS	Aged 10 to 17 years old, M&F, BMI≥99.9th percentile, exclusion criteria NS	27.1% of participa nts had a diagnose d learning disability	The SHINE Program: ongoing care, 3 distinct phases of treatment. Phase 1: assess participants and assign appropriate treatment method. Phase 2: active intense educational program designed to stem	Ν	Global self- esteem improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									weight gain. Phase 3: maintenance phase which educates participants to maintain established behaviour changes. No control group		
O'Conno r et al 2008	Pre-test Post- test (IV)	Australia - Communi ty	SSPA	NS	Adolescen t complete d	100	Aged 13 to 16 years, M&F, BMI z-score range 1.0- 3.5, ethnicity NS, otherwise healthy	NS	The Loozit Trial: 5 months, 7 x 75- minute weekly group sessions. Using cognitive theory to change dietary intake, activity levels through encouraging participants to set weekly goals. Sessions focus	Ν	Global self-worth no change Physical appearanc e improved Romantic appeal improved

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									on benefits of		
									healthy living,		
									increasing		
									physical activity,		
									nealthy eating		
									and building		
									esteem Followed		
									by additional 22		
									months follow		
									up contact		
									consisting of		
									booster sessions		
									and telephone		
									and SMS		
									coaching		
									sessions. Control		
									group: No		
									telephone or		
									SMS coaching		
									sessions		

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
Panagiot opoulos et al 2011	Interrupted time-series controlled study (III-3)	Canada - Communi ty	BYI-II (No. of items NS)	NS	Children/t eens	NS	Aged 6 to17 years old, M&F, BMI ≥85th percentile, mixed ethnicities, at least one co- morbidity	NS	The Centre for Healthy Weights— Shapedown BC Program: 10 weeks, 10 x 2- hour weekly sessions. Sessions in part led by psychologist and dietitian and included 30 minutes of physical activity. Dietitian's provided information on following the Canadian Food Guide, psychologists provided assistance with	Ν	Global self- concept improvem ent
Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
------------------------------	--	-----------------------	--------------------------------------	--	-------------------------------------	--	--	--------------------------------	--	--	--
									goal setting and other behaviour change modifications. No control group		
Pathmasi ri et al 2012	Pre-test Post- test (IV)	US - Communi ty	Rosenberg 10 item scale.	10	Adolescen t complete d	63.2	Aged 12 to 18 years, M&F, BMI ≥95th percentile, White and African American, exclusion criteria NS	NS	The Take off 4 Health Program: 3-week summer camp including personalised nutrition prescription, 6 sessions on nutritional education, 2-4 hours daily physical activity and group and one-on-one cognitive	N	Global self- esteem improvem ent

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									behavioural therapy sessions. No control group		
Quinlan et al 2009	Pre-test Post- test (IV)	US - Communi ty	Rosenberg 10 item scale	10	Child complete d	100	Aged 9 to 18 years, M&F, mean BMI z-score 2.2, mixed ethnicities, exclusion criteria NS	No	Residential camp lasting 1 to 8 weeks. Nutritionally balanced meal provided at 1800 kcal/day with nutrition classes twice a week and cooking classes once a week. Daily physical activity in 5 one- hour sessions a day and group psychological sessions providing psychoeducation	N	Global self- esteem improvem ent

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									and therapeutic support in the sessions which dealt with issues such as body image, emotional eating and self- esteem. No control group		
Sacher et al 2010	RCT (II)	UK - Communi ty	SPPC (No. of items NS)	NS	Child complete d	84	Aged 8 to 12 years, M&F, BMI≥ 98th percentile, mixed ethnicities, otherwise healthy	NS	The MEND program: 10 weeks, 20 sessions delivered over 10 weeks in schools and community health centres. Family-based intervention using a range of professionals to promote the	N	Global self- esteem improvem ent

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									adoption and to sustain a healthy lifestyle. No control group		
Shrewsb ury et al 2011, Nguyen et al 2012, Nguyen et al 2013	RCT (II)	Australia - Communi ty	SPPA 45-item scale	45	Adolescen t complete d	100	Aged 13 to 16 years, M&F, BMI z-score range 1.0- 3.5, ethnicity NS, otherwise healthy	NS	The Loozit Trial: 7 weeks, 7x 75- minute weekly group sessions. Using cognitive theory to change dietary intake, activity levels through encouraging participants to set weekly goals. Sessions focus on benefits of healthy living, increasing physical activity, healthy eating and building positive self-	N - SMS and phone call coaching	Global self-worth improved Scholastic competen ce improved Social competen ce improved Athletic competen ce improved Physical appearanc e improved Job

