The Queensland Twin Adolescent Brain Project, a longitudinal study of adolescent brain development

Lachlan T. Strike[1], Narelle K. Hansell[1], Kai-Hsiang Chuang[1,2], Jessica L. Miller[1], Greig I. de Zubicaray[3], Paul M. Thompson[4], Katie L. McMahon[5], Margaret J. Wright[1,2]

1. The University of Queensland, Queensland Brain Institute, Brisbane, QLD 4072, Australia
2. The University of Queensland, Centre for Advanced Imaging, Brisbane, QLD 4072, Australia
3. School of Psychology and Counselling, Faculty of Health, Queensland University of Technology, Kelvin Grove, QLD 4059, Australia
4. Imaging Genetics Center, Mark & Mary Stevens Institute for Neuroimaging & Informatics, Keck School of Medicine, University of Southern California, Los Angeles, USA
5. School of Clinical Sciences, Centre for Biomedical Technologies, Queensland University of Technology, Brisbane, QLD 4000, Australia

Corresponding author: Lachlan T. Strike (l.strike1@uq.edu.au)

We describe the Queensland Twin Adolescent Brain (QTAB) dataset and provide a detailed methodology and technical validation to facilitate data usage. The QTAB dataset comprises multimodal neuroimaging, as well as cognitive and mental health data collected in adolescent twins over two sessions (session 1: N = 422, age 9-14 years; session 2: N = 304, 10-16 years). The MRI protocol consisted of T1-weighted (MP2RAGE), T2-weighted, FLAIR, high-resolution TSE, SWI, resting-state fMRI, DWI, and ASL scans. Two fMRI tasks were added in session 2: an emotional conflict task and a passive movie-watching task. Outside of the scanner, we assessed cognitive function using standardised tests. We also obtained self-reports of symptoms for anxiety and depression, perceived stress, sleepiness, pubertal development measures, and risk and protective factors. We additionally collected several biological samples for genomic and metagenomic analysis. The QTAB project was established to promote health-related research in adolescence.

Background & Summary

Adolescence is critical for understanding brain changes associated with depression[1-3], as nearly half of lifetime diagnoses begin by age 14[4]. Adolescents who experience depression are more likely as adults to have poor mental and physical health, lower levels of educational attainment, lower salaries, and more relationship difficulties[5-7]. During adolescence, the brain's cognitive control, emotion, and reward-related circuitries are undergoing substantial development[8,9] – developmental changes that can be impaired by restricted sleep, which is common among teenagers[10]. In addition, hormonal surges and consequent physical maturation linked to pubertal development in adolescence are thought to affect multiple aspects of brain development, social cognition, and peer relations[11].

Sex differences in mental health problems emerge in puberty, a developmental period characterised by rapid increases in estrogen in girls and testosterone in boys. Before puberty, depression occurs relatively infrequently in girls and boys. After puberty, there is a sharp increase in the incidence of depression, with adolescent girls around twice as likely to experience depression as boys[12]. Children who start puberty earlier are at greater risk for depressive symptoms and anxiety, especially girls[13]. Adolescent boys, by contrast, are more likely than girls to develop substance use disorders and die by suicide[14]. Importantly, these pubertal hormones cross the blood-brain barrier, influence brain development, and affect various signalling pathways (e.g., neurotransmitter activity) that underlie mood and cognition[15]. Thus, puberty involves transformation across virtually every psychobiological domain—endocrine, neural, physical, cognitive, and socioemotional—and represents a vulnerable time during which depressive symptoms and psychiatric conditions may emerge.

Magnetic resonance imaging (MRI) allows developmental brain changes to be studied *in vivo* while retaining the advantage of being non-invasive, thus providing an invaluable tool for longitudinal research in a population sample. In recent years, terrific

progress has been made towards characterizing typical adolescent structural and functional development[1,2,8,16,17] and understanding brain structure and function changes during adolescence associated with various mental health disorders[1-3,16]. However, there is still little work in this sensitive period of neurodevelopment using genetically informative samples. Further, prior genetic studies have primarily been cross-sectional, and we lack knowledge of how this critical transitional period between childhood and adulthood contributes to optimal cognitive and emotional functioning as well as vulnerability to brain disorders and mental illness. Thus, longitudinal studies are vital to characterise the structural and functional integrity of the brain in genetically informative adolescent samples.

To this end, we present the Queensland Twin Adolescent Brain (QTAB) dataset: a multimodal neuroimaging dataset of Australian adolescent twins with mental health, cognition and social behaviour data collected over two time points (i.e., sessions). Comparisons within and between identical (MZ) and non-identical (DZ) twin pairs - further powered by multiple assessments - provide a rich information source about genetic and environmental contributions to developmental associations and enable stronger tests of causal hypotheses than do comparisons involving unrelated adolescents[18].

The QTAB project reflected a concerted effort to:

1. Assess brain development in a large population sample of adolescent twins and disentangle the influence of genetic and environmental factors on neurodevelopmental trajectories.
2. Investigate whether neurodevelopmental trajectories are the same for adolescent boys and girls and whether any differences are linked with pubertal maturation and genetic and environmental factors.
3. Concurrently assess cognitive function to examine how neurodevelopmental changes are associated with the acquisition of complex tasks (e.g., working memory) and behaviour (e.g., emotional functioning) and whether shared genetic or environmental factors underpin associations.
4. Concurrently assess psychiatric symptoms and examine whether those with anxiety or depressive symptoms follow a different developmental pathway, whether they exhibit early or delayed brain maturation, and whether genetic and environmental factors influence these atypical developmental trajectories.

Here, we make this multimodal QTAB dataset publicly available and provide an overview of the collected imaging, mental health, cognition, and social behaviour data. In doing so, we hope to make a significant and novel contribution to our understanding of the numerous emotional and behavioural health problems that emerge during the developmental period of adolescence. At the same time, we believe the QTAB dataset provides a significant opportunity for combining data with existing adolescent studies such as the Adolescent Brain Cognitive Development (ABCD) [17], Lifespan Human Connectome Project in Development (HCP-D) [19], and the NKI-Rockland Sample Longitudinal Discovery of Brain Development Trajectories[20] projects, as well as consortia efforts such as ENIGMA Lifespan[21,22].

