

Does physical activity in childhood or adolescence predict future anxiety, depression, or wellbeing? A systematic review of 98 prospective cohort studies.

Christopher Knowles^{1*}, Kyle F. Paradis¹, Stephen Shannon¹, Gavin Breslin², Angela Carlin¹

¹Sport and Exercise Sciences Research Institute, School of Sport, Ulster University, Belfast, United Kingdom

²Bamford Centre for Mental Health and Well-being, School of Psychology, Ulster University, Coleraine, United Kingdom

*Corresponding author

Email: knowles-c2@ulster.ac.uk

Abstract

Physical activity (PA) is a modifiable health behaviour that can support and improve mental health. The tendency for activity levels to track over time suggests that through continued participation, those most active in youth may experience better long-term mental health than less active peers. Exploring the extent to which mental health benefits of child/adolescent PA persevere over time helps advocates and policy-makers determine whether PA promotion in youth constitutes effective allocation of public health resources and a viable means of reducing the global burden of common mental disorders and suboptimal wellbeing across the lifespan. This systematic review synthesised evidence for childhood/adolescent PA as a predictor of future anxiety, depression and wellbeing indicators (PROSPERO: CRD42021242555). Systematic searches of CINAHL, Medline, PsycInfo, and Scopus (conducted 27th October 2022) returned 12,703 articles with 98 retained. Included: prospective, quantitative, longitudinal designs; PA measured ages 5-17 years; anxiety, depression, and/or wellbeing indicators measured at least 12 months later; generally healthy populations. Excluded: severe mental illnesses. Risk of bias and quality of evidence were assessed following the GRADE framework. Effect sizes from moderate- and high-quality studies are discussed narratively. Studies are grouped by mental health outcome, PA type, domain, and whether participants adhered to current World Health Organisation PA guidelines. PA was negatively associated with depression in 41/67 studies (61.2%); anxiety in 20/35 studies (57.1%); and positively associated with wellbeing in 25/39 studies (64.1%). Effect sizes for all outcomes were small and evidence quality low across all timeframes. Team sports were the most consistently predictive type of PA. Whether domain-specific activities or guideline (non)adherence have differential effects remains unclear. Heterogeneity was attributed to the vastness of associations tested and psychometric measures used. Current literature offers low-quality partial evidence that childhood/adolescent PA has small beneficial effects for prospective anxiety, depression and wellbeing at least 12 months later.

Keywords: Adolescents, Anxiety, Children, Depression, Physical Activity, Wellbeing

1 Does physical activity in childhood or adolescence predict future
2 anxiety, depression or wellbeing? A systematic review of
3 98 prospective cohort studies

4 **Introduction**

5 **Physical activity & mental health**

6 There is global consensus that Physical Activity (PA) can be a viable low-risk, low-cost
7 mental health intervention throughout the lifespan and across the world. Researchers and leading
8 public health authorities are agreed that PA can be effective at managing depression [1], anxiety
9 [2] and wellbeing [3]. The European Psychiatric Committee declared that PA can be at least as
10 effective as antidepressant drugs in reducing depressive symptoms [4]. A scientific report
11 submitted to the US Secretary of Health and Human Services found strong evidence that moderate-
12 to-vigorous PA (MVPA) can reduce state and trait anxiety in multiple age groups [5]. The
13 Australian Department of Health and Aged Care recommends that following their 24-hour
14 movement guidelines, should result in better wellbeing for all age groups, with low risk of serious
15 injury [6],[7].

16 **A two-continua approach to mental health**

17 Keyes' two-continua model recognises mental illness and wellbeing as two independent,
18 albeit co-related, concepts existing within overall mental health [8]. Keyes' model suggests that
19 wellbeing is not synonymous to the absence of mental illness. Rather, individuals can and should
20 actively pursue social, emotional, and psychological wellbeing outcomes alongside any steps taken
21 to reduce their risk of illness [9]. An individual with high levels of wellbeing and low levels of
22 mental illness may be described as flourishing [10],[11].

23 Flourishing is the epitome of mental health. A nationally representative study from the
24 Netherlands reported that only 37% of the population were flourishing [12] inferring many people

25 are failing to reach their potential in areas such as psychosocial functioning, occupational
26 productivity, and healthcare utilisation [9]. Whilst up-to-date research is required to substantiate
27 these statistics in the present day, it remains likely that sub-optimal mental health is having a
28 greater impact than is commonly understood. Health behaviours such as PA that have potential to
29 increase the prevalence of flourishing through management of mental illness symptomology and
30 wellbeing support may be best placed to help build a more prosperous and healthy society.

31 **Anxiety & depression**

32 The most Common Mental Disorders (CMDs) – Anxiety and Depression – are leading
33 contributors to the global burden of disease and are becoming increasingly prevalent over time
34 [13]. Depression and anxiety affect 280 million (23 million children/adolescents) and 301 million
35 (58 million children/adolescents) globally [13] with one in three individuals likely to experience a
36 CMD at some point in their lives [14]. From an economic-productivity perspective, scaling up
37 treatments for CMDs is projected to have a combined benefit of \$310 billion globally by 2030
38 [15]. From a health perspective, this equates to 43 million extra healthy life years [15],[16]. The
39 investment needed to sufficiently scale-up treatments is considerable, making low-cost methods
40 of managing CMDs including through effective PA promotion, highly desirable.

41 **Wellbeing**

42 Although an historically elusive construct to define, wellbeing can be considered a state of
43 positive feeling and meeting one's full potential in the world [17]. Whilst CMDs consist of
44 symptoms of mal- or sub-optimal-functioning, wellbeing consists of symptoms of positive
45 functioning and/or positive characteristics [18]. Compared to CMDs, indicators of wellbeing have
46 been less heavily researched thus, the economic impact of poor wellbeing is not widely known.
47 That being said, elevated levels of wellbeing have been associated with greater productivity, more
48 meaningful relationships and pro-social behaviour [18],[19] all of which have economic and
49 societal implications.

50 **Children & adolescents**

51 Despite evidence for the benefits of PA, global rates of inactivity remain high. In 2019, the
52 WHO updated their Global Action Plan for PA in which they set out to achieve a 15% reduction
53 in inactivity by 2030 [20]. On their current trajectory, it appears the WHO may fall short of that
54 target [21],[22]. Given that resources dedicated to PA promotion on a national level are often
55 limited, public health would benefit from a better understanding of where investment in PA
56 promotion is likely to see the best return with regards to long-term mental health benefits. To this
57 end, it is possible that children and adolescents are a key target population.

58 Whilst associations between PA and mental health in younger populations are less
59 consistently observed than in adults, previous reviews have concluded there is partial evidence for
60 a causal link between child/adolescent PA, lower levels of mental illness, and greater wellbeing
61 [23]-[25]. One important causal pathway through which child/adolescent PA may support long-
62 term mental health is in the tendency for PA levels to track across the lifespan [26],[27]. Whilst
63 rates of activity generally decrease as we age, those most active early in life tend to remain the
64 most active as adults [27]. This means interventions successful at upscaling PA in children and
65 adolescents not only have potential to benefit young people in the short-term but may improve
66 future rates of participation and subsequently, future mental health [28]. Such an outcome would
67 constitute an effective use of public health resources, particularly for young people – a high risk
68 group due to high rates of inactivity and the early age of onset of symptoms of mental illness [29]-
69 [31].

70 A comprehensive systematic review of the literature investigating associations between
71 child/adolescent PA and future mental health outcomes has not been conducted. To better
72 understand these longitudinal associations, a review should consider relationships measured over
73 lengthy timeframes with outcomes pertaining to both mental illness and wellbeing. A time interval
74 of at least one year between first PA measurement and prospective mental health outcomes

75 represents a significant window in which to capture psychological developmental changes and
76 would strengthen understanding of how long associations persist over time [32],[33].

77 **The current study**

78 The aim of the present systematic review was to synthesise evidence for a longitudinal
79 relationship between PA participation in youth and prospective mental health outcomes measured
80 at least one year later. To align with Keyes' two-continua model, the review included measures of
81 mental illness (i.e., anxiety, depression) and indicators of social, emotional and psychological
82 wellbeing. The review has been reported in line with the recently updated Preferred Reporting
83 Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [34]. The following
84 review questions were written *a priori*:

- 85 1. Can PA performed in youth reduce depressive disorders/symptoms at least 12 months later?
- 86 2. Can PA performed in youth reduce anxiety disorders/symptoms at least 12 months later?
- 87 3. Can PA performed in youth improve indicators of wellbeing at least 12 months later?

88 To offer the most holistic insight of associations observed to date we reviewed PA in its
89 broadest sense, “any form of bodily movement performed by skeletal muscles that results in an
90 increase in energy expenditure” [20],[35]. To make findings useful in a real-world context, we then
91 sorted the PA performed into subgroups where possible. Public health messaging consistently
92 emphasises the importance of children/adolescents achieving WHO guidelines of 60minutes
93 MVPA per day. Additionally, effects are often dependent on the type of activity (i.e., individual or
94 group-based) [36],[37] and the life domain it is performed in (i.e., leisure, occupational/school-
95 based, transport-based, household) [38],[39]. Therefore, as secondary lines of enquiry, we also
96 aimed to synthesise evidence for the effects of (non)adherence to current WHO MVPA guidelines,
97 individual versus group-based activities, and PA in different life domains where possible.

98 **Materials & methods**

99 The review protocol was published on PROSPERO on 15th March 2021 prior to any formal
100 database searches and can be accessed via the registration number: CRD42021242555.

101 **Eligibility criteria**

102 **Participants**

103 Included studies were restricted to those published in English using generally healthy
104 populations [38]. Studies were also restricted to those that collected PA data from school-aged
105 children/adolescents between the ages of 5 and 17 years. Many PA interventions in youths are
106 carried out in-school so using the school years as our upper and lower age thresholds seemed a
107 sensible way to define the target population. Further, the WHO propose PA guidelines for those
108 aged 5 to 17 years old so using these age thresholds best facilitated a review of the effects of
109 adherence to the current youth guidelines. If studies included samples where some participants
110 were within this range and some not (e.g., 15 to 20 years old), they were included if the mean age
111 of the sample fell within the target range. At no point did review authors have access to information
112 that could be used to personally identify participants.

113 **Study length**

114 When first conceptualising the review, the original aim was to capture studies that followed
115 participants as they moved from one key life stage to another. In principle, this was to help answer
116 whether more active young people make more mentally healthy adults. Operationally, this proved
117 difficult and came with the risk that a three-year-long study may be excluded if it were performed
118 between the ages of 6 and 9 (i.e., all within childhood) yet a one-year-long study would be included
119 if it followed participants from age 17 to 18 thus covering what arguably denotes a transition from
120 adolescence into adulthood. Ongoing debate regarding the age at which one life stage ends and
121 another begins added a further layer of complexity to achieving this research aim and made robust

122 methodology elusive [40]. The most inclusive and methodologically coherent approach to achieve
123 our research aim was to target studies that analysed mental health data collected at least 12 months
124 after initial PA measurement. Whilst this meant that there would be studies included that did not
125 follow participants as they moved through different life stages, it would offer a more complete
126 insight into the perseverance of PA-related mental health benefits over time and limit the chance
127 that important studies were overlooked.

128 **Measures**

129 Short-term recall questionnaires typically gather information pertaining to the past few
130 weeks or months whilst quantitative history questionnaires probe memories from several years ago
131 or across an entire lifespan [41]. To minimise the risk of recall bias, measures of PA had to
132 constitute short-term recall (i.e., PA within the last year) or be collected by wearable devices.

133 Included studies investigated anxiety, depression, and/or indicators of wellbeing. As the
134 two-continua model views mental illness and wellbeing as independent concepts, measures were
135 excluded if they combined items targeting depression/anxiety and wellbeing into one overall
136 mental health score. The focus was on CMDs rather than severe mental illnesses (e.g., suicidal
137 ideation, psychosis) due to their comparatively greater contribution to the overall global burden of
138 disease [13].

139 An operational definition of flourishing identifies 10 symptoms said to represent the mirror
140 opposite of CMDs [18]. Competence, engagement, meaning, and positive relationships represent
141 elements of positive functioning whilst emotional stability, optimism, positive emotions,
142 resilience, self-esteem, and vitality represent positive characteristics. Including all 10 factors in a
143 single review would be unwieldy. Instead, three items were identified through group discussion
144 that either represent or heavily contribute to social, emotional, and psychological wellbeing as per
145 the two-continua model and the definition of flourishing discussed. Positive emotions (e.g.,
146 happiness, life satisfaction) are essential criteria for flourishing hence, these measures qualified as

147 indicators of emotional wellbeing [18]. Positive relationships (e.g., social competence, peer-
148 relations) and self-esteem represented indicators of (1) positive functioning/social wellbeing, and
149 (2) positive characteristics/psychological wellbeing, respectively. Self-esteem was selected as a
150 requisite for flourishing over other positive characteristics given its importance to schools and
151 other key stakeholders.

152 **Search strategy, study selection and data collection**

153 Systematic searches of CINAHL Complete, Medline, PsycInfo, and Scopus were
154 conducted from inception to 27th October 2022. Key words used to capture studies of CMDs were
155 derived from DSM-5 definitions of anxiety/depressive disorders. Huppert and So's operational
156 definition of flourishing [18] and Keyes' two-continua model [8] were used to identify key words
157 relating to positive emotions (emotional wellbeing), positive relationships (social wellbeing) and
158 self-esteem (psychological wellbeing). Reference lists of similar existing reviews were manually
159 searched as were authors own libraries. Search results were uploaded to online systematic review
160 software Covidence where eligibility criteria were visibly displayed above each article. Authors
161 confirmed each point sequentially to determine eligibility. Full search strategies for all databases
162 and a succinct version of the eligibility criteria can be found in the published review protocol.

163 Articles first underwent title and abstract screening by a single reviewer removing those
164 that were clearly irrelevant to the focus of the review. It is not essential that this stage be carried
165 out by two reviewers [42] nevertheless, 1270 excluded studies (10%) were screened again by a
166 second reviewer. Dual screening a random subset of excluded studies allowed reviewers to
167 estimate the number of studies mistakenly omitted and assess inter-rater reliability. Next, two
168 authors independently reviewed the full texts of remaining articles to determine which should be
169 included in final analysis. Disputes were resolved through discussion with a third and/or fourth
170 reviewer. Data extraction was conducted by one reviewer using a bespoke data extraction template.

171 **Data items**

172 Data were extracted on study characteristics including authors, location, study design, the
173 name of the dataset used (if applicable), and the data collection schedule, as well as the age,
174 number, and sex/gender split of participants. Further information included: the source of PA and
175 mental health data (e.g., self-report, accelerometers); measures used; association(s) tested;
176 confounding variables; analysis performed; and results. Information provided by study authors
177 and/or the wording of specific items were used to categorise the type and domain of PA, and
178 whether participants adhered to PA guidelines. Further information was extracted to aid quality of
179 evidence and risk of bias assessment, explained in detail below.

180 **Evidence quality & risk of bias**

181 A quality of evidence and risk of bias assessment was made for each study by two
182 independent reviewers using the framework proposed by the GRADE Working Group [43].
183 Evidence was graded as high, moderate, low, or very low. The items used to investigate certainty
184 of evidence, sources of heterogeneity, and risk of bias have been published elsewhere [43].

185 **Synthesis methods**

186 Included studies were heterogeneous with regards to the timespans covered and
187 associations tested. After discussion, authors agreed the review was best suited to an exploratory
188 analysis of potential sources of heterogeneity through narrative synthesis to offer more nuanced
189 conclusions. Results were therefore synthesized according to the Synthesis Without Meta-Analysis
190 (SWiM) reporting guidelines [44].

191 Studies reporting statistically significant associations at the 0.05 level were considered as
192 providing evidence of the direction of an effect. The total number of studies reporting evidence
193 (or lack thereof) an effect were summed and grouped by mental health outcome and length of
194 follow-up in years to provide an overview of the literature as a whole and the timescales across

195 which childhood/adolescent PA was statistically significantly predictive of future mental health.
196 This approach has been taken in existing reviews in other areas of public health[45]. However, the
197 large sample sizes often used meant it was not surprising to find a considerable number of
198 statistically significant associations. To avoid over-reliance on statistical significance as evidence
199 of an effect and the risk that inappropriate conclusions would be drawn from low quality evidence;
200 effect sizes reported in moderate- and high-quality studies were interpreted and discussed using
201 thresholds proposed in existing research [46],[47].

202 Finally, the main analysis was supplemented with important contextual information on
203 studies investigating type/domain of activity and guideline adherence to strengthen depth of
204 understanding. Summary of Findings (SoF) tables for each secondary review question include a
205 brief explanation of the associations tested and effect sizes reported in each study (see supporting
206 information). Again, this approach has been used in existing reviews [48].

207 All studies were eligible for data synthesis provided they did not replicate research from
208 other studies using the same dataset. This was to limit the risk that data from a single source skewed
209 conclusions drawn [49]. Supporting information (File S2) has been provided on the study
210 modifiers used to differentiate between studies using the same datasets to ensure each made a
211 unique contribution to knowledge. Modifiers related to Participants, Exposure, Outcomes and
212 Study Design.