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
									esteem. Followed by additional 22 months follow up contact consisting of booster sessions and telephone and SMS coaching sessions. Control group: No telephone or SMS coaching sessions		competen ce improved Romantic appeal improved Behaviour al conduct improved
Tirlea et al 2016	Cluster RCT (II)	Australia - Communi ty	Rosenberg 10-item Scale	10	Child complete d, assistance was available if required	NS	Aged 10 to 16 years, F only, BMI NS, mixed ethnicities, identified as having poor body	Culturall y diverse area of high social disadvan tage	10 weeks, 10 sessions the first and last session lasting 1 hour and the middle 8 sessions lasting 3 hours. Sessions tailored to	N	Global self- esteem improvem ent

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
							image/self- esteem, underweigh t or overweight		provide team building activities to promote goal setting, sessions on body image and self-esteem, healthy eating, assertiveness, physical activity, trust, celebrations and connections. Control group: wait-list control		
Watson- Jarvis et al 2011	Pre-test Post- test (IV)	Canada - Communi ty	Piers-Harris 2 (No. of items NS)	NS	Child complete d	19	Aged 6 to 12 years, M&F, BMI ≥85th percentile, ethnicity NS, at least one parent	NS	8-12 weeks, 12 x 2-hour weekly sessions consisting of group physical activities and talks on healthy lifestyle and	Ν	Global self- esteem improvem ent (BMI≥98th percentile)

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
							able to participate		nutrition. No control group		
Watson et al 2015	Pre-test Post- test (IV)	UK - Communi ty	SPPC 36-item scale	36	Child complete d	53	Aged 4 to 16 years old, M&F, BMI z-score range 1.53- 4.73, mixed ethnicities, exclusion criteria NS	NS	The Getting Our Active Lifestyles Started Program: 18 weeks, 18 x 2- hour weekly group sessions using social cognitive theory to promote whole family lifestyle changes. Sessions covered physical activity, diet and behaviour change. No control group	N	Global self- esteem no change Social acceptanc e no change Athletic competen ce no change Physical appearanc e no change

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
Weintrau b et al 2008	RCT (II)	US - Communi ty	Rosenberg 10 item scale	10	Child complete d	100	Aged 8 to 11 years old, gender NS, M&F, BMI ≥85th percentile, mixed ethnicities, from low income or ethnically diverse populations , otherwise healthy	Low income commun ity	6 months, 3-4 sessions a week of soccer training built around improving respect for self and others, inclusion and teamwork. Sessions included a homework component followed by 75 minutes of soccer training. Control group: 25 sessions of nutritional education	Ν	Global self- esteem no change

Author, year	Study design (Level of evidence)	Location / setting	Self-esteem measureme- nt tool	Num- ber of items used in tool	Parent/ child complet- ed?	% who comple- ted self- esteem measure	Population characteri- stics	Low SES/ vulnera- ble	Intervention summary	Online only intervent ion (Y/N)	Scale outcomes
Wong et al 2009	Pre-test Post- test (IV)	US- Communi ty	SPPC 36-item scale	36	Child complete d	NS	Aged 10 to 14 years old, M&F, BMI≥ 95th percentile, mixed ethnicities, no conditions that prevented camp participatio n	NS	A 2-week summer camp receiving 6 x behavioural session and 4 x nutritional and physical activity sessions all aimed around improving self- esteem in children and teaching the importance of physical activity and good nutrition. No control group.	N	Global self-worth improved Physical appearanc e improved Athletic competen ce improved Social competen ce improved