Methods

Recruitment of the QTAB Cohort

Participants were recruited through local and national twin registries (i.e., The Queensland Twin (QTwin) Register [https://www.qimrberghofer.edu.au/qtwin] and Twin Research Australia [TRA] [https://www.twins.org.au]) or via flyers and online postings (study website). Families were invited to participate if the twins were aged 9 to 13 years (i.e., year of birth 2004-2010) and lived within approximately 2 hours travelling distance of the study centre in Brisbane (i.e., South-East Queensland). Four hundred and forty families met the inclusion criteria (QTwin Register: 356 families; TRA: 48; study website: 36), of which 63 (14% of families) declined on the first approach (letter followed by phone call) and a further 166 families were excluded (16%

excluded on screening; 22% passive refusals). Exclusion criteria included serious medical, neurological, psychiatric (including diagnosis of autism spectrum disorder [ASD] or attention deficit hyperactivity disorder [ADHD]), or cardiovascular conditions, history of a serious head injury, and any cognitive, physical, or sensory challenges that would limit ability to understand or complete procedures. In addition, both twins (i.e., co-twins) had to have no MRI contraindications (e.g., metal implants or artefact inducing orthodontic braces). Thus, the total available sample comprised 211 families, including 206 twin pairs and 5 triplet sets. Only two individuals from each trio were scanned, and hereafter are referred to as a twin pair.

QTAB Protocol and Participants

Baseline measures (i.e., session 1 data) were collected between June 2017 and October 2019, followed by a second wave of data collection (i.e., session 2 data) from November 2019 to January 2021. Both sessions, each taking ≈3.5 hours, included structural and functional brain imaging, assessments of pubertal status, cognition, anxiety and depressive symptoms plus risk and protective factors, sleep-wake behaviour, and the collection of several biological samples (Fig. 1). While the target interval between sessions was ≈18 months, to enable as many participants as possible to return for their second assessment, and because of the interruption to data collection due to the COVID-19 pandemic, we allowed for some latency jitter, which resulted in an interval of 13 to 30 months (mean 20 months).

Co-twins attended the study centre together and were assessed in parallel, with one twin undergoing brain imaging while the other completed assessments outside of the scanner. In addition, the parent accompanying the twins - usually the mother (94% of families) - completed a parent questionnaire providing family information and parental assessments of the twins. Hair, blood, and urine samples were obtained during the session. Early morning saliva samples were collected at home on the morning of their session visit and brought to the study centre, as were any available baby teeth. At the end of the session, the twins were provided with an accelerometer to objectively measure their sleep behaviour for two weeks and a stool collection pack. The accelerometer and stool sample were returned to the study centre by post and courier, respectively. Informed written consent, including data sharing permission, was obtained from both the twins and the attending parent; families were compensated $150 (AUD) per session. Ethics approval for this study was obtained through the Children's Health Queensland Human Research Ethics Committee (HREC) (reference HREC/16/QRCH/270) and the University of Queensland HREC (reference 2016001784).

Five participants were unable to be scanned in session 1 due to a reluctance to enter the MRI system. Thus, the imaging dataset at session 1 consists of a maximum of 417 participants (48% female, mean age 11.3 ± 1.4 years, age range 9.0 to 14.4 years, 82% right-handed), including 206 twin pairs (109 MZ, 97 DZ) and 5 unpaired twins. In session 2, we focused on pairs with good quality images at session 1 (i.e., both twins had good T1w scans). In total, 59 families did not return for session 2 (39 were no longer eligible (i.e., braces, moved interstate and/or poor T1w scans at session 1) and 20 declined), so the dataset was limited to 304 participants (52% female, mean age 13.04 ± 1.52 years, age range 10.1 to 16.4 years, 85% right-handed), comprising of 152 twin pairs (81 MZ and 71 DZ). Key demographics for QTAB participants are presented in Table 1.

Image Acquisition

Structural and functional scans were acquired on a 3T Magnetom Prisma (Siemens Medical Solutions, Erlangen) paired with a 64-channel head coil at the Centre for Advanced Imaging, University of Queensland. Session 1 scans were collected in a fixed order: localisers, 3D T1-weighted, two runs of resting-state fMRI (rs-fMRI), four runs of diffusion-weighted imaging (DWI), three runs of high-resolution Turbo Spin

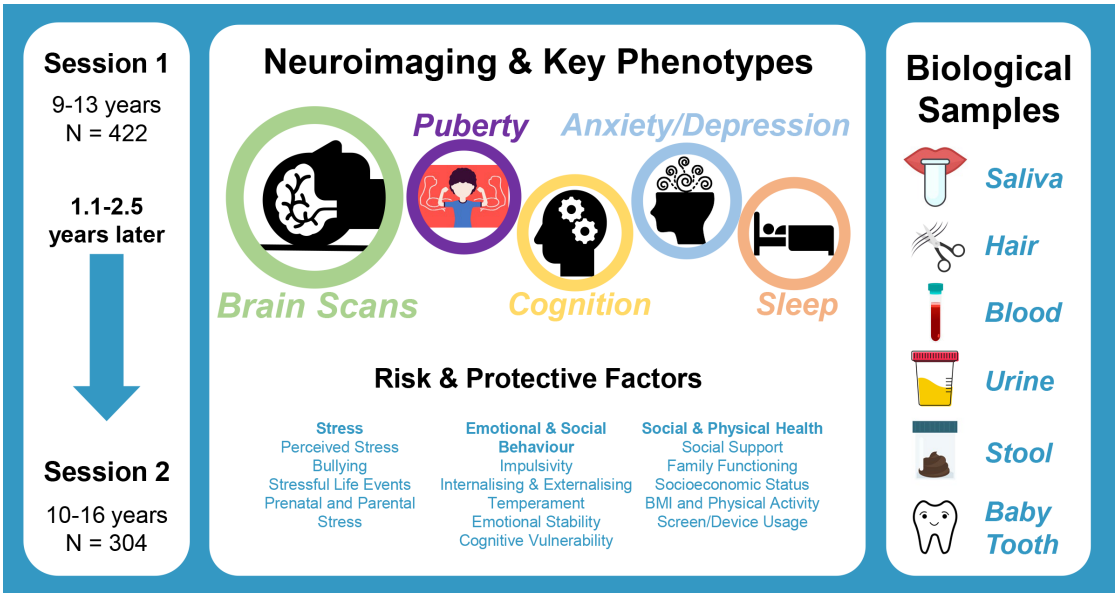

Fig. 1 Data collection overview. Data from imaging and questionnaire/cognition components were collected concurrently between twin pairs (i.e., while one twin underwent brain imaging, their co-twin completed the mental health/behavioural questionnaire and cognitive testing). An early morning saliva sample was collected at home and brought to the session, along with any available baby teeth. Hair, blood, and urine samples were obtained during the session, with faecal samples and accelerometry measures of sleep behaviour collected at home. Neuroimaging and key phenotypes, saliva and hair samples were collected at sessions 1 and 2.
****