213 **Results**

214 **Study selection**

215 After removal of duplicates, systematic searches returned 12,703 articles to be screened by
216 title and abstract. A total of 496 articles underwent full text review. Conflicts were resolved and
217 data was extracted from 100 studies. A further two studies were removed^(74, 86) post-extraction as
218 they replicated research conducted by other included studies using the same data-source.

219 Supporting information has been provided detailing the decision-making process that led to the
220 exclusion of these studies (Supporting Table S1). This left 98 studies eligible for synthesis.

221 Of the 1270 excluded studies reviewed by a second reviewer, 25 (2.3%) were retained for
222 full text review of which only 1 (0.1%) was included in final analysis. This speaks to the rigor with
223 which title and abstract screening was carried out indicating the use of a single reviewer at this
224 stage did not meaningfully detract from review reliability. A detailed report of the number of
225 articles returned per database and the reasons for exclusion are presented in the PRISMA flowchart
226 (Fig 1).

227 *** **Figure 1: PRISMA flowchart** ***

228 **Study characteristics**

229 *Table 1* provides an index of all original 100 articles. Of the 98 articles included in the
230 synthesis, there were 94 observational cohort studies (95.9%) and four randomized trials^{(8, 49, 53,}
231 ⁵⁴⁾. Studies were conducted throughout North America ($n = 48$), Europe ($n = 37$), Australia ($n =$
232 11) and Asia ($n = 2$). Secondary analysis was conducted by 81 studies using data from 53 different
233 longitudinal cohort study projects. The most heavily used data source was the National
234 Longitudinal Study of Adolescent to Adult Health (Add Health) used in 10 separate studies^{(10, 15,}
235 ^{19, 21, 25, 44, 48, 61, 96, 98)} followed by the Nicotine Dependence in Teens Study (NDIT)^(5, 14, 39, 50, 63).
236 The remaining 17 studies appeared to use primary data collected by the lead author or their
237 research partners^(16, 20, 27, 28, 38, 59, 60, 62, 65, 66, 67, 68, 82, 83, 92, 94, 100). The most heavily researched mental
238 health outcome was depression, investigated in 67/98 studies (68.4%) compared to 35/98 (35.7%)
239 investigating anxiety and 39/98 (39.8%) investigating wellbeing. A quality of evidence assessment
240 was conducted for each study: High ($n = 4$), Moderate ($n = 10$), Low ($n = 59$), or Very Low ($n =$
241 25). Overall, the quality of evidence investigating the association between childhood/adolescent
242 PA and mental health outcomes at least 12 months later was low for anxiety, low for depression,
243 and low for indicators of wellbeing.

244 **Table 1**245 *Index of all included studies and the mental health outcomes they investigated*

ID	Authors (Date of Publication)	Mental Illness			Wellbeing			ID	Authors (Date of Publication)	Mental Illness			Wellbeing		
		Dep.	Anx.	Psych.	Soc.	Emo.	Dep.			Anx.	Psych.	Soc.	Emo.		
1	Adachi & Willoughby (2014)			✓			30	Gunnell et al. (2016a)	✓	✓			✓		
2	Agans et al. (2017)	✓					31	Gunnell et al. (2016b)	✓	✓					
3	Ahn et al. (2018)		✓		✓		32	Hamer et al. (2020)	✓						
4	Ames & Leadbeater (2018)	✓					33	Herman et al. (2010)					✓		
5	Ashdown-Franks et al. (2017)		✓				34	Hoegh-Poulsen et al. (2016)	✓						
6	Barber et al. (2001)	✓	✓	✓			35	Howie et al. (2016)	✓	✓					
7	Bedard et al. (2020)				✓		36	Howie et al. (2020)	✓	✓					
8	Bell et al. (2019)		✓	✓	✓	✓	37	Isaksson et al. (2020)	✓	✓					
9	Birkeland et al. (2009)	✓					38	Jerstad et al. (2010)	✓						
10	Bohr et al. (2019)	✓					39	Jewett et al. (2014)	✓						
11	Brière et al. (2020)	✓	✓				40	Kalimo & Vuori (1990)						✓	
12	Brière et al. (2018)	✓	✓				41	Kandola et al. (2020a)	✓						
13	Brunes et al. (2015)	✓	✓			✓	42	Kandola et al. (2020b)		✓					
14	Brunet et al. (2013)	✓					43	Kaseva et al. (2016)	✓						
15	Carter et al. (2015)	✓					44	Kort-Butler & Hageven (2011)				✓			
16	Carvajal (2012)	✓					45	Mahoney & Vest (2012)	✓	✓		✓	✓	✓	
17	Clark et al. (2007)	✓					46	Marsh & Kleitman (2003)				✓			
18	Colman et al. (2014)	✓					47	McKercher et al. (2014)	✓						
19	Daniels & Leaper (2006)				✓		48	McPhie & Rawana (2015)	✓						
20	Denalt et al. (2009)	✓					49	Motl (2004)	✓						
21	Deshpande et al. (2020)	✓	✓				50	Murray et al. (2021)	✓	✓					
22	Doré et al. (2020)			✓	✓	✓	51	Nigg et al. (2021)		✓			✓		
23	Doré et al. (2019)			✓	✓	✓	52	Noordstar et al. (2016)				✓			
24	Duncan et al. (2021)	✓	✓	✓	✓	✓	53	Olive et al. (2019)	✓						
25	Easterlin et al. (2019)	✓	✓				54	Omorou et al. (2016)					✓		
26	Fløtnes et al. (2011)	✓	✓				55	Patte et al. (2020)	✓						
27	Gómez-Baya et al. (2018)					✓	56	Perron et al. (2012)	✓						
28	Gómez-Baya et al. (2020)	✓					57	Piché et al. (2019)	✓						
29	Graupensperger et al. (2021)	✓	✓		✓		58	Rangul et al. (2012)	✓					✓	

ID	Authors (Date of Publication)	Mental Illness					Wellbeing						
		Dep.	Anx.	Psych.	Soc.	Emo.	ID	Authors (Date of Publication)	Dep.	Anx.	Psych.	Soc.	Emo.
59	Raudsepp & Neissaar (2012)	✓					80	Toseeb et al. (2014)	✓				
60	Raudsepp & Vink (2019)	✓					81	Van Dijk et al. (2016)	✓		✓		
61	Rees & Sabia (2010)	✓		✓			82	Visser et al. (2020)			✓	✓	
62	Rongen et al. (2020)				✓		83	Wagnsson et al. (2014)			✓		
63	Sabiston et al. (2016)	✓					84	Wang et al. (2017)	✓				
64	Sagatun et al. (2007)		✓		✓		85	Wiles et al. (2008)		✓			
65	Sánchez-Oliva et al. (2020)	✓				✓	86	Wu et al. (2017)	✓	✓			
66	Schmalz et al. (2007)			✓			87	Wu et al. (2018)	✓				
67	Schumacher-Dimech & Seiler (2011)		✓				88	You et al. (2021)	✓				✓
68	Seeley et al. (2009)	✓					89	Wilson et al., (2022)					✓
69	Shin & You (2013)					✓	90	Redig et al., (2022)	✓				
70	Slutzky & Simpkins (2009)			✓			91	Kandola et al., (2022)		✓			
71	Slykerman et al. (2020)	✓					92	Barth Vedøy et al., (2021)			✓	✓	✓
72	Spreitzer (1994)			✓			93	O'Flaherty et al., (2022)	✓	✓		✓	
73	Stavrakakis et al. (2012)	✓					94	Jetten et al., (2022)	✓	✓			
74	Stavrakakis et al. (2013a)	✓					95	Harbec et al., (2022)	✓	✓			
75	Stavrakakis et al. (2013b)	✓					96	García-Hermoso et al., (2022)	✓				
76	Stein et al. (2007)			✓	✓		97	Buchan et al., (2021)	✓	✓			
77	Ströhle et al. (2007)	✓	✓				98	Iverson et al., (2021)	✓				
78	Suetani et al. (2017)	✓	✓				99	White-Gosselin et al, (2022)	✓	✓			
79	Sund et al. (2011)	✓					100	Latina et al., (2022)	✓		✓		

247 Studies were published from 1990–2022. There were noticeable upticks in publications in
248 2007, 2016, and 2020 signalling a gradual but consistent increase in research output in the area
249 (Fig 2). A total of 75/98 (76.5%) articles were published from 2012-2022 and around half (48/98,
250 49.0%) were from 2017 onward meaning most of the research reviewed was recently published
251 and up to date at the time of writing. Evidence spanned 65 years with studies availing of PA and
252 mental health data collected from 1955⁽⁴⁰⁾ to 2020⁽⁹⁰⁾. Moderating effects of time are discussed
253 herein. All studies used data collected before the outbreak of the Covid-19 pandemic. It remains
254 to be seen whether young people engaging in PA before Covid-19 experienced mental health
255 benefits upon emergence from the pandemic compared to their previously less active peers.

256 *** **Figure 2: Year of publication** ***

257 Self-report measures of PA were used in 88 studies, 11 used mechanical device-measured
258 data^(3, 8, 32, 41, 42, 65, 71, 80, 81, 82, 92). One study used details of a football academy’s weekly training
259 schedule as its PA data source⁽⁶²⁾ and one study used trained observers reports of children in a PE
260 class⁽⁵³⁾. All mental health outcomes were self-reported other than in one study using national
261 medical databases to determine rates of clinical diagnoses of internalizing disorders⁽⁸⁷⁾. A key for
262 all acronyms used for mental health measures can be found at the bottom of Tables 2, 3, and 4,
263 and in the supporting information. Studies used 23 different validated measures of depression of
264 which the CES-D [50] was the most popular with 20/67 depression studies (30.0%) using a version
265 of the scale^(2, 10, 12, 15, 16, 21, 24, 25, 34, 48, 49, 55, 59, 60, 61, 71, 81, 96, 97, 100). There were also 19 validated
266 measures of anxiety and 15 validated scales measuring indicators of wellbeing. The most common
267 measure of anxiety symptoms was the SDQ Emotional Symptoms subscale [51] used in 6/35
268 (17.1%) of anxiety studies^(3, 8, 51, 64, 85, 91).

269 **Table 2**
 270 *Studies investigating associations between physical activity and depression*

ID	Dataset	Participants (n)	Age (Baseline)	Age (Follow-Up)	Study Length (Years)	Measure (Subscale)	Evidence of Association	Quality of Evidence
90	NCDS	18,555	7	62	55	*Malaise (depression)	Y	High
87	CLASS	4861	10	18	8	MSID; CIHIDAD	Y	High
61	Add Health	13,426	12-17	13-18	1	*CES-D	Y	Moderate
96	Add Health	7069	12-17	33-39	21	*CES-D; Diagnosis	Y	Moderate
41	ALSPAC	4257	12	18	6	*CIS	Y	Moderate
63	NDIT	860	12	20	8	*MDI	Y	Moderate
39	NDIT	853	12	20	8	*MDI	Y	Moderate
53	LOOK	821	8	12	4	*CDI	N	Moderate
99	FLUPPY	785	6	10	4	*PBQ; *CBC	Y	Moderate
50	NDIT	733	12	20	8	*MDI	N	Moderate
97	COMPASS	23,395	14-18	15-19	1	*CES-D	Y	Low
10	Add Health	10951	12-17	29	13	*CES-D	Y	Low
25	Add Health	4888	12-17	24-32	13	Diagnosis; *CES-D	Y	Low
49	TEENS	4594	12	14	2	*CES-D	Y	Low
29	LSAC	3956	12	16	4	*SMFQ	Y	Low
93	LSAC	3885	13	15	2	*SMFQ	N	Low
48	Add Health	3676	15	28	13	*CES-D	Y	Low
43	CRYF	3596	9-18	35-50	32	*MDI	Y	Low
78	MUSP	3493	14	21	7	*CIDI	Y	Low
24	COMPASS	3173	14-18	15-19	1	*CES-D	N	Low
77	EDSP	2548	14-24	22-32	8	*CIDI	Y	Low
32	MCS	2416	7	14	7	*SMFQ	Y	Low
55	COMPASS	2292	15-17	16-18	1	*CES-D	N	Low
73	TRAILS	2230	11	16	5	*YSR; *CBC	Y	Low
21	Add Health	2197	12-17	24-32	14	*CES-D	N	Low
79	YAMHS	2175	14	15	1	*SMFQ	Y	Low
88	KCYPS	2092	13	15	2	*KEBPS	Y	Low
26	HUNT	2000	12-17	17-19	5	*HSCL	N	Low

58	HUNT	1869	13-19	23-31	10	*CONOR-MHI	Y	Low
35	Raine	1679	5	20	15	*DASS-21	Y	Low
47	CDAH	1630	12-15	29-35	20	*CIDI	N	Low
34	VestLiv	1589	14-15	20-21	6	*CES-D	Y	Low
57	QLSCD	1516	7	8	1	*SBQ	Y	Low
17	RELACHS	1513	11-14	13-16	2	*SMFQ	N	Low
11	QLSCD	1492	6	12	6	*SBQ; *PBQ	Y	Low
95	QLSCD	1438	5	10	5	*SBQ	Y	Low
37	SALVe	1428	14	17	3	*DSRS	Y	Low
13	HUNT	1417	12-13	17-19	5	*HSCL	N	Low
75	TRAILS	1396	14	19	5	*CIDI	N	Low
56	QLSCD	1250	8	10	2	*PBQ; *CBC	Y	Low
31	REAL	1160	12-15	17-20	11	*CDI	N	Low
65	UP&DOWN	1099	8-18	10-20	2	*PANAS	N	Low
84	CAB	1065	12-15	14-17	2	*HSCL	Y	Low
100	SSP	998	12-14	13-15	1	*CES-D	N	Low
9	NLHB	924	13	23	10	7-items: depressed mood	N	Low
71	ABC	871	11	16	5	*CES-D	N	Low
14	NDIT	816	12	20	8	*MDI	N	Low
80	ROOTS	736	15	18	3	*SMFQ; *SADSSAC	N	Low
4	VHYS	662	12-18	22-29	10	*BCFPI	Y	Low
28	n/a (Primary Data)	525	13-14	16-17	3	*CDI	Y	Low
38	n/a (Primary Data)	496	11-15	17-21	6	*SADSSAC	Y	Low
12	NANS	17550	12-15	13-16	1	*CES-D	N	Very Low
15	Add Health	6504	12-17	18-26	7	*CES-D	Y	Very Low
2	4-H	5961	13	21	8	*CES-D	Y	Very Low
98	Add Health	2353	12-17	24-32	14	*Diagnosis; 1-item: depression	N	Very Low
36	Raine	1628	8	22	14	*DASS-21	Y	Very Low
18	NPHS	1137	16-17	18-23	6	*CIDI	Y	Very Low
45	PSID	1115	12-18	18-24	6	*KPDS	N	Very Low
30	MATCH	932	9-13	13-17	4	*PedsQL	N	Very Low
6	MSLT	900	15	24	9	3-items: depressed mood	N	Very Low
16	Missing	744	12-13	13-14	1.5	*CES-D	Y	Very Low

94	n/a (Primary Data)	558	12-17	13-18	1	*HADS	N	Very Low
68	n/a (Primary Data)	479	16	20	4	*SADSSAC	Y	Very Low
20	n/a (Primary Data)	362	13	16	3	*CDI	N	Very Low
59	n/a (Primary Data)	277	12	15	3	*CES-D	Y	Very Low
60	n/a (Primary Data)	173	13	15	2	*CES-D	Y	Very Low
81	GOALS	158	14	15	1	*CES-D	N	Very Low

271 *Measure has been previously validated. Datasets: 4-H 4-H Study of Positive Youth Development, ABC Auckland Birthweight Collaborative Study, Add Health The
272 National Longitudinal Study of Adolescent Health, ALSPAC Avon Longitudinal Study of Parents And their Children, CAB Childhood and Beyond Study, CDAH
273 Childhood Determinants of Adult Health Study, CLASS 2003 Children's Lifestyle and School Performance Study, COMPASS Cannabis use, Obesity, Mental health,
274 Physical activity, Alcohol use, Smoking, and Sedentary behaviour study, CRYF Cardiovascular Risk in Young Finns Study, EDSP Early Developmental Stages of
275 Psychopathology, FLUPPY The FLUPPY Program, GOALS The GOALS Study, HUNT The Trøndelag Health Study, KCYPS Korean Children and Youth Panel
276 Survey, LOOK Lifestyle of our Kids Study, LSAC Longitudinal Study of Australian Children, MATCH Monitoring Activities of Teenagers to Comprehend their
277 Habits Study, MCS Millennium Cohort Study, MSLT Michigan Study of Life Transitions, MUSP Mater University of Queensland Study of Pregnancy, NANS New
278 Approaches, New Solutions, NCDS The 1958 National Child Development Survey, NDIT Nicotine Dependence in Teens Study, NLHB Norwegian Longitudinal
279 Health Behaviour Study, NPHS National Population Health Survey, PSID Panel Study of Income Dynamics, QLSCD Quebec Longitudinal Study of Child
280 Development, Raine The Raine Study, REAL Research on Eating and Adolescent Lifestyles Study, RELACHS Research with East London Adolescents: Community
281 Health Survey, ROOTS The ROOTS Study, SALVe The SALVe Cohort, SSP The Seven Schools Project, TEENS The TEENS Study, TRAILS The Trails Study,
282 UP&DOWN The Up and Down project, VHYS Victoria Health Youth Survey, YAMHS Youth and Mental Health Study. Measures: BCFPI Brief Child and Family
283 Phone Interview, CBC Child Behaviour Checklist CDI Child Depression Inventory, CES-D Centre for Epidemiological Studies Depression Scale, CIDI Composite
284 International Diagnostic Interview, CIHIDAD Canadian Institute for Health Information Discharge Abstract Database, CIS Clinical Interview Schedule, CONOR-
285 MHI Cohort of Norway Mental Health Index, DASS-21 Depression Anxiety Stress Scales, DSRS Depression Self-Rating Scale, HADS Hospital Anxiety and
286 Depression Scale, HSCL Hopkins Symptom Checklist, KEBPS Korean Emotional Behavioural Problem Scale, KPDS Kessler's Non-Specific Psychological Distress
287 Scale, Malaise The Malaise Inventory, MDI Major Depression Inventory, MSID Medical Services Insurance Database, PANAS Positive and Negative Affect
288 Schedule, PBQ Preschool Behaviour Questionnaire, PedsQL Pediatric Quality of Life Inventory, SADSSAC Schedule for Affective Disorders and Schizophrenia for
289 School-Age Children, SBQ Social Behaviour Questionnaire, SMFQ Short Mood and Feelings Questionnaire, YSR Youth Self-Report