Author,	Title	Conclusion
date		
Butler et al, 2005	Self Esteem/Self Concept Scales for Children and Adolescents: A Review	Overall there has been a shift towards acceptance of multidimensionality with respect to the self, with the latest scales designed around such a construct.
Lowry et al, 2007	The Effects of Weight Management Programs on Self-Esteem in Paediatric Overweight Populations	Of the 21 studies reviewed, 18 studies reported evidence of increases in self-esteem or components of self-esteem from pre- to post-treatment. The authors conclude that overall small positive effects of weight loss on self-esteem in paediatric populations were observed. Components associated with improvements in self-esteem included weight change, parental involvement and a group intervention format.
Griffin et al, 2010	Self-esteem and quality of life in obese children and adolescents: A systematic review	Six of nine cross-sectional studies found lower global self-esteem in obese rather than healthy weight children and adolescents. Generally, weight loss did result in increases in global self-esteem in the majority of studies included that reported on self-esteem.
Ajle et al, 2014	Impact of computer-mediated, obesity-related nutrition education interventions for adolescent: A systematic review	The aim of this review was to examine the effectiveness of nutrition education programs delivered in an online setting to target overweight and obesity. Of the 15 included studies 10 were RCTs. Although a number of the studies purported to measure self-efficacy, tools used did not specifically address self-esteem.
Murray et al, 2017	Systematic Review and Meta-Analysis: The Impact of Multicomponent Weight Management Interventions on Self-Esteem in Overweight and Obese Adolescents	Although changes in self-esteem were observed after behavioural weight-loss interventions, meta- analyses showed no significant change in self-esteem in the 13 included studies (0.27 [-0.04, 0.59]) even though significant changes in BMI z score were reported (-0.17 [-0.22, -0.11]). The authors conclude that multicomponent weight management interventions require a specific focus on self- esteem to improve this outcome in overweight and obese adolescents
Hoare et al, 2017	Systematic review of mental health and wellbeing outcomes following community-based obesity preventions among adolescents.	The review reported on wellbeing outcomes after participation in community-based obesity prevention interventions. Of the two studies that reported on a self-esteem/self-efficacy neither showed an improvement in the outcome of interest.

## Appendix 5: Summary of systematic reviews that examine self-esteem as a component of weight loss interventions

Paper	Tool	Country	Administration	Sample	Population	Setting	Design	Statistics	Reference
Bagley & Mallick 2001	Rosenberg	UK	Self-report, NS whether instruction was given	1310	12-19 years, M&F, middle working class	Secondary Schools and Sixth Form Colleges	Administered once in a classroom setting	Internal consistency: M:0.85 F:0.85	Christopher Bagley & Kanka Mallick (2001) Normative Data and Mental Health Construct Validity for the Rosenberg Self-Esteem Scale in British Adolescents, International Journal of Adolescence and Youth, 9:2-3, 117-126, DOI: 10.1080/02673843.2001.9747871
Marsh & Holmes 1990	SDQI, Piers- Harris, SPC	Australia	Self-report and administered by one of the authors (Homes) in a group setting. The instructions and each question was read aloud to the children and several practise questions were	290	10 years, M&F, upper- middle - working class	Schools	Compared SDQ-I, Piers- Harris and SPC. Each administered once in the classroom with teachers passively present	Factor analysis: target coefficients consistently large median 0.75. Convergent validity: mean = 0.61 statistically significant	Marsh HW, MacDonald Holmes IW. Multidimensional self- concepts: Construct validation of responses by children. American Educational Research Journal. 1990 Mar;27(1):89-117

## Appendix 6: Validation studies in populations of interest of identified tools from search 1

Paper	ΤοοΙ	Country	Administration	Sample	Population	Setting	Design	Statistics	Reference
			also given to the children						
Marsh 1990	Same as abo	ove - this is	a reanalysis					Convergent validity: mean = 0.61 p<0.05	Marsh HW. Confirmatory Factor Analysis of Multitrait- Multimethod Data: The Construct Validation of Multidimensional Self-Concept Responses. Journal of Personality. 1990 Dec 1;58(4):661-92

Paper	ΤοοΙ	Country	Administration	Sample	Population	Setting	Design	Statistics	Reference
Franklin et al 1981	Piers- Harris	US	Self-report, administered in a group setting, NS whether instruction was given	569	9 & 12 years, M&F, mainly Hispanic, some Caucasian and 'other'	Schools	Compared Piers-Harris to Coopersmith. Both were administered to all Grades 4 and 7 pupils in 5 schools. Content of Piers-Harris split evenly into two 40- item forms to test sensitivity to change. 180 pupils were split into control and experimental groups with experimental group receiving tutoring for 4- 6months designed to	Convergent validity: R=0.78 between the two scales for all pupils r=0.75 for 4th grade and r=0.81 for 7th grade Discriminant validity: r= 0.18 for 4th grade r=0.22 for 7th grade compared to academic achievement, has discriminant validity. Sensitivity to change: F=0.56 for interaction and F=0.93 for main effect in control with lack of significance showing stability. Internal consistency =0.92 (p<0.001) for total scale and 0.74	Franklin Jr MR, Duley SM, Rousseau EW, Sabers DL. Construct Validation of the Piers-Harris Children's Self Concept Scale. Educational and Psychological Measurement. 1981 Jul;41(2):439-43