Echo (TSE) imaging, 3D fluid-attenuated inversion recovery (FLAIR), 3D T2-weighted, Arterial Spin Labelling (ASL), and 3D susceptibility-weighted-imaging (SWI). In session 2, scan order was very similar, with the addition of two task-based fMRI scans midway through the imaging session, i.e., localisers, T1-weighted, two runs of rs-fMRI, four runs of DWI, two runs of TSE, passive movie-watching task (Partly Cloudy), emotional conflict fMRI task, ASL, SWI, FLAIR, T2w. Immediately after each scan, the images were checked for apparent artifacts (e.g., head motion), and scans were re-acquired if necessary. Scan parameters are summarised in Table 2**.** Protocol optimisations (e.g., increase in the number of TSE volumes to improve hippocampus and amygdala coverage) and sequence reissues (e.g., to the T1w) that occurred across participants or sessions are noted as a protocol code (detailed in Table 2 and provided in *participants.tsv*; see *Data Records*). MRI acquisition took approximately 1 hour in session 1 and 1:15 hours for session 2.

Image Modalities

Anatomical

A 3D whole-brain T1-weighted image was acquired using a work-in-progress Magnetisation Prepared with 2 Rapid Gradient Echoes (MP2RAGE) sequence. The MP2RAGE sequence combines two individual images acquired at different inversion times (inv-1_MP2RAGE, inv-2_MP2RAGE) to produce an image corrected for B_1 and T2* inhomogeneities[28,29]. This "uniform" image (UNIT1) improves T_1 contrast compared to conventional T1-weighted protocols. However, a by-product of the MP2RAGE protocol is an amplification of background noise (also known as "salt and pepper" noise), which is problematic for automatic registration and segmentation algorithms (see *Usage Notes* for techniques to address this). The sequence generates a denoised uniform image (UNIT1_denoised), though some background noise is still present in this

Participants	Session 1	Session 2
N	422	304
Age (years)	11.3±1.4	13.0±1.5
Sex (% biological female/male)	48.3/51.2	51.6/48.4
Handedness (% right/left)	82.2/17.8	84.9/15.1
Zygosity (% MZ/DZ) **School Grade Year** (% Y3-4/Y5-6/Y7-9/Y10-11)	52.6/47.4 22.3/46/31.7/0	53.3/46.7 2.6/31.6/56.9/8.9
Families		
N	211	152
Ancestry (group identification as % of all reported groups) Most Common: Australian/English Non-Caucasian: All/Asian/First Nations Australian	30.6/24.1 6.9/3.4/0.9	28.8/24.1 5.6/3.1/0.9
Socioeconomic Status (% low/average/high)	9.0/34.1/56.9	9.2/35.5/55.3
Education Mother (% Univ/Trade/School/NA) Father (% Univ/Trade/School/NA)	52.6/28.9/18.5/0 39.3/32.7/22.2/5.8	51.3/27.7/21.1/0 40.7/31.5/22.4/5.4
Birth Information Gestational age (% preterm (<37 wks)/NA) Weight (% low: 1500-2499g/very low: <1500g/NA) Maternal age in years (<20/20-29/30-39/≥40/NA) % Assisted Reproduction Technology (e.g., IVF)	54.5/1.9 36.5/5.7/57.8 0.9/30.3/62.6/3.3/2.8 19.9	52.6/1.3 36.5/5.6/57.9 1.3/29.6/62.5/3.9/2.6 19.7

Table 1. Key demographics for QTAB participants (including family and birth information).

Notes: **Handedness** was based on participant self-report. **Zygosity** in same-sex twins was determined by genotypic data (93% of twin pairs) or parental questionnaire (7% of twin pairs). Genome-wide identity-by-descent (IBD) was estimated for twin pairs using PLINK[23], after quality checking procedures[24]. Pairs whose proportion IBD (PI_HAT) was greater than 0.99 were classified as MZ pairs; remaining pairs (whose proportion IBD [PI_HAT] was between 0.41 and 0.61) were classified as DZ pairs. **School grade year** was classified as follows: Year 3-4 aged 8–10yrs (primary school), Year 5-6 aged 10–12 (last 2yrs of primary school), Year 7-9 aged 12–15 (middle/high school), Year 10-11 aged 15-17 (high school). **Ancestry** was based on self-perceived group/cultural identification with ancestry options obtained from the 2016 Australian Census. Each family identified 1 to 7 cultural groups (M=2.1±0.9). The most-reported ancestries (as a proportion of all reported ancestries) were Australian (30.6%) and English (24.2%), with the most reported non-Caucasian ancestry being Chinese (1.3% of all reported ancestries). Ancestries reported for the Australian population in the 2016 census were similar (33.5% Australian, 36.1% English, and 5.6% Chinese). **Socioeconomic status** refers to Socio-Economic Index for Areas (SEIFA) Index of Relative Socio-Economic Advantage and Disadvantage (IRSAD) deciles (low = deciles 1-3, average = deciles 4-7, high = deciles 8-10). **Education** (parent self/partner report) classifications are based McMillan et al. [25], with post-school diploma and trade Certificate categories collapsed into "Trade". **Birth information** was provided through parental report. The birth statistics available for QTAB (note that birth weight is not available for over half the sample) are similar to those for twins born in Australia between 2001 and 2010[26], of whom 53.6% were born preterm, while 50.2% were low birthweight, and 8.7% were very low birthweight. QTAB maternal age was also similar to national statistics (reported as 2.0/33.2/60.2/4.7%[26], based on the age categories in Table 1). An Assisted Reproductive Technology (ART) reported rate of 19.9% for QTAB twins born 2004-2010 is consistent with the 23.3% reported for 9,831 twin deliveries in Australia from 2007 to 2011[27].
Abbreviations: *DZ* dizygotic; *IVF* in vitro fertilisation; *MZ* monozygotic; *NA* not available; *Univ* university.
****

image. The MP2RAGE sequence was reissued during session 2, producing a denoised uniform image (UNIT1_denoised) without background noise.