290 **Table 3**
 291 *Studies investigating associations between physical activity and anxiety*

ID	Dataset	Participants (n)	Age (Baseline)	Age (Follow-Up)	Study Length (Years)	Measure (Subscale)	Evidence of Association	Quality of Evidence
87	CLASS	4861	10	18	8	MSID; CIHIDAD	Y	High
8	AHEAD	673	13	15-16	3	*SDQ (Emotional Symptoms)	Y	High
3	MCS	6153	7	11	4	*SDQ (Emotional Symptoms)	N	Moderate
91	MCS	4599	14	17	3	*SDQ (Emotional Symptoms)	Y	Moderate
99	FLUPPY	785	6	10	4	*TRF	Y	Moderate
50	NDIT	733	12	20	8	CCHS	Y	Moderate
97	COMPASS	23,395	14-18	15-19	1	*GAD-7	Y	Low
25	Add Health	4888	12-17	24-32	13	Clinical Diagnosis	Y	Low
42	ALSPAC	4257	12	18	6	*CIS	Y	Low
29	LSAC	3956	12	16	4	*SCAS	Y	Low
93	LSAC	3885	13	15	2	*PedsQL (Emotional)	Y	Low
78	MUSP	3493	14	21	7	*CIDI	N	Low
24	COMPASS	3173	14-18	15-19	1	*GAD-7	N	Low
77	EDSP	2548	14-24	22-32	8	*CIDI	Y	Low
64	HUBRO	2433	15-16	18-19	3	*SDQ (Emotional Symptoms)	Y	Low
21	Add Health	2197	12-17	24-32	14	Clinical Diagnosis	N	Low
88	KCYPS	2092	13	15	2	*KEBPS	Y	Low
26	HUNT	2000	12-17	17-19	5	*HSCL	N	Low
58	HUNT	1869	13-19	23-31	10	*CONOR-MHI	Y	Low
35	Raine	1679	5	20	15	*DASS-21	N	Low
11	QLSCD	1492	6	12	6	*SBQ; *PBQ	Y	Low
85	n/a (Secondary Data)	1446	11-14	12-15	1	*SDQ (Emotional Symptoms)	Y	Low
95	QLSCD	1438	5	10	5	*SBQ	Y	Low
37	SALVe	1428	14	17	3	*SCAS	N	Low
13	HUNT	1417	12-13	17-19	5	*HSCL	N	Low
31	REAL	1160	12-15	17-20	11	*MASC	N	Low
5	NDIT	781	12	20	8	CCHS	Y	Low
12	NANS	17550	12-15	13-16	1	*SCAS (Social Anxiety)	Y	Very Low

36	Raine	1628	8	22	14	*DASS-21	N	Very Low
45	PSID	1115	12-18	18-24	6	*KPDS	N	Very Low
30	MATCH	932	9-13	13-17	4	*PedsQL (Emotional)	N	Very Low
6	MSLT	900	15	24	9	3-items measuring worry	N	Very Low
51	MoMo	686	6	17	11	*SDQ (Emotional Symptoms)	N	Very Low
94	n/a (Primary Data)	558	12-17	13-18	1	*HADS	N	Very Low
67	n/a (Primary Data)	145	7-9	8-10	1	*SPAIC; *PQSACY; *TRF	Y	Very Low

292 *Measure has been previously validated. Datasets: *Add Health* The National Longitudinal Study of Adolescent Health, *AHEAD* Activity and Healthy Eating in
293 Adolescence, *ALSPAC* Avon Longitudinal Study of Parents And their Children, *CLASS* 2003 Children's Lifestyle and School Performance Study, *COMPASS*
294 Cannabis use, Obesity, Mental health, Physical activity, Alcohol use, Smoking, and Sedentary behaviour study, *EDSP* Early Developmental Stages of
295 Psychopathology, *FLUPPY* The FLUPPY Program, *HUBRO* Oslo Health Study, *HUNT* The Trøndelag Health Study, *KCYPS* Korean Children and Youth
296 Panel Survey, *LSAC* Longitudinal Study of Australian Children, *MATCH* Monitoring Activities of Teenagers to Comprehend their Habits Study, *MCS*
297 Millennium Cohort Study, *MoMo* German Motorik-Modul Longitudinal Study, *MSLT* Michigan Study of Life Transitions, *MUSP* Mater University of
298 Queensland Study of Pregnancy, *NANS* New Approaches, New Solutions, *NDIT* Nicotine Dependence in Teens Study, *PSID* Panel Study of Income
299 Dynamics, *QLSCD* Quebec Longitudinal Study of Child Development, *Raine* The Raine Study, *REAL* Research on Eating and Adolescent Lifestyles Study,
300 *SALVe* The SALVe Cohort. Measures: *CCHS* Canadian Community Health Survey 2003, *CIDI* Composite International Diagnostic Interview, *CIHIDAD*
301 Canadian Institute for Health Information Discharge Abstract Database, *CIS* Clinical Interview Schedule, *CONOR-MHI* Cohort of Norway Mental Health
302 Index, *DASS-21* Depression Anxiety Stress Scales, *GAD-7* Generalized Anxiety Disorder Assessment, *HADS* Hospital Anxiety and Depression Scale, *HSCL*
303 Hopkins Symptom Checklist, *KEBPS* Korean Emotional Behavioural Problem Scale, *KPDS* Kessler's Non-Specific Psychological Distress Scale, *MASC*
304 Multidimensional Anxiety Scale for Children, *MSID* Medical Services Insurance Database, *PBQ* Preschool Behaviour Questionnaire, *PedsQL* Pediatric
305 Quality of Life Inventory, *PQSACY* Parents Questionnaire for Social Anxiety in Children and Youth, *SBQ* Social Behaviour Questionnaire, *SCAS* Spence
306 Children's Anxiety Scale, *SDQ* Strengths and Difficulties Questionnaire, *SPAIC* Social Phobia and Anxiety Inventory for Children, *TRF* Teacher Report Form

307 **Table 4**
 308 *Studies investigating associations between physical activity and indicators of wellbeing*

ID	Dataset	Participants (n)	Age (Baseline)	Age (Follow-Up)	Study Length (Years)	Measure (Subscale)	Evidence of Association	Quality of Evidence
54 ^b	PRALIMAP	1445	14-18	16-20	2	*Duke Health Profile (Social)	Y	High
8 ^{b, d}	AHEAD	673	13	15-16	3	*SDQ (Peer Relations); *WEMWBS	N	High
61 ^c	Add Health	13,426	12-17	13-18	1	*RSE	Y	Moderate
3 ^b	MCS	6153	7	11	4	*SDQ (Peer Relations)	Y	Moderate
72 ^c	HSBS	10536	15-17	21-23	6	*RSE	N	Low
19 ^c	Add Health	10500	12-17	13-21	1	4 items: Self-Esteem	Y	Low
89 ^a	Active NZ	6725	11-17	14-21	3	1-item: Subjective Wellbeing	Y	Low
44 ^c	Add Health	5399	12-17	21-26	7	4 items: Self-Esteem	Y	Low
29 ^b	LSAC	3956	12	16	4	*SDQ (Peer Relations)	N	Low
93 ^b	LSAC	3885	13	15	2	*PedsQL (Social)	Y	Low
69 ^a	KCYPS	3188	14	15	1	3 items: Life Satisfaction	Y	Low
24 ^d	COMPASS	3173	14-18	15-19	1	*FS-8	Y	Low
64 ^b	HUBRO	2433	15-16	18-19	3	*SDQ (Peer Relations)	N	Low
7 ^b	PHAST	2278	10	14	4	*HSPP (Social Competence)	Y	Low
88 ^a	KCYPS	2092	13	15	2	*SLSS	N	Low
58 ^a	HUNT	1869	13-19	23-31	10	1 item: Life Satisfaction	Y	Low
1 ^c	n/a (Primary Data)	1492	14	17	3	*RSE	N	Low
13 ^a	HUNT	1417	12-13	17-19	5	3 items: Subjective Wellbeing	Y	Low
83 ^c	n/a (Primary Data)	1174	10-16	13-19	3	*HSPP (Global Self-Worth)	Y	Low
65 ^a	UP&DOWN	1099	8-18	10-20	2	*PANAS	Y	Low
100 ^c	SSP	998	12-14	13-15	1	*RSE	Y	Low
27 ^a	n/a (Primary Data)	714	13-16	14-17	1	*SLS	Y	Low
92 ^d	n/a (Primary Data)	599	13	15	2	*WEMWBS	N	Low
23 ^d	MATCH	318	10	16	6	*MHC-SF	Y	Low
66 ^c	NEEDS	197	9	13	4	*HSPP (Global Self-Worth)	Y	Low
52 ^c	MAS	184	4-7	7-9	2	*HSPP (Global Self-Worth)	Y	Low
62 ^b	n/a (Primary Data)	115	13	14	1	*KIDSCREEN-27	N	Low
76 ^{b, c}	GUTS	8670	10-15	12-18	2	*HSPP (GSW; Social Competence)	Y	Very Low

46 ^c	NELS	4757	15	17	2	7 items: Self-Esteem	Y	Very Low
45 ^d	PSID	1115	12-18	18-24	6	*MHC-SF	Y	Very Low
70 ^c	CAB	987	7-10	8-11	1	*HSPP (Global Self-Worth)	Y	Very Low
30 ^b	MATCH	932	9-13	13-17	4	*PedsQL (Social)	Y	Very Low
6 ^c	MSLT	900	15	24	9	3 items: Self-Esteem	N	Very Low
40 ^a	HCS	706	8-21	31-44	24	3 items: Life Satisfaction	N	Very Low
51 ^b	MoMo	686	6	17	11	*SDQ (Peer Relations)	Y	Very Low
22 ^d	MATCH	424	10	16	6	*MHC-SF	N	Very Low
33 ^b	PALS	310	7-18	29-41	22	*SF-36 (Social Functioning)	N	Very Low
81 ^c	GOALS	158	14	15	1	*RSE	N	Very Low
82 ^{b, c}	n/a (Primary Data)	134	7-9	8-10	1	*KINDLR	N	Very Low

309 *Measure has been previously validated. ^aPositive Emotions, ^bPositive Relations, ^cSelf-Esteem, ^dOverall Wellbeing/Flourishing. Datasets: Active-NZ Active New
310 Zealand Survey, Add Health The National Longitudinal Study of Adolescent Health, AHEAD Activity and Healthy Eating in Adolescence, CAB Childhood and
311 Beyond Study, COMPASS Cannabis use, Obesity, Mental health, Physical activity, Alcohol use, Smoking, and Sedentary behaviour study, GOALS The GOALS
312 Study, GUTS Growing Up Today Study, HCS The Healthy Child Study, HSBS The High School and Beyond Survey, HUBRO Oslo Health Study, HUNT The
313 Trøndelag Health Study, KCYPS Korean Children and Youth Panel Survey, LSAC Longitudinal Study of Australian Children, MAS The Move-Along Study,
314 MATCH Monitoring Activities of Teenagers to Comprehend their Habits Study, MCS Millennium Cohort Study, MoMo German Motorik-Modul Longitudinal
315 Study, MSLT Michigan Study of Life Transitions, NEEDS The Girls' Needs Project, NELS The National Education Longitudinal Study of 1988, PALS Physical
316 Activity Longitudinal Study, PHAST Physical Health and Activity Study Team project, PRALIMAP PRomotion de l'ALIMentation et de l'Activité Physique trial,
317 PSID Panel Study of Income Dynamics, SSP The Seven Schools Project, UP&DOWN The Up and Down project. Measures: FS-8 Flourishing Scale, HSPP Harter's
318 Self-Perception Profile, KINDLR KINDer Lebensqualitätsfrageogen, MHC-SF Mental Health Continuum-Short Form, PANAS Positive and Negative Affect
319 Schedule, PedsQL Pediatric Quality of Life Inventory, RSE Rosenberg Self-Esteem Scale, SDQ Strengths and Difficulties Questionnaire, SF-36 Medical Outcomes
320 Study Short Form-36, SLS Satisfaction with Life Scale, SLSS Students' Life Satisfaction Scale, WEMWBS Warwick-Edinburgh Mental Well-being Scale

321 Indicators of wellbeing were mostly measured using subscales of larger questionnaires
322 typically used to quantify a range of outcomes both related and unrelated to mental health. The
323 FS-8 [52], KINDL-R [53], WEMWBS [54] and MHC-SF [10] were the only dedicated
324 wellbeing/flourishing scales that targeted elements of social, emotional and psychological
325 wellbeing simultaneously. These scales were relatively underutilised, appearing in only 7/39
326 (17.9%) wellbeing studies combined^(8, 22, 23, 24, 45, 82, 92). The RSE^(1, 61, 72, 81, 100) [55] and/or HSPP
327 Global Self-Worth subscale^(52, 66, 70, 76, 83) [56], and the SDQ Peer Relationships subscale^{(3, 8, 29, 51,}
328 ⁶⁴⁾ [51] were the most common measures of self-esteem (psychological wellbeing) and positive
329 relationships (social wellbeing), respectively. Positive emotions (emotional wellbeing) were only
330 investigated using validated questionnaires in three studies using the SLS⁽²⁷⁾ [57], SLSS⁽⁸⁸⁾ [58]
331 and the PANAS Positive Affect subscale⁽⁶⁵⁾ [59]. Additional studies measured emotional
332 wellbeing via (to the best of our knowledge) unvalidated items, so findings should be treated with
333 caution^(13, 40, 58, 69, 89).

334 **Results of synthesis**

335 To address our primary research questions, studies are grouped by mental health outcome
336 into SoF tables structured in order of quality of evidence and sample size (see Tables 2-7). Studies
337 investigating multiple outcomes are reported in all relevant tables. Baseline age represents the age
338 at which PA data was first collected whilst follow-up age represents the age participants last
339 reported on their mental health.