Paper	ΤοοΙ	Country	Administration	Sample	Population	Setting	Design	Statistics	Reference
							improve self- concept	(p<0.001) for 1st 40-item scale and 0.77 (p<0.001) for 2nd	
Stewart et al 2010	SPPC	USA	Self-report administered in a group setting on the first day of a summer camp for at risk teens, NS whether instructions were given	92	12 years, M&F, lower SES, African American	Summer camp	Compared to Rosenberg to establish the stability of the construct assessment in an African- American population. Questionnaires were completed once	Internal reliability: alpha = 0.79 Convergent validity: r=0.37 compared to Rosenberg	Stewart PK, Roberts MC, Kim KL. The psychometric properties of the Harter self-perception profile for children with at-risk African American females. Journal of Child and Family Studies. 2010 Jun 1;19(3):326-33. for Children with At-Risk African American Females

Paper	ΤοοΙ	Country	Administration	Sample	Population	Setting	Design	Statistics	Reference
Welk & Eklund 2005	CY-PSPP	US^	Self-report administered in a group setting at school during PE class. Survey was read to children and assistants circulated throughout the room to provide extra assistance	754	8-12 years, M&F, middle class, majority Caucasian	Schools	Administered once in a school class setting along with assessment of physical activity and physical fitness tests	Factorial validity: all models differed significantly from their reference independence models. The comparative fit index and non- normed fit index exceeded the 0.90 criterion in all instances. Median loadings for the total sample, boys subsample and girls subsample were 0.69 (range 0.41–0.82), 0.67 (range 0.32–0.85), and 0.70 (range 0.43–0.83)	Welk GJ, Eklund B. Validation of the children and youth physical self-perceptions profile for young children. Psychology of Sport and Exercise. 2005 Jan 31;6(1):51-65

M = male, F = female, NS = not stated, SES = Socioeconomic Status, Rosenberg = Rosenberg Self-Esteem Scale, SDQ-I = Marsh's Self-Description Questionnaire I, Piers-Harris = Piers-Harris Children's Self Concept Scale, SPC = Harter's Perceived Competence SPPC = Harter's Self-Perception Profile for Children, Coopersmith = The Coopersmith Self-Esteem Inventory, CY-PSPP = Children and Youth Physical Self-Perception Profile ^ not explicitly stated, inferred from article.

Tool	RSES	SPPC	Piers-Harris-2	SDQ	BYI-II	CY-PSPP	IRISE_C
No of items	10	36	60	76	Up to 100	36	40
Multi/One-	One	Multidimensiona	Multidimensiona	Multidimensiona	Multidimensional	Multidimensiona	Multidimensional
dimensional	dimensiona	1	1	1		1	
	1						
Subscales	None	Scholastic Competence, Athletic Competence, Social Competence, Physical Appearance, and Behavioural Conduct. A separate, sixth subscale, taps Global Self- Worth (or self- esteem)	Behavioural adjustment, Intellectual and school status, Physical appearance and attributes, Freedom from anxiety, Popularity and Happiness and Satisfaction	Physical Ability, Physical Appearance, Peer Relations, Total Academic, General Self, Reading, Mathematics, All school subjects and Parent Relations	Depression, Anxiety, Anger, Disruptive behaviour, and Self- concept	Global self- esteem, Physical self-worth, Sport competence, Body attractiveness, Physical strength and Physical condition	Knowledge/experienc e of identity, Salience placed upon identity, Knowledge/experienc e of culture, Salience placed upon culture
Completed by	Children and adolescents	Children	Children and Adolescents	Children	Children/interviewe r	Children with assistance from trained personnel	Children with RA
Year of validation	1965	1985	2002	1988	2001	1989	2015