T2-weighted and FLAIR whole-brain 3D sequences were acquired to examine brain pathology and to improve the visibility of brain regions bordering CSF (e.g., cortical pial surface, periventricular structures). Due to time constraints, the T2-weighted and FLAIR sequences were excluded for some participants (≈5% in session 1, ≈10% in session 2). T2-weighted TSE scans (slab aligned orthogonally to the hippocampus) were acquired to examine the structural properties of the hippocampus and amygdala at the subdivision level (three runs collected at session 1, reduced to two runs at session 2).

Susceptibility

SWI scans were acquired to study the pathological mechanisms underlying paediatric brain diseases and disorders. In addition to a combined magnitude and phase image (*_swi), the SWI sequence produces magnitude (*_part-mag_GRE), phase (*_part-phase_GRE), and minimum intensity projection (*_minIP) images.

Diffusion
Multi-shell diffusion weighted images were acquired in four runs with opposing phase encode directions (two A>>P and two P>>A runs). Each run consisted of 23 diffusion weighted volumes (which includes b = 0 s/mm^2 (3 volumes), b = 1000 s/mm^2 (5 volumes), b = 3000 s/mm^2 (15 volumes).

Functional
Resting-state scans were acquired consecutively over two runs with reversed-phase encoding directions (anterior-posterior (A>>P), posterior-anterior (P>>A)). Participants were instructed to relax with their eyes open, with an abstract visual stimulus displayed to improve resting-state compliance[30]. The stimulus for functional scans was presented on a back-projection screen that the participants viewed via a mirror attached to the head coil.

In session 2, the emotional conflict task was based on a previously characterised emotional conflict task[31,32] and probed emotion-relevant neural processes engaging the amygdala and other structures that represent the processing of emotions and faces. In each trial, participants were presented with a face displaying a fearful or happy expression and the words "happy" or "fear" written across the face to create congruent and incongruent emotional stimuli. Participants were asked to identify the facial emotion while ignoring the word overlayed on the face. Responses were made via a button press (left button = fearful face, right button = happy face). Stimuli were presented for 1,000 ms, with a varying interstimulus interval of 2,000 – 4,000 ms (average 3,000 ms; 163 trials in total), presented in a fixed order, counterbalanced across trial types for expression, word, and sex. Participants completed an in-scanner practice task (8 trials) immediately prior to the task. Stimuli were presented using the Cogent toolbox (www.vislab.ucl.ac.uk/cogent_2000.php) implemented in the MATLAB (www.mathworks.com) programming environment.

Session 2 participants also completed a passive movie-watching task involving the short animated film Partly Cloudy[33] (silent version). This task was selected based on prior work examining cortical and cognitive changes in adolescents' social development[34,35]. Participants were instructed to remain still and watch the film. The stimulus was transcoded from a PAL format DVD (see *Usage Notes*) and presented using E-Prime (Psychology Software Tools, Pittsburgh, PA); there was no rest period prior to stimulus presentation. We provide condition timings based on a prior study[34] (see *Data Records* and *Usage Notes*); however, we note that such a naturalistic movie-watching task can be used to examine a range of different cognitive, emotional, and perceptual processes.

Perfusion
ASL scans provide a measure of cerebral blood flow without the requirement of contrast agents or radiation. In session 1 we used a work-in-progress 2D pseudo-continuous ASL (PCASL) sequence. The product sequence was available at the start of data collection for session 2. In both sessions, a separate M0 scan was acquired using a shortened version of the ASL sequence to produce an improved measure of M0 for calibration purposes.

QTAB Non-Imaging Phenotypes
Table 3 provides a detailed list of the scales and tests used to assess *puberty*, *cognition*, *anxiety and depressive symptoms*, *emotional and social behaviours*, *social support and family functioning*, *stress*, *sleep and physical health*, *early life and family demographic factors*, *dietary behaviour*, *and COVID-19 pandemic specific assessments*. In general, the same measures were collected in sessions 1 and 2, and for several scales, we included both adolescent and parent versions. Selection of tests and self-report scales for phenotypic characterisation was based on validation for use