340 **Table 5**
 341 *Effects sizes in moderate- and high-quality depression studies*

ID	Predictor	Outcome	Measure of Effect	Results (unadjusted)	Results (adjusted)	Comments
90	Inactive (ref) vs Active age 7 years	Depressive Symptoms from 23-62	Hazard Ratio (HR), 95%CI	.85 [.68 to 1.08]	.81 [.55 to 1.18]	No difference
	Inactive (ref) vs Active age 11 years	Depressive Symptoms from 23-62	Hazard Ratio (HR), 95%CI	.91 [.78 to 1.07]	.96 [.75 to 1.22]	No difference
	Inactive (ref) vs Active age 16 years	Depressive Symptoms from 23-62	Hazard Ratio (HR), 95%CI	.75 [.68 to .83]	.73 [.62 to .85]	27% less risk of elevated depressive symptoms
87	Weekly PA without coach: Never (ref) vs. 1-3 times age 10 (boys)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.78 [.54 to 1.14]	No difference
	Weekly PA without coach: Never (ref) vs. 4+ times age 10 (boys)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.98 [.69 to 1.38]	No difference
	Weekly PA with coach: Never (ref) vs. 1-3 times age 10 (boys)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.85 [.68 to 1.08]	No difference
	Weekly PA with coach: Never (ref) vs. 4+ times age 10 (boys)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.73 [.55 to .98]	27% less risk of receiving a diagnosis of an internalizing disorder
	Weekly PA without coach: Never (ref) vs. 1-3 times age 10 (girls)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.80 [.61 to 1.06]	No difference
	Weekly PA without coach: Never (ref) vs. 4+ times age 10 (girls)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.77 [.59 to 1.02]	No difference
	Weekly PA with coach: Never (ref) vs. 1-3 times age 10 (girls)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	1.03 [.86 to 1.23]	No difference

	Weekly PA with coach: Never (ref) vs. 4+ times age 10 (girls)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.92 [.71 to 1.18]	No difference
61	Times engaging in PA in past week during adolescence (12-17) (males)	Depressive Symptoms after one year	Unstandardized Beta (B), (SE), <i>p</i>	n/a	-.125 (.027) <i>p</i> < .01	Each additional bout of weekly PA had a small negative effect on depressive symptoms
	Times engaging in PA in past week during adolescence (12-17) (females)	Depressive Symptoms after one year	Unstandardized Beta (B), (SE), <i>p</i>	n/a	-.040 (.039) <i>p</i> > .10	No difference
96	MVPA Guidelines: Didn't Meet (ref) vs. Met in Adolescence (12-17)	Depressive symptoms or disorder age 33-39	Incidence Rate Ratio (IRR), 95%CI	n/a	.84 [.72 to .98]	16% lower incidence of depression
	MVPA Guidelines: Never Met (ref) vs. Met in Adolescence (12-17)	Depressive symptoms or disorder age 33-39	Incidence Rate Ratio (IRR), 95%CI	n/a	1.02 [.86 to 1.08]	No difference
	MVPA Guidelines: Never Met (ref) vs. Met in Adulthood (33-39)	Depressive symptoms or disorder age 33-39	Incidence Rate Ratio (IRR), 95%CI	n/a	.91 [.77 to .99]	9% lower incidence of depression
	MVPA Guidelines: Never Met (ref) vs. Consistently Met (12-17 & 33-39)	Depressive symptoms or disorder age 33-39	Incidence Rate Ratio (IRR), 95%CI	n/a	.56 [.40 to .79]	44% lower incidence of depression
41	PA: Total counts age 12 (per 100)	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.910 [.882 to .939]	.941 [.910 to .972]	Additional 100 counts of daily activity associated with 5.9% reduction in depression
	PA: Total counts age 14 (per 100)	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.933 [.902 to .965]	.965 [.932 to .999]	Additional 100 counts of daily activity associated with 3.5% reduction in depression
	PA: Total counts age 16 (per 100)	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.939 [.896 to .983]	.984 [.937 to 1.033]	No difference in fully adjusted model
	Light PA age 12 (per 60mins)	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.883 [.834 to .933]	.904 [.850 to .961]	Additional 60mins light activity associated with 9.6% reduction in depression

	Light PA age 14 (per 60mins)	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.908 [.851 to .970]	.922 [.857 to .992]	Additional 60mins light activity associated with 7.8% reduction in depression
	Light PA age 16 (per 60mins)	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.882 [.810 to .961]	.889 [.809 to .974]	Additional 60mins light activity associated with 11.1% reduction in depression
	MVPA age 12 (per 15mins)	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.848 [.863 to .965]	.910 [.857 to .966]	Additional 15mins MVPA associated with 9.0% reduction in depression
	MVPA age 14 (per 15mins)	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.913 [.863 to .965]	.960 [.905 to 1.018]	No difference in fully adjusted model
	MVPA age 16 (per 15mins)	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.938 [.883 to .997]	1.001 [.936 to 1.071]	No difference in fully adjusted model
	Total counts: Persistently Low vs. Persistently High from age 12-16	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.934 [.823 to 1.060]	n/a	No difference
	Light: Persistently Low (ref) vs. Persistently Average from age 12-16	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.936 [.843 to 1.040]	n/a	No difference
	Light: Persistently Low (ref) vs. Persistently High from age 12-16	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.804 [.652 to .990]	n/a	Persistently high levels of light activity associated with 19.6% fewer depressive symptoms
	MVPA: Persistently Low (ref) vs. Persistently Above Average from age 12-16	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	1.006 [.884 to 1.145]	n/a	No difference
	MVPA: Persistently Low (ref) vs. Persistently High from age 12-16	Depressive symptoms age 18	Incidence Rate Ratio (IRR), 95%CI	.699 [.515 to .950]	n/a	Persistently high levels of MVPA associated with 30.1% fewer depressive symptoms
63	Years of school sport (team) from age 12-17	Depressive Symptoms age 20	Standardized Beta (β), <i>p</i>	n/a	-.09 <i>p</i> < .05	Small negative association between years of team sport and depressive symptoms

	Years of school sport (individual) from age 12-17	Depressive Symptoms age 20	Standardized Beta (β), p	n/a	.01 $p > .05$	No effect
39	Years of adolescent school sport (any) from age 12-17	Depressive Symptoms age 20	Standardized Beta (β), p	n/a	-.09 $p < .05$	Years of sport had small negative association with depressive symptoms
53	Specialist Taught vs. Regular PE age 8: end of 12-month intervention	Change in depressive symptoms age 9	p	$p > .05$	n/a	No difference
	Specialist Taught vs. Regular PE age 8: after 4 years	Change in depressive symptoms age 12	p	$p > .05$	n/a	No difference
99	All Sport: Low (ref) vs. High trajectory from age 6-10	Depressive symptoms age 10	Standardized Beta (β), 95%CI	n/a	-.10 [-.18 to -.02]	Small negative effect of high sport trajectory on depressive symptoms
	High Team (ref) vs. High Individual sport participation from age 6-10	Depressive symptoms age 10	Standardized Beta (β), 95%CI	n/a	.05 [-.05 to .16]	No difference
	Team Sport: Low (ref) vs High trajectory from age 6-10	Depressive symptoms age 10	Standardized Beta (β), 95%CI	n/a	-.12 [-.20 to -.03]	Small negative effect of high team sports trajectory on depressive symptoms
	Individual Sport: Low (ref) vs. High trajectory from age 6-10	Depressive symptoms age 10	Standardized Beta (β), 95%CI	n/a	-.05 [-.13 to .03]	No difference
50	Team Sport: Non-Participants (ref) vs. Discontinuers from age 12-20	Depressive Symptoms age 20	Odds Ratio (OR), 95%CI	n/a	.73 [-.66 to 3.53]	No effect
	Team Sport: Non-Participants (ref) vs. Sustainers from age 12-20	Depressive Symptoms age 20	Odds Ratio (OR), 95%CI	n/a	.41 [-1.30 to 2.13]	No effect

343 **Table 6**
 344 *Effects sizes in moderate- and high-quality anxiety studies*

ID	Predictor	Outcome	Measure of Effect	Results (unadjusted)	Results (adjusted)	Comments
87	Weekly PA without coach: Never (ref) vs. 1-3 times age 10 (boys)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.78 [.54 to 1.14]	No difference
	Weekly PA without coach: Never (ref) vs. 4+ times age 10 (boys)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.98 [.69 to 1.38]	No difference
	Weekly PA with coach: Never (ref) vs. 1-3 times age 10 (boys)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.85 [.68 to 1.08]	No difference
	Weekly PA with coach: Never (ref) vs. 4+ times age 10 (boys)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.73 [.55 to .98]	27% lower hazard of receiving a diagnosis of an internalizing disorder
	Weekly PA without coach: Never (ref) vs. 1-3 times age 10 (girls)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.80 [.61 to 1.06]	No difference
	Weekly PA without coach: Never (ref) vs. 4+ times age 10 (girls)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.77 [.59 to 1.02]	No difference
	Weekly PA with coach: Never (ref) vs. 1-3 times age 10 (girls)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	1.03 [.86 to 1.23]	No difference
	Weekly PA with coach: Never (ref) vs. 4+ times age 10 (girls)	(Time until) diagnosis of internalizing disorder or age 18	Hazard Ratio (HR), 95%CI	n/a	.92 [.71 to 1.18]	No difference
8	PA volume: per 100 count increase in mean cpm age 13	Emotional Symptoms age 16	Standardized Beta (β), 95%CI	-.12 [-.24 to -.01]	-.11 [-.23 to -.00]	100count increase in mean daily cpm associated with small decrease in emotional symptoms
	PA intensity: per 60min increase in daily MVPA age 13	Emotional Symptoms age 16	Standardized Beta (β), 95%CI	-.54 [-1.08 to -.01]	-.49 [-1.02 to .04]	No effect in fully adjusted model

3	Per 30min increase in daily light PA (boys) age 7	Emotional Symptoms age 11	Standardized Beta (β), 95%CI	n/a	-.039 [-.106 to .028]	No effect
	Per 30min increase in daily light PA (girls) age 7	Emotional Symptoms age 11	Standardized Beta (β), 95%CI	n/a	-.026 [-.094 to .041]	No effect
	Per 15min increase in daily MVPA (boys) age 7	Emotional Symptoms age 11	Standardized Beta (β), 95%CI	n/a	-.048 [-.111 to .015]	No effect
	Per 15min increase in daily MVPA (girls) age 7	Emotional Symptoms age 11	Standardized Beta (β), 95%CI	n/a	-.049 [-.115 to .016]	No effect
91	Team Sport: Replacing 60mins of television at age 14	Emotional symptoms age 17	Standardized Beta (β), 95%CI	n/a	-.166 [-.300 to -.033]	Small reduction in emotional symptoms by replacing 60mins of TV with team sport
	Team Sport: Replacing 60mins of video games at age 14	Emotional symptoms age 17	Standardized Beta (β), 95%CI	n/a	-.040 [-.178 to .098]	No effect
	Team Sport: Replacing 60mins of social media at age 14	Emotional symptoms age 17	Standardized Beta (β), 95%CI	n/a	-.145 [-.282 to -.008]	Small reduction in emotional symptoms by replacing 60mins of social media with team sport
	Team Sport: Replacing 60mins of computer at age 14	Emotional symptoms age 17	Standardized Beta (β), 95%CI	n/a	-.069 [-.215 to .077]	No effect
	Individual Sport: Replacing 60mins of television at age 14	Emotional symptoms age 17	Standardized Beta (β), 95%CI	n/a	-.041 [-.092 to .009]	No effect
	Individual Sport: Replacing 60mins of video games at age 14	Emotional symptoms age 17	Standardized Beta (β), 95%CI	n/a	.016 [-.020 to .051]	No effect
	Individual Sport: Replacing 60mins of social media at age 14	Emotional symptoms age 17	Standardized Beta (β), 95%CI	n/a	-.019 [-.049 to .011]	No effect
	Individual Sport: Replacing 60mins of computer at age 14	Emotional symptoms age 17	Standardized Beta (β), 95%CI	n/a	.009 [-.028 to .047]	No effect
99	All Sport: Low (ref) vs. High trajectory age 6-10	Anxiety symptoms age 10	Standardized Beta (β), 95%CI	n/a	-.04 [-.12 to .05]	No difference

	High participation in Team (ref) vs. Individual sport age 6-10	Anxiety symptoms age 10	Standardized Beta (β), 95%CI	n/a	.05 [-.05 to .15]	No difference
	Team Sport: Low (ref) vs High trajectory from age 6-10	Anxiety symptoms age 10	Standardized Beta (β), 95%CI	n/a	-.09 [-.18 to -.01]	High involvement in team sports had small negative effect on anxiety symptoms
	Individual Sport: Low (ref) vs. High trajectory from age 6-10	Anxiety symptoms age 10	Standardized Beta (β), 95%CI	n/a	-.03 [-.11 to .05]	No difference
50	Team Sport: Non-Participators (ref) vs. Discontinuers from ages 12-20	Anxiety symptoms age 20 (panic disorder)	Odds Ratio (OR), 95%CI	n/a	.80 [.54 to 1.19]	No difference
	Team Sport: Non-Participators (ref) vs. Discontinuers from ages 12-20	Anxiety symptoms age 20 (generalized anxiety)	Odds Ratio (OR), 95%CI	n/a	.78 [.52 to 1.17]	No difference
	Team Sport: Non-Participators (ref) vs. Discontinuers from ages 12-20	Anxiety symptoms age 20 (social anxiety)	Odds Ratio (OR), 95%CI	n/a	.90 [.61 to 1.31]	No difference
	Team Sport: Non-Participators (ref) vs. Discontinuers from ages 12-20	Anxiety symptoms age 20 (agoraphobia)	Odds Ratio (OR), 95%CI	n/a	.63 [.40 to .98]	Discontinuers had 37% lesser odds of experiencing symptoms of agoraphobia
	Team Sport: Non-Participators (ref) vs. Sustainers from ages 12-20	Anxiety symptoms age 20 (panic disorder)	Odds Ratio (OR), 95%CI	n/a	.57 [.34 to .94]	Sustainers had 43% lesser odds of experiencing panic disorder symptoms
	Team Sport: Non-Participators (ref) vs. Sustainers from ages 12-20	Anxiety symptoms age 20 (generalized anxiety)	Odds Ratio (OR), 95%CI	n/a	.65 [.39 to 1.08]	No difference
	Team Sport: Non-Participators (ref) vs. Sustainers from ages 12-20	Anxiety symptoms age 20 (social anxiety)	Odds Ratio (OR), 95%CI	n/a	1.05 [.66 to 1.67]	No difference
	Team Sport: Non-Participators (ref) vs. Sustainers from ages 12-20	Anxiety symptoms age 20 (agoraphobia)	Odds Ratio (OR), 95%CI	n/a	.77 [.43 to 1.36]	No difference

346 **Table 7**
 347 *Effects sizes in moderate- and high-quality wellbeing studies*

ID	Predictor	Outcome	Measure of Effect	Results (unadjusted)	Results (adjusted)	Comments
54	Consistently Not Meeting (ref) vs. Meeting WHO PA recommendations one year in 14 to 18-year-olds	Social Health after two years (16-20)	Unstandardized Beta (B), (SE), <i>p</i>	n/a	5.8 (1.3) <i>p</i> < .0001	Meeting guidelines at one timepoint associated with "clinically meaningful" improvement in social health
	Consistently Not Meeting (ref) vs. Meeting WHO PA recommendations both years in 14 to 18-year-olds	Social Health after two years (16-20)	Unstandardized Beta (B), (SE), <i>p</i>	n/a	9.3 (1.2) <i>p</i> < .0001	Consistently meeting guidelines associated with a moderate, "clinically meaningful" improvement in social health
8	PA volume: per 100 count increase in mean counts per minute age 13	Wellbeing age 16	Standardized Beta (β), 95%CI	-.06 [-.47 to .35]	-.02 [-.44 to .40]	No effect in fully adjusted model
	PA intensity: per 60min increase in daily MVPA age 13	Wellbeing age 16	Standardized Beta (β), 95%CI	-.54 [-2.49 to 1.41]	.03 [-1.98 to 2.05]	No effect in fully adjusted model
	PA volume: per 100 count increase in mean counts per minute age 13	Peer Relations age 16	Standardized Beta (β), 95%CI	-.02 [-.09 to .06]	-.02 [-.09 to .05]	No effect
	PA intensity: per 60min increase in daily MVPA age 13	Peer Relations age 16	Standardized Beta (β), 95%CI	-.15 [-.49 to .20]	-.16 [-.51 to .19]	No effect
61	Times engaging in PA in past week in 12 to 17-year-olds (males)	Self-Esteem after one year (13-18)	Unstandardized Beta (B), SE, <i>p</i>	n/a	.063 (.014) <i>p</i> < .01	Each additional bout of weekly PA had a small positive effect for self-esteem one year later
	Times engaging in PA in past week in 12 to 17-year-olds (females)	Self-Esteem after one year (13-18)	Unstandardized Beta (B), SE, <i>p</i>	n/a	.062 (.015) <i>p</i> < .01	Each additional bout of weekly PA had a small positive effect for self-esteem one year later
3	Per 30min increase in daily light PA (boys) age 7	Peer Relations age 11	Standardized Beta (β), 95%CI	n/a	-.086 [-.150 to -.021]	Small reduction in peer relationship problems
	Per 30min increase in daily light PA (girls) age 7	Peer Relations age 11	Standardized Beta (β), 95%CI	n/a	-.071 [-.130 to -.013]	Small reduction in peer relationship problems
	Per 15min increase in daily MVPA (boys) age 7	Peer Relations age 11	Standardized Beta (β), 95%CI	n/a	-.077 [-.133 to -.022]	Small reduction in peer relationship problems
	Per 15min increase in daily MVPA (girls) age 7	Peer Relations age 11	Standardized Beta (β), 95%CI	n/a	-.007 [-.057 to .071]	No effect

349 Fig 3 clusters studies by mental health outcome reported, length of follow-up (i.e., from first
350 PA measurement to last mental health measurement) and whether they reported a statistically
351 significant beneficial effect. Across nearly all timeframes and all outcomes, a greater proportion of
352 studies reported a beneficial effect than did not, providing partial evidence that child/adolescent PA
353 may predict anxiety, depression and wellbeing at least 12 months later and as far as 55 years across
354 the lifespan⁽⁹⁰⁾.

355 *** **Figure 3:** Length of studies in years grouped by outcome ***

356 A total of 41/67 depression studies (61.2%) reported significant negative associations (Table
357 2). Ten of these studies were moderate- or high-quality^(39, 41, 50, 53, 61, 63, 87, 90, 96, 99) and observed 1/42
358 moderate, 17/42 small, and 24/42 non-significant effects (Table 5). In comparison, 19/35 studies
359 (54.3%) reported significant negative associations with anxiety (Table 3). Of these, 6 were moderate-
360 or high-quality^(3, 8, 50, 87, 91, 99) and reported 1/34 moderate, 6/34 small and 27/34 non-significant effects
361 (Table 6). Finally, 25/39 studies (64.1%) investigating indicators of wellbeing reported significant
362 positive associations (Table 4). Of these, 4 were moderate- or high-quality^(3, 8, 54, 61) and presented
363 evidence of 2/12 moderate, 5/12 small and 5/12 non-significant effects (Table 7). A key is provided
364 next to the study ID in Table 4 to indicate which indicator of wellbeing/flourishing was investigated.
365 Further details on the 88 associations explored in moderate- or high-quality studies ($n = 14$) can be
366 found in Tables 5, 6, 7.