### Appendix 7: Overview of self-esteem measurement tools

Tool	RSES	SPPC	Piers-Harris-2	SDQ	BYI-II	CY-PSPP	IRISE_C
Validation population (Country, age, males and females, ethnicity, SES)	US, males and females, 8- 18, ethnically diverse, range of SES	US	US, 6-16 years, males and females, ethnically diverse, range of SES	Australia, mean age 10.5 years, males and females, range of SES from working class to upper middle class	US, 7-12 years, males and females, majority Caucasian, middle to middle- upper class, diagnosed with a mental health disorder	US, 19.7 years, males and females, college students, ethnicity and SES NS	Australia, 6-13 years old males and females, Aboriginal
Tested in Australia (Y/N)	Y	Y	N	Y	N	N	Y
Tested in Population of interest (Y/N)	Y	Y	Y	Y	Y	Y	Y
Tested in Vulnerable Population (Y/N)	Y	Y	Y	N	Y	Y	Y
Face Validity	acceptable	acceptable	acceptable	acceptable	acceptable	acceptable	NA
Convergent Validity	acceptable	acceptable	acceptable	acceptable	acceptable	acceptable	NA
Discriminant Validity	acceptable	acceptable	acceptable	acceptable	acceptable	acceptable	NA
Content Validity	acceptable	acceptable	acceptable	acceptable	acceptable	acceptable	NA
Factor Analysis	acceptable	acceptable	acceptable	acceptable	acceptable	acceptable	NA
Internal Consistency (Reliability)	acceptable	acceptable	acceptable	acceptable	acceptable	acceptable	NA
Test-Retest Reliability	> 0.70	> 0.70	> 0.70	> 0.70	> 0.70	> 0.70	Ν

Tool	RSES	SPPC	Piers-Harris-2	SDQ	BYI-II	CY-PSPP	IRISE_C
Short vs Long Form	Y	N	N	Y	Ν		Ν
Respondent burden	low	low/moderate	moderate	moderate/high	moderate/high	low/moderate	low
Alternate Versions available (i.e. translation/adaptation )	Y	Y	Y	Y	Y	Y	N
Freely Available	Y	Y	N	Y	Ν	ТВС	ТВС
Reference	Rosenberg & Morris, 1989	Harter, 1985, 2012 (SPPC); Harter, 1988, 2012 (SPPA)	Piers et al, 2002	Marsh, 1988	Steer et al, 2001	Welk & Eklund, 2005; Fox & Corbin, 1989	Kickett-Tucker et al, 2015
Additional Notes	NA	Recommended that a parent/carer reads the questions to the younger age children	Answers and scores should be interpreted by an appropriately trained psychological professional	NA	Completion by a professional Level B (i.e. registered allied health)	NA	NA

RSES - Rosenberg Self-Esteem Scale; SPPC - Harter's Self Perception Profile for Children; Piers-Harris-2- Piers-Harris Self Concept Scale for Children 2; SDQ - Marsh's Self-Description Questionnaire; BYI-II - Beck Youth Self-Concept Inventory; CY-PSPP – Children and Youth Physical Self-Perception Profile; IRISE\_C - Racial Identity and Self-Esteem of children; RA – research assistant; Y – yes; N – no; Appendix 8: Multicomponent weight management interventions that assessed self-esteem in overweight and obese adolescents from vulnerable population groups

Reference	Study design	Location	Setting	Self- esteem measurem ent tool	Parent/chi ld complete d	n	% who complete d self- esteem	Age	Vulnerable population	Interventio n duration	Self-esteem outcome	Acceptabili ty and Usability
Annesi et al 2007 Annesi et al 2008	Non-randomised experimental trial Non-randomised experimental trial	US	Communit y-based Communit y-based	SDQ-I SDQ-I	Child completed Child completed	231 269	Unclear 100	9 to 12 mean 10.6	African American Lower to lower-middle SES strata African American Lower to lower-middle SES strata	12 weeks	General-self improved Physical appearance improved Physical self-concept improved General self- improved Physical appearance improved	NS
Christinso n et al 2016	Pre-test post-test	US	Communit y-based	CY-PSPP 6 domains	Child completed	84	45.8	8 to 16	Located near neighbourho ods with poverty	10 weeks	Physical self-concept improved Global self- worth improved Physical	NS