Modality	Voxel Size	Key Parameters	NIfTI Filename
Anatomical			
T1w	0.8 x 0.8 x 0.8 mm	3D MP2RAGE, sagittal acquisition, TR/TE = 4000/2.99 ms, TI1/TI2 = 700/2220 ms, FA1/FA2 = 6/7 deg, slices = 192, FOV = 256 x 240 mm, in-plane acceleration = 3, duration = 6:04 min:sec. *Protocol change*: T1w reissue during session 2 (protocol code 2c).	*_inv-1_MP2RAGE, *_inv-2_MP2RAGE, *_UNIT1, *_UNIT1_denoised
T1w (reissue)	As above, except: TE = 3.03 ms		
T2w	1.0 x 1.0 x 1.0 mm	3D T2, sagittal acquisition, TR/TE = 3200/408 ms, slices = 176, FOV = 256 x 256 mm, in-plane acceleration = 2, duration = 3:31 min:sec	*_T2w
FLAIR	1.0 x 1.0 x 1.0 mm	3D T2-FLAIR, sagittal acquisition, TR/TE = 5000/388 ms, slices = 160, FOV = 256 x 256 mm, in-plane acceleration = 2, duration = 5:55 min:sec	*_FLAIR
TSE	0.5 x 0.5 x 1.0 mm (interpolated in k-space to 0.25 x 0.25 x 1.0 mm)	2D TSE slab orthogonal to the hippocampus, TR/TE = 8460/67 ms, FA = 150 deg, slices = 60, FOV = 192 x 192 mm, duration = 4:48 min:sec (each). *Protocol change:* Number of slices was initially 50, increased to 55 to improve hippocampal coverage (protocol code 1b), increased again to 60 to improve amygdala coverage (protocol code 1c).	*_T2w_TSE_run-<index>
Susceptibility			
SWI	1.1 x 1.1 x 2.0 mm	3D SWI, TR/TE = 27/20 ms, FA = 15 deg, slices = 60, FOV = 220 x 198 mm, in-plane acceleration = 2, duration = 2:51 min:sec	*_swi, *_part-mag_GRE , *_part-phase_GRE, *_minIP
Diffusion			
dMRI	2.0 x 2.0 x 2.0 mm	4 acquisitions (2 A>>P, 2 P>>A), TR/TE = 3800/70 ms, FA = 90 deg, in-plane acceleration = 2, slice acceleration = 2, slices = 68, FOV = 244 x 244 mm, b-values: 0 (x 3), 1000 (x 5), 3000 (x 15) s/mm^2, duration = 1:47 min:sec (each)	*_dir-AP_run-<index>_dwi, *_dir-PA_run-<index>_dwi
Functional			
rs-fMRI	2.0 x 2.0 x 2.0 mm	2 acquisitions (1 A>>P, 1 P>>A), TR/TE = 930/30 ms, FA = 52 deg, slices = 72, FOV = 208 x 208 mm, slice acceleration = 6, volumes = 327 (each), duration = 5:13 min:sec (each)	*_task-rest_dir-AP_bold, *_task-rest_dir-PA_bold
fMRI	2.4 x 2.4 x 2.4 mm	TR/TE = 800/30 ms, FA = 52 deg, A>>P, slices = 60, FOV = 216 x 216 mm, slice acceleration = 6, volumes = 380 (Partly Cloudy), volumes = 837 (emotional conflict), duration = 5:11 min:sec (Partly Cloudy), duration = 11:17 min:sec (emotional conflict). *Protocol change:* number of volumes in the emotional conflict task was initially 830, increased to 837 (protocol code 2b).	*_task-partlycloudy_bold, *_task-emotionalconflict_bold
Perfusion			
ASL	1.7 x 1.7 x 4.0 mm	2D PCASL, TR/TE = 4800/13 ms, slices = 24, labelling time = 1500ms, post-labelling delay = 1500ms, volumes = 31 (1 x M0, 15 label-control pairs), slices = 24, FOV = 220 x 220 mm, duration = 2:30 min:sec	*_asl
ASL M0	1.7 x 1.7 x 4.0 mm	As above, except TR = 8000 ms, volumes = 3, duration = 0:26 min:sec (session 1), 0:34 min:sec (session 2)	*_m0scan
Field Maps			
Field maps	2.4 x 2.4 x 2.4 mm	2 acquisitions (1 A>>P, 1 P>>A), TR/TE = 6350/63 ms, slices = 60, FOV = 216 x 216 mm, volumes = 4 (each), duration = 0:33 min:sec (each)	*_dir-AP_epi, *_dir-PA_epi

Table 2. MRI acquisition parameters for the QTAB project.
Notes: Voxel size was optimised for each modality. The scanning protocol code for each participant is defined in *participants.tsv* (see *Data Records*).
Abbreviations: *A>>P* anterior-posterior phase encoding; *deg* degree; *FA* flip angle; *FOV* field of view; *ms* milliseconds; P>>*A* posterior-anterior phase encoding; *TE* time to echo; *TI* inversion time; *TR* repetition time
<index> refers to run number.
* Full NifTI filename begins with sub-<participant_id>_ses-<session_id>
****

with adolescents (i.e., gold standard tasks with known reliability and applicability to the age range of the QTAB cohort) and ease of administration (i.e., suitability or adaptation for online assessment with an iPad or computer, and whether scales and tests were freely available or provided at a modest cost [e.g., NIH Toolbox Cognition Battery[36]]). Tests and scales that were widely used were also prioritised and highly considered.

Puberty

While clinician assessment of sexual maturation is considered superior to self-report, the validation of self-assessment methods has supported their use as an acceptable alternative for research purposes[37]. In QTAB, we assessed pubertal status using a combination of self- and parent-report based on line drawings corresponding to the Tanner stages of pubertal development[38-40] and complementary questions regarding the emergence of secondary sexual characteristics[41]. Further, to aid in tracking

pubertal development, saliva samples were collected at each visit and stored for future analysis of sex hormones (see *Biological Samples* for more information on saliva collection).

Cognition

The cognitive tests focussed on 5 domains: Executive Function, Memory, Language, Processing Speed, and Spatial Processing – abilities that are related to general adolescent development, as well as anxiety and depression. For ease of administration, we selected the iPad version of the NIH Toolbox Cognition Battery (NIHTB-CB) [36] and the Penn Computerized Judgement of Line Orientation (CJOLO) [42,43]. Both are now widely used and facilitate the comparison of QTAB data to independent adolescent cohorts, such as the ABCD and HCP-D. The electronic tests were supplemented with 3 well-validated pencil and paper tests from the Wechsler Intelligence Scale for Children – Fifth Edition[44] (Digit Span Forward & Backward) and Delis Kaplan Executive Function System[45] (FAS Letter Fluency), as well as the parental report Prospective and Retrospective Memory Questionnaire for Children (PRMQ) [46]. In session 2, a novel inclusion was the assessment of social cognition using the Children's Reading Mind in the Eyes Task (RME) [47] and the Empathy Questionnaire for Children and Adolescents (EmQue-CA) [48]. In addition, we obtained consent to access their National Assessment Program – Literacy and Numeracy (NAPLAN) scores (https://www.nap.edu.au/home). NAPLAN is a standardised national assessment of reading, spelling, grammar and punctuation, writing, and numeracy skills undertaken by Australian children in grades 3, 5, 7, and 9 in both government and non-government schools. Currently, NAPLAN scores are not provided as part of the QTAB dataset; please contact authors Greig de Zubicaray and Katie McMahon.