367 **Individual vs. group-based activities**

368 With just one exception⁽⁵³⁾, all studies researching the effects of group-based or individual
369 activities investigated sport participation. Team sports were found to be beneficial in 7/13 depression
370 studies (53.8%), 8/9 anxiety studies (88.9%) and 1/2 wellbeing studies (50%). Individual sports were
371 beneficial in 1/8 depression studies (12.5%), 4/6 anxiety studies (66.7%) and 0/1 wellbeing studies
372 (0%). Effect sizes were consistently reported as small on all but two occasions^(25, 50) wherein,
373 compared to no participation, involvement in team sport throughout adolescence⁽²⁵⁾ with sustained

374 participation in early adulthood⁽⁵⁰⁾ was associated with 24% lower odds of depressive disorder
375 diagnosis⁽²⁵⁾, 30% lower odds of anxiety disorder diagnosis⁽²⁵⁾, and 43% lower odds of panic disorder
376 symptoms⁽⁵⁰⁾ in early adulthood. Moderate/high-quality studies reported that team but not individual
377 sport participation in childhood/adolescence had small protective effects against prospective
378 depressive and anxiety symptoms up to 3 years later^(50, 63, 91, 99). For further information on study
379 lengths see Fig 4. For further information on all studies targeting either individual or group-based
380 activities see Supporting Table S3.

381 *** **Figure 4:** Length of studies in years grouped by type of activity ***

382 **Domains of activity**

383 No studies focused on the effects of travel or household activities. In total, leisure-time
384 activities were found to be beneficial in 7/12 depression studies (58.3%), 3/6 anxiety studies (50%)
385 and 5/7 wellbeing studies (71.4%). School-based activities were beneficial in 5/8 depression studies
386 (62.5%), 2/5 anxiety studies (40%) and 2/5 wellbeing studies (40%). Three studies reported moderate
387 effects for depression, anxiety and wellbeing when leisure activity participation was high in
388 adolescence and maintained into adulthood^(47, 50, 58). Specifically, 65% reduced relative risk of
389 depression⁽⁴⁷⁾, 47% reduced odds of panic disorder symptoms⁽⁵⁰⁾, and twice the likelihood high life
390 satisfaction⁽⁵⁸⁾ in early adulthood. Moderate and high-quality studies presented evidence that small
391 beneficial effects of school sports (but not PE)⁽⁵³⁾ and leisure activities for anxiety and depressive
392 symptoms were still observable after 3 years⁽⁵⁰⁾ and up to 55 years⁽⁹⁰⁾ in the future. For further
393 information on study lengths see Fig 5. For further information on the effects of PA performed in
394 leisure- or school-time see Supporting Table S4.

395 *** **Figure 5:** Length of studies in years grouped by domain of activity ***

396 **WHO MVPA guideline adherence**

397 In total, 12 studies used data referring to the mean daily amount of MVPA performed across
398 the past 7 days (hence facilitating categorisation) or directly compared the effects of guideline
399 adherence versus non-adherence. Positive effects of guideline adherence were reported in 2/4
400 depression studies (50%), 1/3 anxiety studies (33.3%) and 4/5 wellbeing studies (80%). Findings
401 were mixed across studies comparing adherence and non-adherence with 1/3 (33.3%) finding
402 additional benefits of achieving meeting guidelines for future depressive symptoms, 0/2 (0%) for
403 anxiety symptoms and 2/2 (100%) for wellbeing-related outcomes. All significant effect sizes
404 reported were small. One study was high quality⁽⁵⁴⁾, presenting evidence of a dose response
405 relationship between years of WHO MVPA guideline adherence in adolescence and social quality of
406 life measured one year later. For further information on study lengths see Fig 6. For further evidence
407 of the prospective effects of WHO MVPA guideline (non)adherence, see Supporting Table S5.

408 *****Figure 6:** Length of studies in years grouped by WHO MVPA guideline adherence ***

409 **Discussion**

410 The present study was the first comprehensive systematic review to assess the association
411 between PA performed in childhood/adolescence for future anxiety, depression and wellbeing using
412 a longitudinal one-year time interval. Evidence of associations with anxiety, depression and wellbeing
413 was deemed to be low quality due to heavy reliance on observational study designs. Moderate and
414 high-quality studies reported small or non-significant effects of child/adolescent PA persisted for at
415 least 12 months suggesting there may be some minor long term mental health benefits that emerge
416 from participation in certain activities in youth. Directing PA promotion towards youth as a method
417 of reducing long-term risk of mental illness or suboptimal wellbeing likely requires careful
418 consideration of important contextual factors in order to be successful. For example, the longest
419 running study in the review was high quality and investigated associations in 18,555 participants over

420 55 years⁽⁹⁰⁾. PA performed age 16 but not age 7 or 11 had protective effects against depressive
421 symptoms that grew in magnitude as participants progressed through mid-life and into old age. This
422 study supports existing literature suggesting that the relationship between PA and prospective mental
423 health outcomes may not be uniformly linear [23]. Inherent complexity in the relationship may be
424 somewhat explained through consideration of type, domain and volume of the activity in question.

425 **Individual vs. group-based activities**

426 Psychosocial influences can play a key role in the relationship between PA and mental health.
427 In PA research and in particular sports studies, psychosocial factors are commonly investigated by
428 drawing comparisons between individual and group-based activities. Currently, the general
429 hypothesis is that group-based or team-sport environments are more conducive to the development
430 of interpersonal connections and social support networks than individual ones [36],[37]. The current
431 review supports this stance with evidence that team sports are more consistently beneficial for long-
432 term mental health outcomes than individual sports [60]. It is possible that group-based environments
433 offer more opportunities for positive interpersonal interactions than individual environments, thus
434 fulfilling an individual's need for relatedness and by extension, social wellbeing [61]. In addition, the
435 readily available and naturally forming social support networks, social identification, and sense of
436 belonging that can emerge through team membership may play a role in the reduction of depressive
437 and anxiety symptoms [62].

438 Included studies investigated a number of psychosociological factors that influence the
439 association between sport participation and mental health. Two studies controlled for mediating
440 effects of sport self-concept and perceived sporting competence^(70, 83) reporting that links with future
441 self-esteem were fully dependent on positive sport-related self-beliefs which aligns with existing
442 literature [62]. These studies, among others, present important points for PA promotion. Long-term
443 mental health benefits of sport may not be felt equally for all individuals [60]. For many, PA, whether
444 accumulated through physical education, school sport, or otherwise, can be perceived as a chore and

445 even unpleasant [63],[64] in which case, positive mental health outcomes are likely to be limited.
446 Activities that can take place outside of competitive environments or those that do not lend themselves
447 as easily to comparison with others may be more suitable for individuals with negative self-
448 perceptions, at least in the first instance.

449 There are also some risks with sport involvement that were not addressed in the review such
450 as the young person's goal-orientation or athletic identity. Goal Achievement Theory posits that ego-
451 orientated goals can negatively impact wellbeing when playing poorly [65]. Likewise, strong athletic
452 identities have been linked with anxiety during non-normative transitions [66]. It is important to be
453 mindful of these potentially harmful aspects of participation throughout PA promotion and
454 intervention design.

455 **Domains of activity**

456 Studies in the current review assessed only the effects of leisure or school-based activities
457 leaving the prospective effects of active transport and household activities in young people to be
458 determined. Whether there are differential effects associated with leisure or school-based activities in
459 childhood/adolescence for prospective mental health outcomes remains unclear. One high-quality
460 study⁽⁹⁰⁾ found no association between leisure activities performed age 7 or 11 while those performed
461 at age 16 were associated with a 27% reduced risk of depressive symptoms between the ages of 23-
462 62. This could infer the existence of a critical period during which being active yields greater long
463 term mental health effects. Alternatively, such an outcome could be a product of the significant drop
464 off in activity levels that typically coincides the transition from childhood to adolescence [67]. If
465 activity levels reported age 7 or 11 did not track well over time they are less likely to predict future
466 mental health [27].

467 Two moderate quality studies tested associations with school-based activities^(39, 53). One
468 reported that participating in school sports in adolescence had a small negative association with
469 depressive symptoms in early adulthood⁽³⁹⁾ whilst the other reported that specialist taught PE classes

470 at age 8 had no effect on depressive symptoms after 4 years and was even associated with less decline
471 in symptoms compared to controls⁽⁵³⁾. Again, inconsistent effects reported by these studies could be
472 linked to poor tracking of PA levels across the transition from childhood to adolescence. They may
473 also illuminate the importance of intrinsic motivation to participate which was more likely a
474 contributing factor for those playing extracurricular sports than in compulsory PE classes [61].

475 Many explanatory mechanisms for domain-specific effects are rooted in self-determination
476 theory [61]. Research in adults would suggest that household activities are seldom challenging
477 enough to experience a sense of task-mastery, nor are they often fuelled by intrinsic motivation, and
478 hence may not benefit mental health [38]. Indeed, in certain circumstances, these activities might
479 even cause stress. In contrast, it is plausible that consistently connecting with the natural environment
480 by walking to school with friends coupled with intrinsic motivation to be active could simultaneously
481 satisfy one's need to connect with nature and each other [68],[69]. Ease of access to green space is
482 cited alongside increased surveillance of life domains in the WHO's Global Action Plan for PA [20].
483 For these reasons, more research on domain-specific PA, particularly active transport is required.

484 In adults, leisure and transport activities are more likely to be autonomy-supportive and
485 provide opportunities to experience a sense of belonging than occupational or household activities
486 [70]. Whilst adult-based literature can be used to inform hypotheses in children/adolescents,
487 researchers should be careful not to generalise findings across these groups. Domain-specific effects
488 on wellbeing can vary across age ranges with evidence that active transport is beneficial for adults
489 but not adolescents [38]. Benefits may materialise for adults through autonomous endorsement of
490 transport PA for enjoyment or health reasons compared to young people who may have to use active
491 means of transport due to lack of an alternative [38]. Improving understanding of domains could help
492 individuals working in public health develop more strategic interventions best suited to promote PA
493 and mental health in different age groups and communities.

494 **WHO MVPA guideline adherence**

495 A moderate-quality study using the National Longitudinal Study of Adolescent Health (Add
496 Health) data found that adherence vs. non-adherence to WHO MVPA guidelines in adolescence was
497 associated with a 16% reduction in the incidence of depressive symptoms when aged 33 to 39⁽⁹⁶⁾.
498 Trajectory analysis from the same study found evidence of an added benefit for those who
499 continuously met guidelines (i.e., throughout adolescence and adulthood) compared those who started
500 to meet guidelines later in life. Whilst findings do not dispute the commonly held belief that it is
501 never too late to start being active, they do infer that the younger we start, the better.

502 Guidelines dictate that across the week, children and adolescents should engage in an average
503 of 60 minutes MVPA per day to experience health benefits. It is important to note that current MVPA
504 guidelines are not specific to mental health and also aim to support physical health through
505 cardiorespiratory fitness [39]. As such, focusing on the effects of 60 minutes of MVPA per day could
506 mean that benefits associated with lesser amounts of PA are overlooked. Whilst in the current review
507 beneficial effects emerged from guideline adherence, there were several incidences where benefits
508 were achieved through lesser amounts^(30, 51) or where there were no differences between those who
509 met and did not meet guidelines^(55, 85). The current review presents partial evidence that WHO MVPA
510 guideline adherence may be beneficial for long-term mental health. However, given the inconsistency
511 of findings, perhaps a more realistic or achievable goal would be for young people to simply ‘get
512 moving’. This approach aligns with a growing body of evidence that any activity is better than no
513 activity [71],[72].

514 Research should continue to explore the benefits of WHO MVPA guideline adherence but
515 broaden its focus to also consider the long-term effects of lesser volumes. Some low-quality studies
516 reported that light-intensity PA performed in childhood/adolescence was associated with better social
517 wellbeing⁽³⁾ and fewer prospective depressive/anxiety symptoms^(8, 41, 42). Increasing light-intensity
518 activity or introducing small bouts of MVPA through programmes such as ‘Snacktivity’ [73] aligns

519 with the small-change approach to habit-formation [74] and may present a more achievable goal for
520 particularly inactive young people. Additionally, PA guidelines for children have only recently begun
521 to incorporate muscle-strengthening activities (e.g., bodyweight exercises, resistance training) into
522 recommendations. In our review, few studies measured such behaviours in young people^(16, 26). Given
523 emerging research in adults showing the benefits of muscle strengthening activities for mental health
524 beyond those associated with aerobic PA [75],[76], closer inspection of their association in
525 children/adolescents' mental health is warranted.

526 **Heterogeneity & other observations**

527 Heterogeneity in the items used to measure mental health limited the comparability of study
528 findings. For example, aside from two studies using the SMFQ to measure depression^(29, 93), every
529 study investigating the effects of individual or group-based activities used a different measure of
530 mental health. In total, 23 validated measures of depression, 19 validated measures of anxiety, and
531 15 validated measures of wellbeing were used across the 98 studies alongside a plethora of non-
532 validated items. Some scales have been validated for use in the general population however, to the
533 best of our knowledge are yet to undergo formal validation in child/adolescent samples. In attempting
534 to advance the field, psychometric validation studies should assess a range of metrics specifically
535 designed for use in young people for concurrent/divergent validity, test-retest reliability, factor
536 structure, and item comprehension. Such examples exist for health-related quality of life measures in
537 children (e.g., KIDSCREEN-27)[77] but few have undergone the same rigorous assessments for
538 mental illness.

539 Conversely, depression and anxiety are both unitary constructs with widely accepted
540 definitions, so there is a greater likelihood that different scales converge on the same latent construct
541 than for an ill-defined concept such as wellbeing. Two high quality studies with similar aged
542 participants reported inconsistent effects between comparable levels of daily MVPA and social
543 wellbeing^(8, 54). These studies are said to be measuring two broadly similar constructs in social quality

544 of life and peer relationships, but the lack of conceptual clarity makes it difficult to determine if they
545 were comparing like for like. Concurrent validity studies have attempted to address this issue in the
546 past [78] but up-to-date comparisons of wellbeing measures are scant. Rigorous validation studies for
547 wellbeing measures in younger populations are necessary to improve study comparability going
548 forward.

549 Historical shifts in research focus were observed. Studies using datasets commencing prior to
550 the year 2000 displayed a clear tendency to investigate depression over anxiety and indicators of
551 wellbeing. It could be argued that to reduce the subsequent disbalance in the literature, wellbeing
552 and/or anxiety should be the focus of research going forward. However, technological advancements
553 and ‘The Physical Activity Transition’ [79] toward more sedentary-type jobs/pastimes have given
554 rise to a generational shift in PA patterns. This shift means the relationship between PA and mental
555 health has evolved over time [63] making older data less representative of modern-day associations.
556 PA participation now offers a way to reduce total screentime; an increasingly prevalent lifestyle factor
557 independently linked to depression, anxiety, and self-esteem in young people that did not exist prior
558 to the invention of smart phones [80]. Research must continue to investigate the extent to which PA
559 can influence both mental illness and wellbeing to maintain up-to-date understanding of these
560 perpetually fluid relationships.

561 **Limitations of included studies**

562 Due to heavy reliance on longitudinal cohort studies, most evidence was low quality. The
563 resources required to monitor the effects of PA at one timepoint for mental health outcomes at least
564 one year later are considerable, meaning the reliance on longitudinal cohort studies and other
565 observational designs is likely to continue in the future. Whilst there is a need for more randomized
566 trials, a more pertinent recommendation would be for those overseeing the management of
567 longitudinal cohort studies to come to a consensus on a theoretically informed understanding of
568 mental health and the measurement tools they use to improve physical activity and mental health

569 epistemology in young people. Such a move would reduce some limitations associated with
570 secondary analysis and enhance study comparability whilst accommodating newly emerging high-
571 level forms of statistical analyses such as retrospective data harmonization [81].

572 **Limitations of the review design**

573 Several studies used the same data-source to answer similar research questions, posing the
574 risk that associations (or lack thereof) from a single data-source could skew findings [82]. A
575 systematic process was followed [49] to compare studies using the same data-source on a number of
576 study modifiers (see supporting files). Except for two studies^(74, 86), those included in data synthesis
577 were all deemed to provide a unique insight into the association between child/adolescent PA and
578 outcomes related to anxiety, depression or wellbeing at least 12 months later (Supporting Table S1).
579 Further, best efforts have been made to ensure the source of observed associations is clear with SoF
580 tables noting the dataset used in each article.

581 Aligning with Keyes' two-continua model meant that whilst the review was structured within
582 a sound theoretical framework, a small number of studies had to be excluded as they analysed
583 psychometric scales that combined aspects of mental illness and wellbeing into an overall mental
584 health score. It is the authors' belief that the increased heterogeneity that would have arisen through
585 inclusion of these additional measures would have obfuscated trends further and that the benefits of
586 using Keyes' model as a conceptual framework outweighed its drawbacks, strengthening the quality
587 of the review as a whole.