Reference	Study design	Location	Setting	Self- esteem	Parent/chi Id	n	% who complete	Age	Vulnerable population	Interventio n duration	Self-esteem outcome	Acceptabili ty and
				measurem ent tool	complete d		d self- esteem					Usability
											self-worth improved	
Nobles et al 2016	Pre-test post-test	UK	Communit y-based	RSES	Adolescent completed	347	unclear	10 to 17	27.1% or participants had a diagnosed learning disability	Ongoing care	Global self- esteem improved	NS
Tirlea et al 2016	RCT	Australia	Communit y-based	RSES 10 item	Child completed, assistance if required	62, and 60	98	10 to 16	Culturally diverse area of high social disadvantage	10 weeks	Global self- esteem improved	NS
Weintraub et al 2008	RCT	US	Communit y-based	RSES 10 item	Child completed	21	100	8 to 11	low income community	6 months	No change in global self-esteem	NS

SDQ-I - Marsh's Self-Description Questionnaire-I; CY-PSPP - Children and Youth Physical Self-Perception Profile; RSES - Rosenberg Self-Esteem Scale; SES - socioeconomic status; NS - not stated

# Appendix 9: Summary of NHMRC Guidelines for working with Aboriginal and Torres Strait Islander populations

In considering validation of a self-esteem measure in Aboriginal and Torres Strait Islander populations, we recommend that the NHMRC guidelines for ethical conduct in Aboriginal and Torres Strait Islander Health Research be considered in establishing a process for validation that respectfully engages the Aboriginal and Torres Strait Islander population in the research and provides capacities, outcomes or opportunities that are of benefit and of value to them. This involves consideration of:

- Consultation with the community so that participating communities have had equal and respectful input into the development of the project and how it will be implemented and have understood and expressed satisfaction with the project, its potential benefits and how study findings will be disseminated
- Values and cultures being respectfully considered in the development of the self-esteem tool in addition to consideration of differences in values, norms and aspirations
- Attention being given to involving the community in decisions regarding the project and that local structures are acknowledged and used in these processes
- Recognition of different people's input into the project
- Information provided as a result of the project being understood and usable in decision-making by the participating communities.

#### **Reference:**

NHMRC. Values and Ethics: Guidelines for ethical conduct in Aboriginal and Torres Strait Islander health research. Commonwealth of Australia, 2003. <u>http://www.nhmrc.gov.au</u>

#### Appendix 10: Recommendations for validating a self-esteem tool

#### What are psychometric properties?

In psychology research, using self-report questionnaires are central to collecting data about the constructs we are interested in. These self-report measurement instruments rest on underlying assumptions of validity and reliability. Validity concerns whether a measurement instrument is accurately and comprehensively measuring the construct of interest. For example, a self-esteem measurement instrument would be valid if an individual who was scored as low on self-esteem did in fact have low self-esteem. Reliability concerns whether a measurement instrument is a consistent manner across items and time. For example, a self-esteem measurement reliable if an individual with low self-esteem measurement instrument would be deemed reliable if an individual with low self-esteem was scored as having low self-esteem across all the items within the instrument, and at different time points (assuming that their self-esteem remained the same during the interim period).

#### Why is investigating psychometric properties important?

In order for the data we collect to be accurate, assumptions of validity and reliability must be met. Therefore, it is important that we conduct research using measurement instruments that have established psychometric adequacy (i.e., evidence of validity and reliability). In some case, previous studies may have established psychometric adequacy of a measurement instrument in a population that is different from the population we intend to use in our research (e.g., adolescences vs adults). In this case, psychometric adequacy cannot be assumed to extend to the population we intend to use our measurement instrument in. This is because the way a construct operates may be different across populations.

#### How can we establish psychometric adequacy?

Testing the psychometric properties of an instrument is a relatively straightforward process. Most often this can be achieved through two administrations of an online survey containing the measurement instrument of interest and several other related measurement instruments that have established validity (to compare the primary measurement instrument to). Additionally, a relatively small sample size can be used (a common rule-of-thumb is seven participants per item, with at least 100 participants total). Researchers can use demographic information, scores from the measurement instrument of interest, and scores for comparator instruments to investigate psychometric properties. Established guidelines can be used to ensure high-quality psychometric property research is being undertaken. Two leading psychometric property guidelines are the COSMIN checklist (Mokkink et al., 2010) and Terwee et al. (2007) criteria. The COSMIN checklist provides detailed guidelines for the methodology of psychometric property research, informing researchers how to design their studies and statistically analyse their data (Mokkink et al., 2010). The Terwee et al. (2007) criteria assists researchers interpret the results of their psychometric property tests, providing cut-offs for determining whether the results of research demonstrate psychometric adequacy. Both these guidelines are widely used to investigate the psychometric properties of health-related measurement instruments.