Mental Health

We focused on symptoms of *anxiety* and *depression*, both of which increase in prevalence during adolescence. *Depressive* symptoms were assessed by self-report with the Short Moods and Feelings Questionnaire (SMFQ) [50] and Somatic and Psychological Health Report (SPHERE-21) [52]. *Anxiety* symptoms were measured using the Spence Children's Anxiety Scale (SCAS) [49]. We also assessed traits linked to mental health, which may indicate risk of progression or vulnerability. *Emotional and social behaviour* measures included: *impulsivity* (Short UPPS-P Impulsive Behaviours Scale in Children [IBS] [53]), *internalising and externalising* (e.g., emotional and conduct problem scales from the Strength and Difficulties Questionnaire [SDQ] [54]), *temperament* (Early Adolescent Temperament Questionnaire-Revised [EATQ] [59]), *emotional stability* (Australian-adapted Hierarchical Personality Inventory for Children [HiPIC-A] [55]; also extraversion, amenability, conscientiousness, and imagination), symptoms of *autism* (Autism Spectrum Quotient – 10 items [AQ-10] [Adolescent]) [56]), and *cognitive vulnerability* (rumination subscale from the Children's Response Styles Questionnaire [CRSQ] [57] and the Children's Attributional Style Questionnaire-Revised [CASQ-R] [58], both of which can improve screening of adolescent depression[82]). These dimensional measures investigate constructs relevant to general adolescent maladaptive behaviours (e.g., problem behaviours, risky and impulsive decision making). In addition, *Social Support and Family Functioning* was assessed with the Multidimensional Scale of Perceived Social Support (MSPSS) [60], which captures support from family, friends, and significant other[83] and two scales answered by the parent – the Alabama Parenting Questionnaire (APQ) [61] to capture parenting style and the McMaster Family Assessment Device (FAD) [62], which provides both dimensional and overall measures of family functioning[84]. Peer influences and family factors are posited to influence adolescent brain development[85-87] and moderate adolescent risk of experiencing mental health problems[88-91].

Domain	Measurement Source	Adolescent Session 1	Adolescent Session 2	Parent Session 1	Parent Session 2
Puberty					
	Pubertal Development Sale (PDS)[41]	×	×	×	×
	Sexual Maturation Scale (SMS)[38]	×	×	×	×
Cognition					
Executive Function	NIH Toolbox Cognition Battery (NIHTB-CB)[36] Flanker & Card Sort	×	×		
Memory	NIHTB-CB Picture Sequence	×	×		
	NIHTB-CB List Sorting & Auditory Verbal Learning	×			
	Wechsler Intelligence Scale for Children – Fifth Edition, Digit Span Forward & Backward (DS)[44]	×			
	Prospective and Retrospective Memory Questionnaire for Children (PRMQ)[46]			×	×
Language	NIHTB-CB Picture Vocabulary & Oral Reading Recognition	×	×		
	Delis Kaplan Executive Function System, FAS Letter Fluency (FAS)[45]	×			
Processing Speed	NIHTB-CB Pattern Comparison	×	×		
Spatial Processing	Penn Computerized Judgement of Line Orientation (CJOLO)[42,43]	×			
Social Cognition	Empathy Questionnaire for Children and Adolescents (EmQue-CA)[48]		×		
	Children's Reading the Mind in the Eyes Task (RME)[47]		×		
Literacy & Numeracy	National Assessment Program – Literacy and Numeracy (NAPLAN)	School Grade Years 3,5,7,9			
	School Grade Year			×	×
Anxiety and/or Depression					
	Spence Children's Anxiety Scale (SCAS)[49]	×	×	×	×
	Short Moods and Feelings Questionnaire (SMFQ)[50]	×	×	×	×
	Somatic and Psychological Health Report (SPHERE-21)[51,52]	×	×		
Emotional and Social Behaviours					
	Short UPPS-P Impulsive Behaviours Scale in Children (IBS)[53]		×		
	Strength and Difficulties Questionnaire (SDQ)[54]			×	×
	Australian-adapted Hierarchical Personality Inventory for Children (HiPIC-A)[55]				×
	Autism Spectrum Quotient – 10 items (AQ-10) (Adolescent)[56]				×
	Children's Response Styles Questionnaire – rumination subscale (CRSQ)[57]	×	×		
	Children's Attributional Style Questionnaire – Revised (CASQ-R)[58]	×	×		
	Early Adolescent Temperament Questionnaire – Revised (EATQ-R)[59]			×	×
Social Support and Family Functioning					
	Multidimensional Scale of Perceived Social Support (MSPSS)[60]	×	×		
	Alabama Parenting Questionnaire (APQ; Parental Monitoring session 1 only)[61]			×	×
	McMaster Family Assessment Device (FAD)[62]			×	×
Stress					
	Daily Life Stressors Scale (DLSS)[63]	×	×		
	Gatehouse Bullying Scale (GBS)[64,65]	×	×		
	Child Life Events Questionnaire (CLEQ)[66]				×
	Prenatal Stress Exposure Scale (PSES) – adapted from[67]			×	
	Parental Stress Scale (PaSS)[68]			×	×
	List of Threatening Experiences (LTE)[69]			×	×
Sleep and Physical Health					
	Pediatric Daytime Sleepiness Scale (PDSS)[70]	×	×		
	Sleep Disturbances Scale for Children (SDSC)[71]				×
	Sleep behaviours across early childhood – items adapted from Generation R[72,73]				×
	Sleep/wake times	×	×		
	Chronotype		×		×
	Shared or own bedroom				×
	Actigraphy (2-week period following other assessments – subsample only*)	×	×		
	Activity (e.g., sport participation), screen/device usage		×		
	Body measurements (height, weight, blood pressure, pulse)	×	×		
Early Life and Family Demographics					
	Twin pregnancy and birth (including breastfeeding, maternal smoking/drinking)			×	
	Early twin development (including gestation age, birth weight, length, head circumference)			×	
	Socioeconomic Status (Socio-Economic Index for Areas [SEIFA][74], Australian Socioeconomic Index [AUSE106][25], MacArthur Scale of Subjective Social Status[75])			×	×†
	Parental education, maternal/paternal age at twin birth, maternal/paternal height/weight, family medical history, ancestry (adapted from the 2016 Australian Census)			×	
Dietary Behaviour					
	Australian Child and Adolescent Eating Survey (AES)[76]			×	
COVID-19 Pandemic Specific Assessments (subsample only)					
	Experiences and worries‡		×		×
	Active and Passive Social Media Use (APSMU)[77,78]		×		
	UCLA Brief COVID-19 Screen for Child/Adolescent PTSD (PTSD)[79]		×		
	Perceived Stress Scale (PSS)[80]				×
	Brief Resilience Scale (BRS)[81]		×		×

Table 3. QTAB Non-Imaging Phenotypes (domain and measurement source).