588 **Conclusions**

589 PA promotion in young people may represent a viable means of achieving a small reduction
590 in the overall burden of CMDs and improving indicators of wellbeing in some individuals however,
591 research is disjointed. The current review presents low quality partial evidence for PA participation
592 in childhood/adolescence having small benefits for mental health outcomes related to anxiety,
593 depression and wellbeing at least 12 months later. Team sport participation is among the most

594 consistently predictive type of PA. Long-term effects of domain-specific activities require further
595 research with particular focus needed on the effects of active transport and household activities.
596 Adhering to WHO MVPA guidelines in youth may be associated with small improvements in long-
597 term mental health however, the benefits of lesser volumes and intensities of activity warrant further
598 discussion. A vast range mental health measures used across studies led to heterogeneous operational
599 definitions and conceptual ambiguity. Researchers must strive to achieve consensus on their
600 theoretical approach and instruments used to measure depression, anxiety, and wellbeing in young
601 people to overcome epistemological concerns and improve study comparability going forward.

602 **Acknowledgements**

603 We would like to thank all authors who provided additional information used for synthesis.

604 **References**

- 605 1. Schuch FB, Vancampfort D, Firth J, Rosenbaum S, Ward PB, Silva ES, et al. Physical activity
606 and incident depression: a meta-analysis of prospective cohort studies. *Am J Psychiatry*.
607 2018;175(7):631–48. doi: 10.1176/appi.ajp.2018.17111194
- 608 2. McDowell CP, Dishman RK, Gordon BR & Herring MP. Physical activity and anxiety: a
609 systematic review and meta-analysis of prospective cohort studies. *Am J Prev Med*.
610 2019;57(4):545–556. doi: 10.1016/j.amepre.2019.05.012
- 611 3. Wiese CW, Kuykendall, L & Tay L. Get active? A meta-analysis of leisure-time physical
612 activity and subjective well-being. *J Posit Psychol*. 2018;13(1):57–66.
613 doi: 10.1080/17439760.2017.1374436
- 614 4. Stubbs B, Vancampfort D, Hallgren M, Firth J, Veronese N, Solmi M et al. EPA guidance on
615 physical activity as a treatment for severe mental illness: a meta-review of the evidence and
616 Position Statement from the European Psychiatric Association (EPA), supported by the
617 International Organization of Physical Therapists in Mental Health (IOPTMH). *Eur Psychiatry*.
618 2018;54:124-144. doi: 10.1016/j.eurpsy.2018.07.004
- 619 5. 2018 Physical Activity Guidelines Advisory Committee Scientific Report. Available from:
620 [https://health.gov/our-work/nutrition-physical-activity/physical-activity-guidelines/current-](https://health.gov/our-work/nutrition-physical-activity/physical-activity-guidelines/current-guidelines/scientific-report)
621 [guidelines/scientific-report](https://health.gov/our-work/nutrition-physical-activity/physical-activity-guidelines/current-guidelines/scientific-report); 2018. Accessed: 19/06/2023
- 622 6. Brown WJ, Bauman AE, Bull F & Burton NW (2013). Development of Evidence-based
623 Physical Activity Recommendations for Adults (18-64 years). Report prepared for the
624 Australian Government Department of Health. 2012.
- 625 7. Okely AD, Ghersi D, Loughran SP, Cliff D, Shilton T & Jones R. (2019) Australian 24-hour
626 movement guidelines for children (5-12 years) and young people (13-17 years): an integration
627 of physical activity, sedentary behaviour, and sleep – research report: Australian Government,
628 The Department of Health; 2019. Available from: <https://www.health.gov.au/topics/physical->

- 629 activity-and-exercise/physical-activity-and-exercise-guidelines-for-all-australians/for-children-
630 and-young-people-5-to-17-years?language=und. Accessed: 19/06/2023
- 631 8. Keyes CL. Mental illness and/or mental health? Investigating axioms of the complete state
632 model of health. *J Consult Clin Psychol.* 2005;73(3):539-548.
- 633 9. Westerhof GJ & Keyes CL. Mental illness and mental health: The two continua model across
634 the lifespan. *J Adult Dev.* 2010;17(2):110-119. doi: 10.1007/s10804-009-9082-y
- 635 10. Keyes CL. The mental health continuum: From languishing to flourishing in life. *J Health Soc*
636 *Behav.* 2002;43(2):207-222. doi: 10.2307/3090197
- 637 11. Keyes CL. Promoting and protecting mental health as flourishing: a complementary strategy for
638 improving national mental health. *Am Psychol.* 2007;62(2):95-108.
639 doi: 10.1037/0003-066X.62.2.95
- 640 12. Schotanus-Dijkstra M, Pieterse ME, Drossaert CH, Westerhof GJ, De Graaf R, ten Have M et
641 al. What factors are associated with flourishing? Results from a large representative national
642 sample. *J Happiness Stud.* 2016;17(4):1351-1370. doi: 10.1007/s10902-015-9647-3
- 643 13. Global Burden of Diseases Study. 2019. Available from: [https://ghdx.healthdata.org/gbd-](https://ghdx.healthdata.org/gbd-results-tool?params=gbd-api-2019-permalink/d780dffbe8a381b25e1416884959e88b)
644 [results-tool?params=gbd-api-2019-permalink/d780dffbe8a381b25e1416884959e88b](https://ghdx.healthdata.org/gbd-results-tool?params=gbd-api-2019-permalink/d780dffbe8a381b25e1416884959e88b).
645 Accessed: 19/06/2023
- 646 14. Steel Z, Marnane C, Iranpour C, Chey T, Jackson JW, Patel V et al. The global prevalence of
647 common mental disorders: a systematic review and meta-analysis 1980–2013. *Int J Epidemiol.*
648 2014;43(2):476-493. doi: 10.1093/ije/dyu038
- 649 15. Chisholm D, Sweeny K, Sheehan P, Rasmussen B, Smit F, Cuijpers P et al. Scaling-up
650 treatment of depression and anxiety: a global return on investment analysis. *Lancet Psychiatry.*
651 2016;3(5):415-424. doi: 10.1016/S2215-0366(16)30024-4
- 652 16. Katzmarzyk PT, Friedenreich C, Shiroma EJ & Lee IM. Physical inactivity and non-
653 communicable disease burden in low-income, middle-income and high-income countries. *Br J*
654 *Sports Med.* 2022;56(2):101–106. doi: 10.1136/bjsports-2020-103640

- 655 17. Simons G & Baldwin DS. A critical review of the definition of ‘wellbeing’ for doctors and their
656 patients in a post Covid-19 era. *Int J Soc Psychiatry*. 2021;67(8):984-991.
657 doi: 10.1177/00207640211032259
- 658 18. Huppert FA & So TT. Flourishing across Europe: Application of a new conceptual framework
659 for defining wellbeing. *Soc Indic Res*. 2013;110(3):837-861. doi: 10.1007/s11205-011-9966-7
- 660 19. Diener E & Ryan K. Subjective well-being: A general overview. *S Afr J Psychol*.
661 2009;39(4):391-406. doi: 10.1177/008124630903900402
- 662 20. World Health Organisation (WHO). *Global Action Plan on Physical Activity 2018-2030: More
663 Active People for a Healthier World*. Geneva: WHO; 2019.
- 664 21. Guthold R, Stevens GA, Riley LM & Bull FC. Worldwide trends in insufficient physical
665 activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1.9 million
666 participants. *Lancet Glob Health*. 2018;6(10):e1077–e1086. doi: 10.1016/S2214-
667 109X(18)30357-7
- 668 22. Guthold R, Stevens GA, Riley LM & Bull FC. Global trends in insufficient physical activity
669 among adolescents: a pooled analysis of 298 population-based surveys with 1.6 million
670 participants. *Lancet Child Adolesc Health*. 2020;4(1):23–35. doi: 10.1016/S2352-
671 4642(19)30323-2
- 672 23. Biddle SJ, Ciaccioni S, Thomas G & Vergeer I. Physical activity and mental health in children
673 and adolescents: An updated review of reviews and an analysis of causality. *Psychol Sport
674 Exerc*. 2019;42:146-155. doi: 10.1016/j.psychsport.2018.08.011
- 675 24. Dale LP, Vanderloo L, Moore S & Faulkner G. Physical activity and depression, anxiety, and
676 self-esteem in children and youth: An umbrella systematic review. *Ment Health Phys Act*.
677 2019;16:66–79. doi: 10.1016/j.mhpa.2018.12.001
- 678 25. Rodriguez-Ayllon M, Cadenas-Sánchez C, Estévez-López F, Muñoz NE, Mora-Gonzalez J,
679 Migueles JH et al. Role of physical activity and sedentary behavior in the mental health of

- 680 preschoolers, children and adolescents: a systematic review and meta-analysis. *Sports Med.*
681 2019;49(9):1383-1410. doi: 10.1007/s40279-019-01099-5
- 682 26. Kristensen PL, Møller N, Korsholm L, Wedderkopp N, Andersen LB & Froberg K. Tracking of
683 objectively measured physical activity from childhood to adolescence: the European youth heart
684 study. *Scand J Med Sci Sports.* 2008;18(2):171–178. doi: 10.1111/j.1600-0838.2006.00622.x
- 685 27. Telama R, Yang X, Leskinen E, Kankaanpää A, Hirvensalo M, Tammelin T et al. Tracking of
686 physical activity from early childhood through youth into adulthood. *Med Sci Sports Exerc.*
687 2014;46(5):955–962. doi: 10.1249/mss.0000000000000181
- 688 28. Connolly S, Carlin A, Johnston A, Woods C, Powell C, Belton S et al. Physical Activity, Sport
689 and Physical Education in Northern Ireland School Children: A Cross-Sectional Study. *Int J*
690 *Environ Res Public Health.* 2020;17(18):6849. doi: 10.3390/ijerph17186849
- 691 29. Lawrence D, Hafekost J, Johnson SE, Saw S, Buckingham WJ, Sawyer MG et al. Key findings
692 from the second Australian child and adolescent survey of mental health and wellbeing. *Aust N*
693 *Z J Psychiatry.* 2016;50(9):876-886. doi: 10.1177/0004867415617836
- 694 30. Inchley J, Currie D, Budisavljevic S, Torsheim T, Jåstad A, Cosma A et al. Spotlight on
695 adolescent health and well-being: Findings from the 2017/2018 Health Behaviour in School-
696 Aged Children (HBSC) survey in Europe and Canada. International Report. 2020. 1. Available
697 from: <https://apps.who.int/iris/handle/10665/332104>. Accessed: 19/05/2023
- 698 31. Solmi M, Radua J, Olivola M, Croce E, Soardo L, Salazar de Pablo G. et al. Age at onset of
699 mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. *Mol*
700 *Psychiatry.* 2022;27(1):281–295. doi: 10.1038/s41380-021-01161-7
- 701 32. Berk L. *Child development* (9th Edition). Pearson Higher Education AU; 2015
- 702 33. Donaldson C, Moore G & Hawkins J. A systematic review of school transition interventions to
703 improve mental health and wellbeing outcomes in children and young people. *School Ment*
704 *Health.* 2023;15:19-35. doi: 10.1007/s12310-022-09539-w

- 705 34. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD et al. The PRISMA
706 2020 statement: an updated guideline for reporting systematic reviews. *Int J Surg.*
707 2021;88:105906. doi: 10.1016/j.ijsu.2021.105906
- 708 35. Caspersen CJ, Powell KE & Christenson GM. Physical activity, exercise, and physical fitness:
709 definitions and distinctions for health-related research. *Public Health Rep.* 1985;100(2):126.
710 PMID: 3920711
- 711 36. Eime RM, Young JA, Harvey JT, Charity MJ & Payne WR. A systematic review of the
712 psychological and social benefits of participation in sport for children and adolescents:
713 informing development of a conceptual model of health through sport. *Int J Behav Nutr Phys*
714 *Act.* 2013;10(1):1-21. doi: 10.1186/1479-5868-10-98
- 715 37. Evans MB, Allan V, Erickson K, Martin LJ, Budziszewski R & Côté J. Are all sport activities
716 equal? A systematic review of how youth psychosocial experiences vary across differing sport
717 activities. *Br J Sports Med.* 2017;51(3):169-176. doi: 10.1136/bjsports-2016-097326
- 718 38. White RL, Babic MJ, Parker PD, Lubans DR, Astell-Burt T & Lonsdale C. Domain-specific
719 physical activity and mental health: a meta-analysis. *Am J Prev Med.* 2017;52(5):653-666.
720 doi: 10.1016/j.amepre.2016.12.008
- 721 39. Teychenne M, White RL, Richards J, Schuch FB, Rosenbaum S & Bennie JA. Do we need
722 physical activity guidelines for mental health: What does the evidence tell us? *Ment Health Phys*
723 *Act.* 2020;18:100315. doi: 10.1016/j.mhpa.2019.100315
- 724 40. Sawyer SM, Azzopardi PS, Wickremarathne D & Patton GC. The age of adolescence. *Lancet*
725 *Child Adolesc Health.* 2018;2(3):223-228. doi: 10.1016/S2352-4642(18)30022-1
- 726 41. Ainsworth B, Cahalin L, Buman M & Ross R. The current state of physical activity assessment
727 tools. *Prog Cardiovasc Dis.* 2015;57(4):387-395. doi: 10.1016/j.pcad.2014.10.005
- 728 42. Gunnell KE, Belcourt VJ, Tomasone JR & Weeks LC. Systematic review methods. *Int Rev*
729 *Sport Exerc Psychol.* 2022;15(1):5-29. doi: 10.1080/1750984X.2021.1966823

- 730 43. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P et al. GRADE: an
731 emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*.
732 2008;336(7650):924-926. doi: 10.1136/bmj.39489.470347.AD
- 733 44. Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S et al. Synthesis
734 without meta-analysis (SWiM) in systematic reviews: Reporting guideline. *BMJ*.
735 2020;368:16890. doi: 10.1136/bmj.l6890
- 736 45. Henry S, Ospina MB, Dennett L & Hicks A. Assessing the Risk of Respiratory-Related
737 Healthcare Visits Associated with Wildfire Smoke Exposure in Children 0–18 Years Old: A
738 Systematic Review. *Int J Environ Res Public Health*. 2021;18(16):8799. doi:
739 10.3390/ijerph18168799
- 740 46. Cohen J. The effect size. *Statistical Power Analysis for the Behavioral Sciences*. 1988;77-83.
741 ISBN 0-8058-0283-5
- 742 47. Fey CF, Hu T & Delios A. The Measurement and communication of effect sizes in management
743 research. *Manag Organ Rev*. 2023;19(1):176-197. doi: 10.1017/mor.2022.2
- 744 48. Breslin G, Shannon S, Haughey T, Donnelly P & Leavey G. A systematic review of
745 interventions to increase awareness of mental health and well-being in athletes, coaches and
746 officials. *Syst Rev*. 2017;6(1):1-15. doi: 10.1186/s13643-017-0568-6
- 747 49. Dishman RK, McDowell CP, Herring MP. Customary physical activity and odds of depression:
748 a systematic review and meta-analysis of 111 prospective cohort studies. *Br J Sports Med*.
749 2021;55(16):926–34. doi: 10.1136/bjsports-2020-103140
- 750 50. Radloff LS. The CES-D scale: A self-report depression scale for research in the general
751 population. *Appl Psychol Meas*. 1977;1(3):385-401. doi: 10.1177/014662167700100306
- 752 51. Goodman R. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol*
753 *Psychiatry*. 1997;38(5):581-586. doi: 10.1111/j.1469-7610.1997.tb01545.x

- 754 52. Diener E, Wirtz D, Tov W, Kim-Prieto C, Choi D, Oishi S, et al. New well-being measures:
755 Short scales to assess flourishing and positive and negative feelings. *Soc Indic Res.*
756 2010;97(2):143–56. doi: 10.1007/s11205-009-9493-y
- 757 53. Bullinger M. Entwicklung und anwendung von instrumenten zur erfassung der lebensqualität.
758 *Lebensqualitätsforschung.* 1997;1-6.
- 759 54. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh
760 mental well-being scale (WEMWBS): development and UK validation. *Health Qual Life*
761 *Outcomes.* 2007;5(1):1–13. doi: 10.1186/1477-7525-5-63
- 762 55. Rosenberg M. Rosenberg self-esteem scale (RSE). *Acceptance and commitment therapy.*
763 *Measures Package.* 1965;61(52):18.
- 764 56. Harter S. Self-perception profile for children. *Hisp J Behav Sci.* 1985. doi: 10.1037/t05338-000
- 765 57. Diener ED, Emmons RA, Larsen RJ & Griffin S. The satisfaction with life scale. *J Pers Assess.*
766 1985;49(1):71-75. doi: 10.1207/s15327752jpa4901_13
- 767 58. Huebner ES. Initial development of the student's life satisfaction scale. *School Psychol Int.*
768 1991;12(3):231-240. doi: 10.1177/0143034391123010
- 769 59. Watson D, Clark LA & Tellegen A. Development and validation of brief measures of positive
770 and negative affect: The PANAS scales. *J Pers Soc Psychol.* 1988;54(6):1063–1070.
771 doi: 10.1037/0022-3514.54.6.1063
- 772 60. Lubans D, Richards J, Hillman C, Faulkner G, Beauchamp M, Nilsson M et al. Physical activity
773 for cognitive and mental health in youth: a systematic review of mechanisms. *Pediatr.*
774 2016;138(3):e20161642. doi: 10.1542/peds.2016-1642
- 775 61. Ryan RM & Deci EL. *Self-determination theory: basic psychological needs in motivation,*
776 *development, and wellness.* New York: Guilford Press; 2017.
- 777 62. Vella SA, Benson A, Sutcliffe J, McLaren C, Swann C, Schweickle MJ et al. Self-determined
778 motivation, social identification and the mental health of adolescent male team sport
779 participants. *J Appl Sport Psychol.* 2021;33(4):452-466. doi: 10.1080/10413200.2019.1705432