# Appendix 11: Harter's Self Perception Profile Children (SPPC)

# What I Am Like

Na	ame		Age	Birthday		[	Boy 🗌 Girl	
					Month	Day	(check one)	
	Really True for me	Sort of True for me					Sort of True for me	Really True for me
			San	nple Sent	tence			
a.			Some kids would rather play outdoors in their spare time	BUT	Other watch	kids would rathe T.V.	ș <b>r</b>	
1.			Some kids feel that they are very good at their school work	BUT	Other whether school them	kids worry abou er they can do th work assigned	t to	
2.			Some kids find it hard to make friends	BUT	Other easy to	kids find it pretty o make friends		
3.			Some kids do very well at all kinds of sports	BUT	Other they a when i	kids don't feel th re very good t comes to sport	nat Is	
4.			Some kids are happy with the way they look	BUT	Other happy look	kids are <i>not</i> with the way the	ey 🗌	
5.			Some kids often do not like the way they behave	вит	Other the wa	kids usually like y they behave		
6.			Some kids are often unhappy with themselves	BUT	Other please thems	kids are pretty d with elves		
7.			Some kids feel like they are just as smart as other kids their age	BUT	Other sure a are as	kids aren't so nd wonder if the smart	y	
8.			Some kids know how to make classmates like them	BUT	Other how to classm	kids don't know make nates like them		
9.			Some kids wish they could be a lot better at sports	BUT	Other good e	kids feel they an enough at sports	e	
10.			Some kids are happy with their height and weight	BUT	Other height differe	kids wish their or weight were nt		
11.			Some kids usually do the right thing	BUT	Other do the	kids often don't right thing		

	Really True for me	Sort of True for me				Sort of True for me	Really True for me
12.			Some kids don't like the		Other kids do like the		
			way they are leading their life	BUT	way they are leading their life		
13.			Some kids are pretty slow in finishing their school work	BUT	Other kids can do their school work quickly		
14.			Some kids don't have the social skills to make friends	BUT	Other kids do have the social skills to make friends		
15.			Some kids think they could do well at just about any new sports activity they haven't tried before	BUT	Other kids are afraid they might not do well at sports they haven't ever tried		
16.			Some kids wish their body was different	BUT	Other kids like their body the way it is		
17.			Some kids usually act the way they know they are supposed to	BUT	Other kids often don't act the way they are supposed to		
18.			Some kids are happy with themselves as a person	BUT	Other kids are often not happy with themselves		
19.			Some kids often forget what they learn	BUT	Other kids can remember things easily		
20.			Some kids understand how to get peers to accept them	BUT	Other kids don't understand how to get peers to accept them		
21.			Some kids feel that they are better than others their age at sports	BUT	Other kids don't feel they can play as well		
22.			Some kids wish their physical appearance (how they look) was different	BUT	Other kids like their physical appearance the way it is		
23.			Some kids usually get in trouble because of things they do	BUT	Other kids usually don't do things that get them in trouble		
24.			Some kids like the kind of person they are	BUT	Other kids often wish they were someone else		

	Really True for me	Sort of True for me				Sort of True for me	Really True for me
25.			Some kids do very well at their classwork	BUT	Other kids don't do very well at their classwork		
26.			Some kids wish they knew how to make more friends	BUT	Other kids know how to make as many friends as they want		
27.			In games and sports some kids usually watch instead of play	BUT	Other kids usually play rather than just watch		
28.			Some kids wish something about their face or hair looked different	BUT	Other kids like their face and hair the way they are		
29.			Some kids do things they know they shouldn't do	BUT	Other kids hardly ever do things they know they shouldn't do		
30.			Some kids are very happy being the way they are	BUT	Other kids wish they were different		
31.			Some kids have trouble figuring out the answers in school	BUT	Other kids almost always can figure out the answers		
32.			Some kids know how to become popular	BUT	Other kids do not know how to become popular		
33.			Some kids don't do well at new outdoor games	BUT	Other kids are good at new games right away		
34.			Some kids think that they are good looking	BUT	Other kids think that they are not very good looking		
35.			Some kids behave themselves very well	BUT	Other kids often find it hard to behave themselves		
36.			Some kids are not very happy with the way they do a lot of things	BUT	Other kids think the way they do things is fine		