*The actigraphy sample was largely restricted (*N* session 1 = 188, session 2 = 134) due to the availability of actigraphy watches rather than a refusal to participate.

†SEIFA only.

‡ Adapted from multiple COVID-19 survey sources, including the Swinburne University of Technology (Melbourne) and the NIH Office of Behavioral and Social Sciences Research (OBSSR).

****

Stress

We included structured assessments to capture *perceived stress* (Daily Life Stress Scale [DLSS] [63]) and *bullying* (Gatehouse Bullying Scale [GBS] [64,65], capturing both overt and covert types of victimisation), supplemented with parental report of *stressful*

life events (Child Life Events Questionnaire [CLEQ] [66] and List of Threatening Experiences [LTE] [69], i.e., stressful life events over the last year) and measures of *prenatal* (Prenatal Stress Exposure Scale [PSES] adapted from[67]) and *parental* (Parental Stress Scale [PaSS] [68]) stress. Adolescence is a period of increased vulnerability to stressors and to stress-related psychopathology, including anxiety and depression – relationships that may be mediated through the effects of stress on the developing adolescent brain[3,92].

Sleep and Physical Health

During each session, we assessed *sleepiness* using the Pediatric Daytime Sleepiness Scale (PDSS) [70], which has robust psychometric properties in adolescents aged 11 to 15 years and is associated with academic achievement and mood. Participants completed a sleep diary, recording their sleep and wake times (i.e., sleep duration) on weekdays and weekends. In addition, we obtained several anthropometric measures (e.g., height, weight). In session 2, we also asked whether they were a morning or evening person (i.e., to provide a proxy measure of chronotype) and assessed general physical activity and device usage (questions from the literature[93-95] and the Youth Risk Behavior Survey, United States, 2019). Further, the parent provided information about sleep disturbances (Sleep Disturbances Scale for Children [SDSC] [71] and sleep behaviours across early childhood items, as assessed in the Generation R study[72,73]).

In addition, at the end of each session, a sub-sample of participants were given a wrist-mounted accelerometry recording device (GENEActiv, Activinsights, Kimbolton, UK) to wear for two weeks on the wrist of their non-dominant hand. Accelerometry devices detect motion and estimate sleep from decreased movement. Participants completed a sleep diary every morning to consolidate the accelerometry data, providing information on bedtimes, wake times, and restorative sleep[96]. On day 15, the devices were returned via postal service and the data was downloaded using GENEActiv software. Currently, shared actigraphy data is restricted to measures of mean sleep onset, wake time, duration, midpoint, and restorative sleep (session 1 only). The raw actigraphy data for sessions 1 and 2 are held by Kathleen Merikangas, National Institute of Mental Health, and Ian Hickie, University of Sydney, Australia, and may be able to be shared by contacting them.

Early Life and Demographic Factors

Perinatal and postnatal information (e.g., maternal smoking and drinking; gestation age, birth weight, breastfeeding) were provided by the mother. Questions were adapted from earlier work[97,98] and other freely available sources (The World Health Organisation Global Adult Tobacco Survey [GATS, 2nd Edition]). Family demographic information was also provided by the parent attending session 1. These included self-report of ancestry (based on questions included in the 2016 Australian Census) and socioeconomic indexes from Census neighbourhood classifications (Socio-Economic Index for Areas [SEIFA] [74]), occupation or education (Australian Socioeconomic Index [AUSE106] [25]), and social status (The MacArthur Scale of Subjective Social Status[75]), as well as age at twin birth, and current height and weight.

Dietary Behaviour

In session 1 the attending parent completed the Australian Child and Adolescent Eating Survey (AES) [76] for each twin. This online food frequency questionnaire assesses the dietary intake of children and adolescents aged 2 to 17 years. It includes the frequency of consumption of 120 common foods, use of supplements, as well as eating and behaviours. Although not suitable for estimating absolute intakes, individuals can be classified into quintiles of intake for categories including total energy, protein, carbohydrate, sugars, fibre, and vitamin and mineral intake[76]. For a sub-sample of QTAB participants, these dietary intake assessments have been analysed and supplement gut microbiome measures obtained from stool samples[99].

COVID-19 Interruption to Session 2 Data Collection and Pandemic Specific Assessments

Session 2 data collection, which began in November 2019, was paused on the 25th of March 2020 as Brisbane entered a nation-wide shutdown (including the closure of schools). In late June 2020, session 2 data collection resumed, and continued until the conclusion of the QTAB study in January 2021. 47 families completed session 2 prior to the COVID19 interruption, and 105 families following the resumption of data collection for session 2 (see variable *lockdown_ses02* in the COVID-19 dataset (*10_covid-19.tsv*). There was no change to the study protocol or collection methodology post-interruption (i.e., twin and parent assessments were completed in-person at the study centre). Due to the COVID-19 interruption, the interval between session 1 and 2 was extended to a maximum of 30 months. Further, to maximise participation while constrained by funding deadlines, the minimum interval was reduced to 13 months. Preference was given to "overdue" families following the resumption of data collection; otherwise, there was no specific criteria when approaching families to return for their session 2 visit.

Pandemic specific assessments via questionnaire were acquired between August 2020 and January 2021. During this time, Brisbane schools were open and restrictions eased; however, various non-pharmaceutical interventions remained[100]. All families who participated in session 1 were contacted via email and invited to complete the questionnaire online at home (administered using Qualtrics software [Provo, UT, USA, https://www.qualtrics.com]). Those families who returned for session 2 but had not yet completed the COVID-19 questionnaire at home, completed the survey at the end of their session 2 visit. In total, we received responses from at least one family member (i.e., twin or parent) for 108 families, of which approximately half completed the questionnaire at home.