- 780 63. Biddle SJ, Mutrie N & Gorely T. Psychology of Physical Activity: Determinants, Well-Being
781 and Interventions. United States: Taylor & Francis; 2021. ISBN: 9781000367553
- 782 64. Ekkekakis P, Zenko Z & Vazou S. Do you find exercise pleasant or unpleasant? The Affective
783 Exercise Experiences (AFFEXX) questionnaire. Psychol Sport Exerc. 2021;55:101930.
784 doi: 10.1016/j.psychsport.2021.101930
- 785 65. Reinboth M & Duda JL. Effects of competitive environment and outcome on achievement
786 behaviors and well-being while engaged in a physical task. Sport Exerc Perform Psychol.
787 2016;5(4):324–336. doi: 10.1037/spy0000075
- 788 66. Knowles C, Shannon S, Prentice G & Breslin G. Comparing mental health of athletes and non-
789 athletes as they emerge from a COVID-19 pandemic lockdown. Front Sports Act Living.
790 2021;3:612532. doi: 10.3389/fspor.2021.612532
- 791 67. O'Donovan G, Blazeovich AJ, Boreham C, Cooper AR, Crank H, Ekelund U et al. The ABC of
792 Physical Activity for Health: a consensus statement from the British Association of Sport and
793 Exercise Sciences. J Sports Sci. 2010;28(6):573–591. doi: 10.1080/02640411003671212
- 794 68. Beyer KM, Kaltenbach A, Szabo A, Bogar S, Nieto FJ & Malecki KM. Exposure to
795 neighborhood green space and mental health: evidence from the survey of the health of
796 Wisconsin. Int J Environ Res Public Health. 2014;11(3): 3453-3472. doi:
797 10.3390/ijerph110303453
- 798 69. Bikomeye JC, Balza J & Beyer KM. The impact of schoolyard greening on children's physical
799 activity and socioemotional health: A systematic review of experimental studies. Int J Environ
800 Res Public Health. 2021;18(2):535. doi: 10.3390/ijerph18020535
- 801 70. Teychenne M, Abbott G, Lamb KE, Rosenbaum S & Ball K. Is the link between movement and
802 mental health a two-way street? Prospective associations between physical activity, sedentary
803 behaviour and depressive symptoms among women living in socioeconomically disadvantaged
804 neighbourhoods. Prev Med. 2017;102:72–78. doi: 10.1016/j.ypmed.2017.07.005

- 805 71. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G et al. World Health
806 Organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med.*
807 2020;54(24):1451-1462. doi: 10.1136/bjsports-2020-102955
- 808 72. Department of Health and Social Care. *UK Chief Medical Officers' Physical Activity*
809 *Guidelines*. London: Department of Health and Social Care, 2019. Available from:
810 [https://www.gov.uk/government/publications/physical-activity-guidelines-uk-chief-medical-](https://www.gov.uk/government/publications/physical-activity-guidelines-uk-chief-medical-officers-report)
811 [officers-report](https://www.gov.uk/government/publications/physical-activity-guidelines-uk-chief-medical-officers-report). Accessed: 19/06/2023
- 812 73. Sanders JP, Biddle SJ, Gokal K, Sherar LB, Skrybant M, Parretti HM et al. 'Snackitivity™' to
813 increase physical activity: Time to try something different? *Prev Med.* 2021;153: 106851.
814 doi: 10.1016/j.ypmed.2021.106851
- 815 74. Fogg BJ. *Tiny habits: The small changes that change everything*. London: Houghton Mifflin
816 Harcourt; 2019. ISBN: 9780358003328
- 817 75. Bennie JA, Teychenne MJ, De Cocker K & Biddle SJ. Associations between aerobic and
818 muscle-strengthening exercise with depressive symptom severity among 17,839 US adults. *Prev*
819 *Med.* 2019;121:121-127. doi: 10.1016/j.ypmed.2019.02.022
- 820 76. Shannon S, Carlin A, Woods C, Nevill AM, Murphy N & Murphy MH. Adherence to aerobic
821 and muscle-strengthening components of the physical activity guidelines and mental health.
822 *Health Promot Int.* 2022;37(5):1-11. doi: 10.1093/heapro/daac083
- 823 77. Ravens-Sieberer U, Auquier P, Erhart M, Gosch A, Rajmil L, Bruil J et al. The KIDSCREEN-
824 27 quality of life measure for children and adolescents: psychometric results from a cross-
825 cultural survey in 13 European countries. *Qual Life Res.* 2007;16(8):1347-1356.
826 doi: 10.1007/s11136-007-9240-2
- 827 78. Hagborg WJ. The Rosenberg Self-Esteem scale and Harter's Self-Perception profile for
828 adolescents: a concurrent validity study. *Psychol Sch.* 1993;30(2):132-136.
829 doi: 10.1002/1520-6807(199304)30:2<132::aid-pits2310300205>3.0.co;2-z

- 830 79. Katzmarzyk PT & Mason C. The physical activity transition. *J Phys Act Health*. 2009;6(3):269–
831 280. doi: 10.1123/jpah.6.3.269
- 832 80. Stiglic N & Viner RM. Effects of screentime on the health and well-being of children and
833 adolescents: a systematic review of reviews. *BMJ Open*. 2019;9(1):e023191.
834 doi: 10.1136/bmjopen-2018-023191
- 835 81. Fortier I, Raina P, Van den Heuvel ER, Griffith LE, Craig C, Saliba M et al. Maelstrom
836 Research guidelines for rigorous retrospective data harmonization. *Int J Epidemiol*.
837 2017;46(1):103-105. doi: 10.1093/ije/dyw075
- 838 82. Wood JA. Methodology for dealing with duplicate study effects in a meta-analysis. *Organ Res*
839 *Methods*. 2008;11(1):79-95. doi: 10.1177/1094428106296638

840 **Supporting information**

- 841 **S1 Table: References of all studies included in the review**
- 842 **S2 File: Study modifiers**
- 843 **S3 Table: Duplicated datasets**
- 844 **S4 Table: Study characteristics**
- 845 **S5 Table: Studies of individual and group-based physical activity**
- 846 **S6 Table: Studies of leisure-time and school-based physical activity**
- 847 **S7 Table: Studies investigating effects of WHO MVPA guideline adherence**
- 848 **S8 File: Abbreviations for datasets and measures used**
- 849 **S9 File: Database search terms**
- 850 **S10 File: Review protocol**

Identification of studies via databases and registers

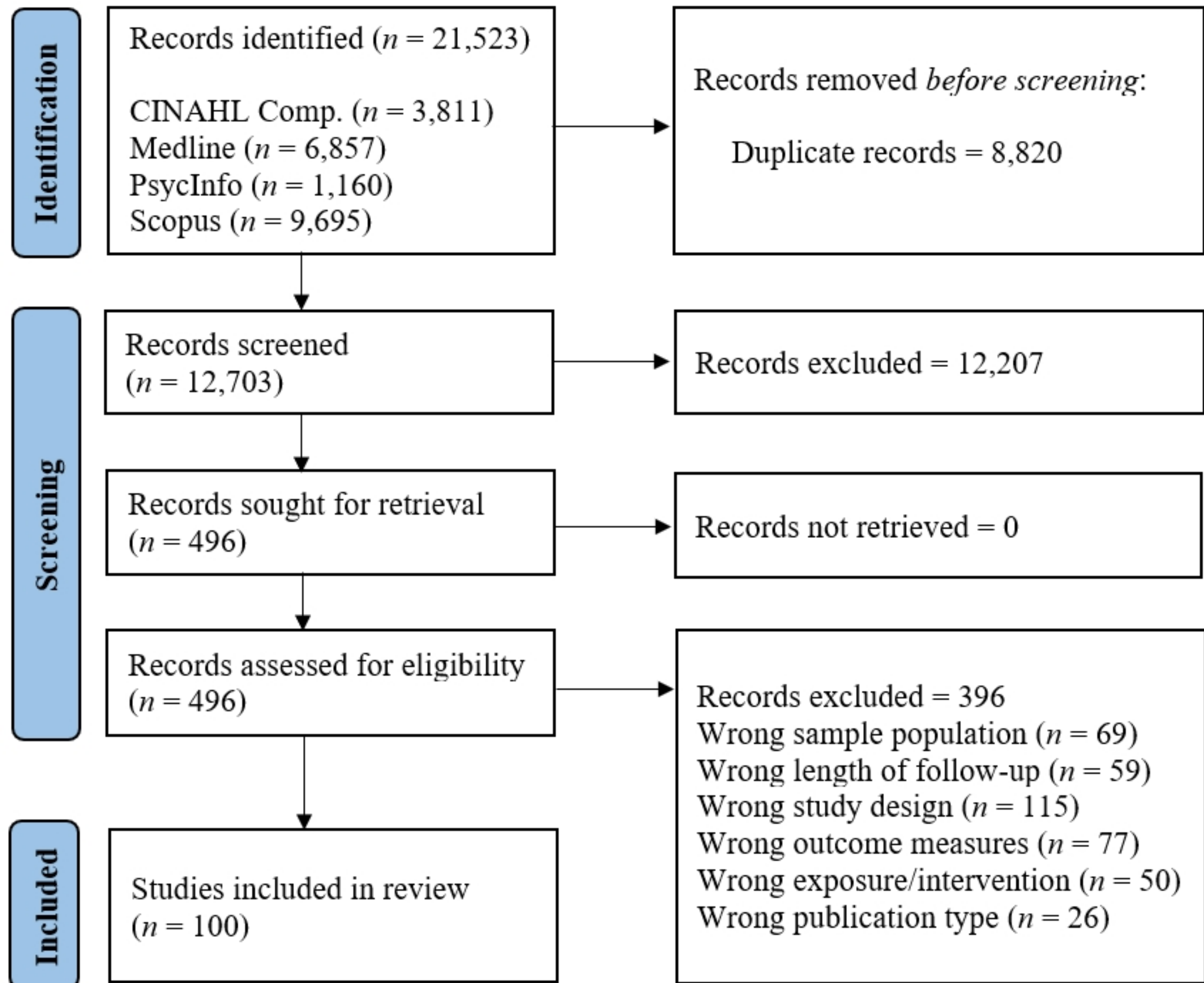


Fig 1

Year of Publication

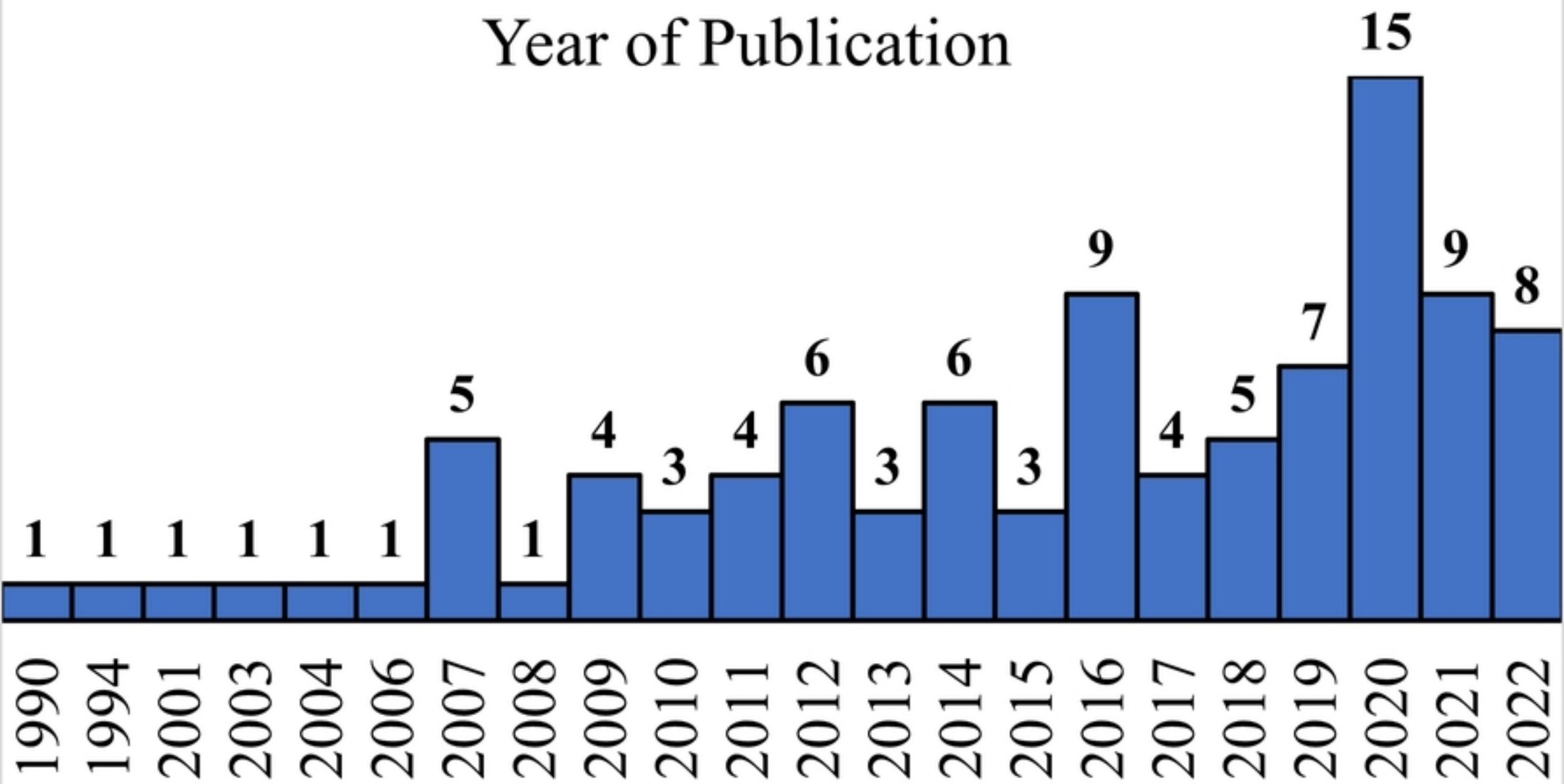


Fig2

Length of Studies (Years)

■ Beneficial Effect
■ No Effect

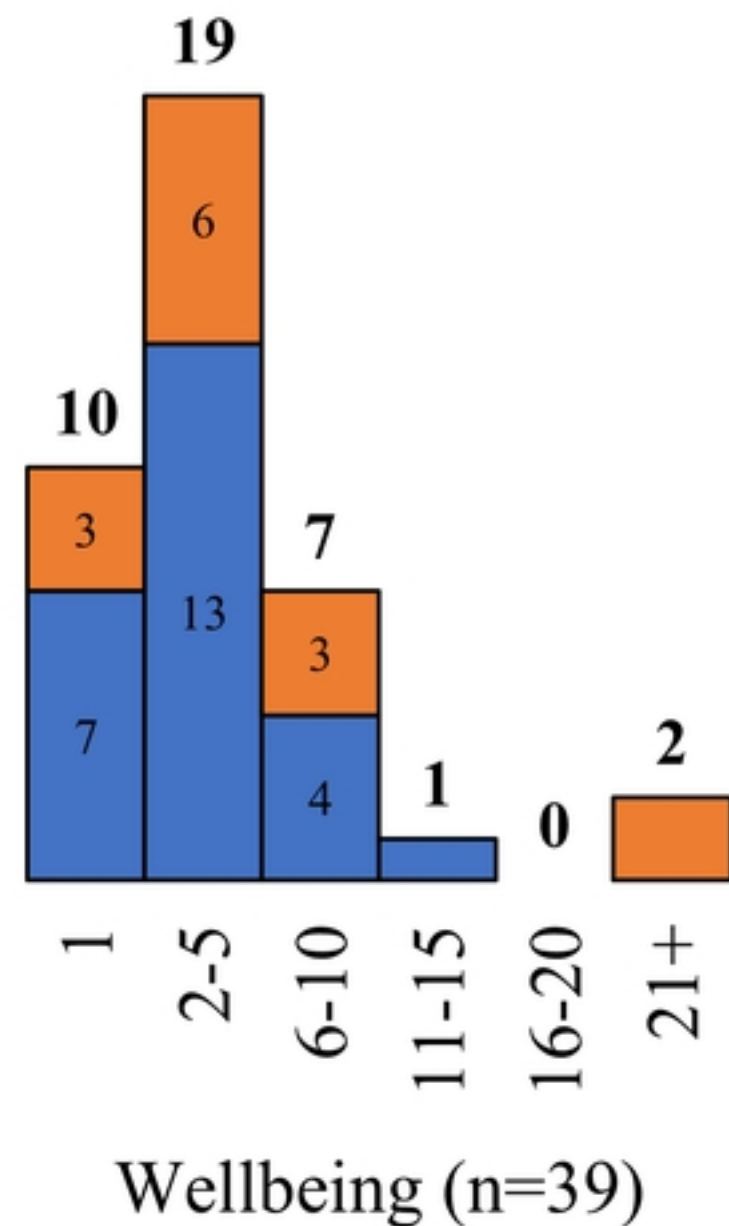
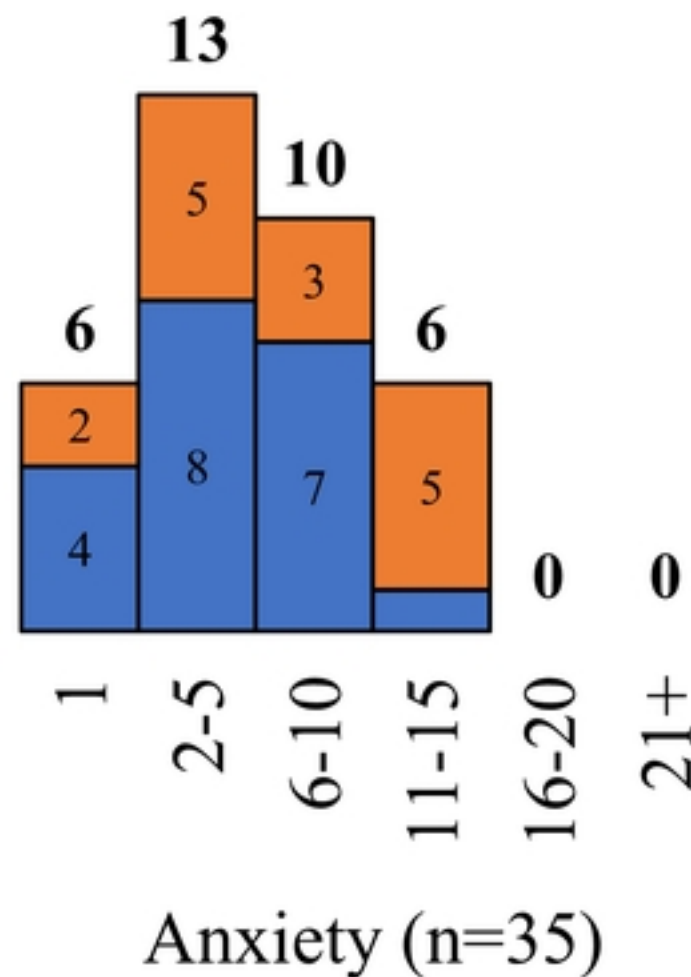
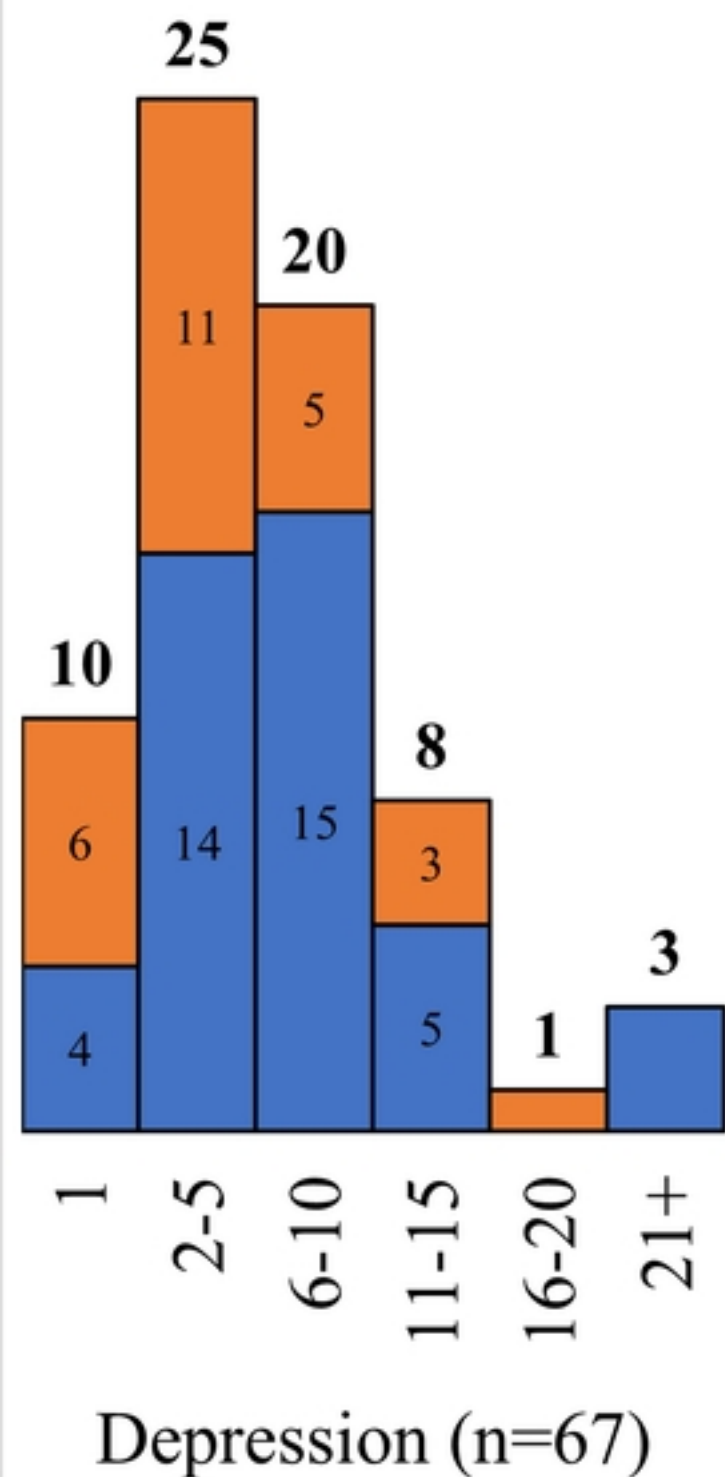


Fig3

Length of Studies by Type (Years)

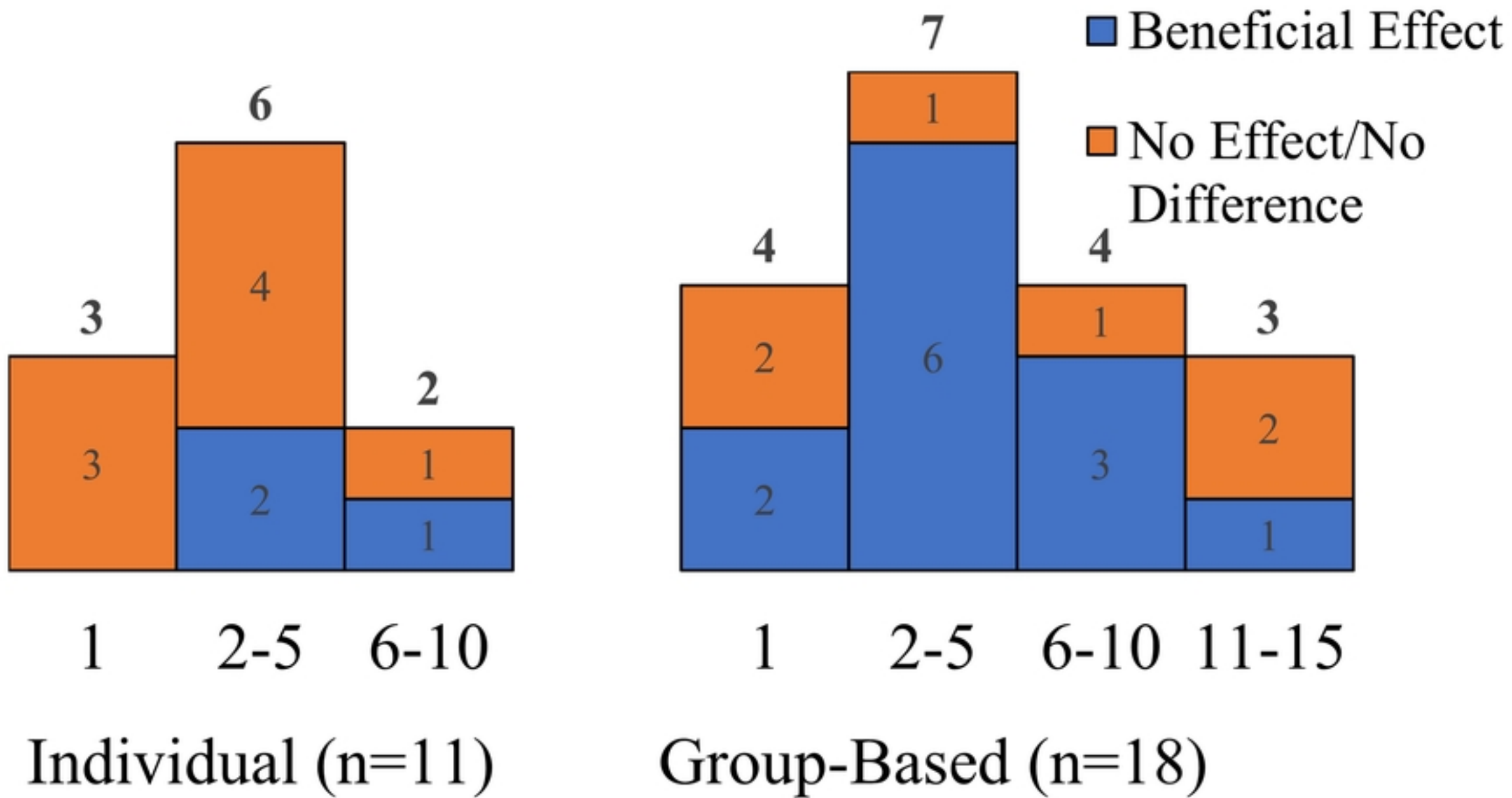


Fig4

Length of Studies by Life Domain (Years)

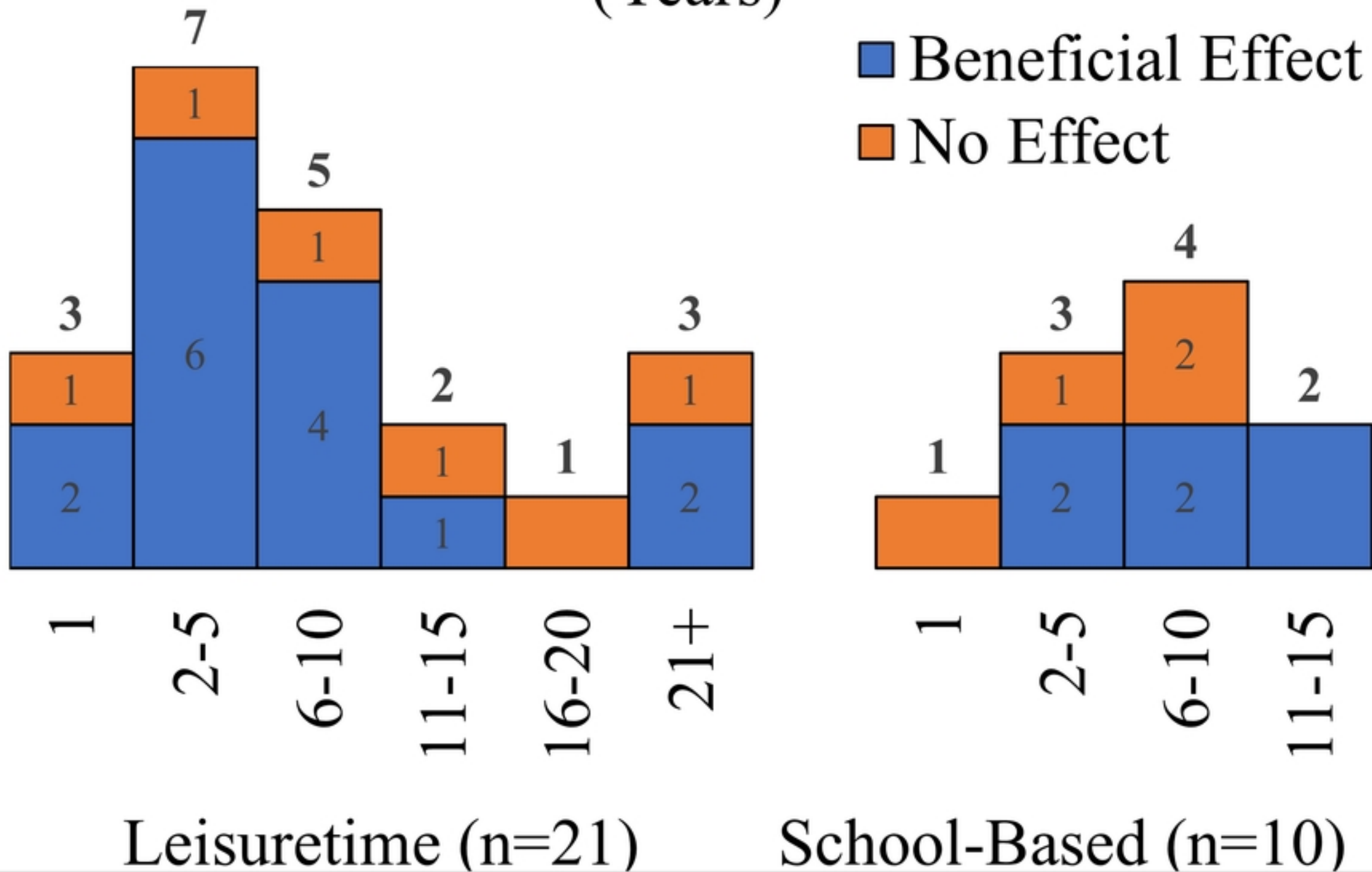


Fig5

Length of Studies Investigating WHO MVPA Guidelines (Years)

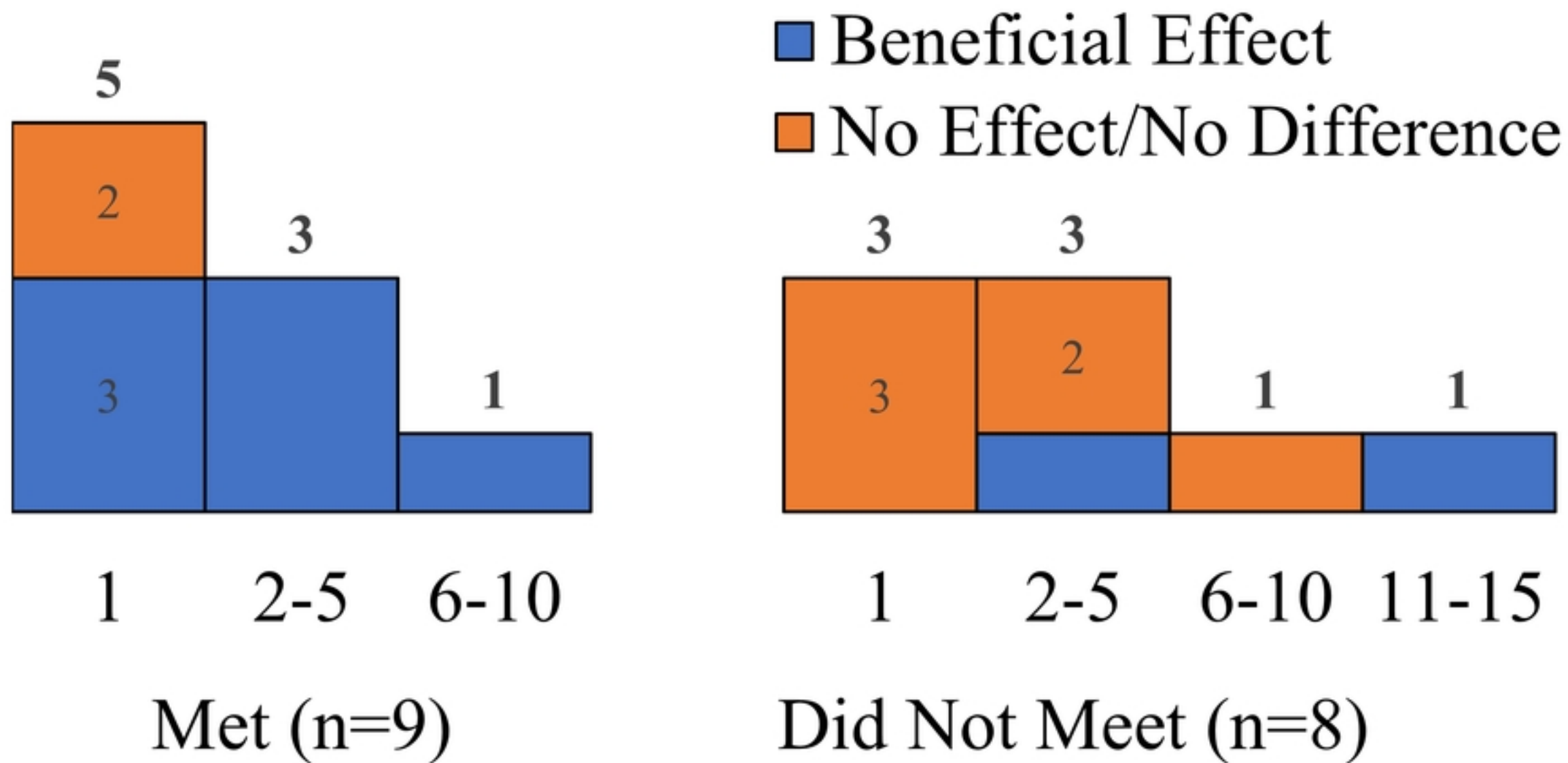


Fig6