We adapted questions from multiple COVID-19 survey sources, including the Swinburne University of Technology (Melbourne) and the NIH Office of Behavioral and Social Sciences Research (OBSSR). We also assessed post-traumatic stress (Child PTSD Symptom Scale for children 8-18 years[101]), *resilience* (Brief Resilience Scale[81]; Posttraumatic Growth Inventory X[102]) and active and passive social media use (Multi-dimensional Scale of Facebook Use[78]; modified for all types of social media[77]). Resilience is implicated in mediating the development of mental illness following trauma[103], while type of social media use (i.e., active "connection promoting" versus passive "non-connection promoting") has been shown to mediate symptoms of anxiety and depressed mood in adolescents[77], and active usage may be protective against negative consequences of social distancing. Further, a parent (usually the mother) reported on pandemic-related concerns and the impact on their work and home situation, personal finances, general wellbeing, and mental health – all factors that may mediate or moderate adolescent responses to the pandemic.

Biological Samples

While we collected several biological samples, most samples have not yet been processed due to budgetary constraints; however, collection statistics (i.e., yes/no) are provided for each biological sample (see *Data Records*).

Salivary Sex-Steroid Collection

Participants collected a saliva sample upon minutes of wakening on the day of their visit. A saliva kit, including a collection tube, icepack, and step-by-step written instructions, was mailed in advance of the study session. The procedure was explained to the parent on the phone when booking their session visit. Participants were instructed to generate some saliva in their mouth and then slowly drool it into the tube until the desired amount was collected (i.e., 2 ml). Then, they placed it on the icepacks for transport. Upon arrival at the QTAB session, saliva samples were immediately

placed on fresh ice and transferred to a -80C freezer. 99% of participants provided a saliva sample at session 1 and 100% at session 2. The long-term aim was to process the samples at the Stress Physiology Investigative Team (SPIT) laboratory at Iowa State University. Estradiol and testosterone concentrations (pg/mL) would be indexed using the Salimetrics Salivary Testosterone Enzyme Immunoassay ELISA kit.

Hair Cortisol
At sessions 1 and 2, a research assistant used scissors to cut 10–50 mg of hair from the posterior vertex region of the scalp (1 to 3 cm – enough to assess cortisol for the previous 3 months). Hair specimens have been stored at room temperature and have not yet been processed. Hair cortisol has recently emerged as a promising biomarker for long-term retrospective HPA activation[104], with extremes of cortisol concentration levels (i.e., both lowest and highest) predicting depressive symptoms in adolescents[105] and longitudinal change in concentration levels predicting later social anxiety[106]. Puberty is a period of HPA-axis plasticity, and the effects of stress on cortisol regulation may depend on developmental stage/pubertal maturation. A longer-term aim of QTAB was to track how cortisol functioning impacts the brain regions processing emotion and whether HPA reactivity interacts with pubertal stage and subsequent associations with depression. Hair cortisol data are held by associate researchers exclusively until the 30th of June, 2027. After this date, data will be shared via an appropriate repository; please contact Liza van Eijk and Zoltan Sarnyai, James Cook University, after the exclusivity period ends.

Blood
Genomic DNA was extracted from a blood sample provided at session 1 using standard procedures. DNA samples for available twin participants were genotyped using the Infinium Global Screening Array-24 v3.0 BeadChip (Illumina, California, USA). Genotyping data is expected to be available in 2023. These data will be stored in a dedicated genotype repository under different participant identifiers. Researchers will require access approval to the QTAB non-imaging phenotypes dataset (*see Data Records - Non-Imaging Phenotype Data*) before participant identifiers can be linked between the imaging/phenotypic and genotypic datasets.

Data Records

Imaging Data
The QTAB imaging dataset is publicly available through the OpenNeuro data sharing platform[107] (accession number ds004146; please use the latest version as updates may exist). The dataset is organised as per the Brain Imaging Data Structure (BIDS) specification v1.6.0[108]. BIDS specifies a hierarchical organisational format, with participant data stored under sub-folders denoting session (ses-01, ses-02) and then image modality: anat (structural), dwi (diffusion), func (functional), perf (asl), swi (susceptibility). An overview of the data record is available in Supplementary Table 1. Before sharing, all personally identifiable information was removed from the dataset, including facial features from the 3D whole-brain images (T1-weighted, T2-weighted, and FLAIR). We have made all data available, regardless of data quality (see *Technical Validation*).

Basic demographic data (age [rounded down], sex, handedness, family_id [to denote non-independent participants]) is provided at the top-level of the dataset directory in participants.tsv (variables and properties described in participants.json). Each scan is stored as a compressed NifTI file (.nii.gz) with an accompanying sidecar JSON file (.json) describing scan acquisition parameters. In addition, for the diffusion scans, gradient orientation information is provided in **_dwi.bvec and *_dwi.bval,* and for the asl scans, volume type information (i.e., m0scan, control, label) is provided in **_aslcontext.tsv*. For the emotional conflict and movie-watching (Partly Cloudy) tasks, event details (e.g., facial emotion responses for the emotional conflict task, condition

timings for the Partly Cloudy task) are provided in *_task-emotionalconflict_events.tsv* and *_task-partlycloudy_events.tsv* respectively (variables and properties described in accompanying *_task-emotionalconflict_events.json* and *_task-partlycloudy_events.json*). The defacing masks used to deface the 3D whole-brain images are provided alongside participant anatomical data in *_inv-2_MP2RAGE_defacemask.nii.gz*. Lastly, the *derivatives* folder at the top-level of the dataset directory contains quality checking metrics for imaging data (see *Technical Validation*) and MP2RAGE UNIT1 images with background noise removed (see *Usage Notes - MP2RAGE Acquisitions*).

Non-Imaging Phenotype Data

The QTAB non-imaging phenotypes dataset is available through the Zenodo data sharing platform as a restricted access dataset[109]. Access requests can be submitted through the Zenodo platform. Please note, applicants are required to complete a data usage agreement (DUA)[110] as part of the application process. To protect and assure the confidentiality and privacy of QTAB participants, all Recipients who are granted access to these data are expected to adhere to all terms of use outlined in the agreement. Access to the non-imaging phenotypic dataset will be granted to qualified investigators for appropriate use, as determined by authors Lachlan Strike and Katie McMahon.

The QTAB non-imaging phenotypes dataset contains restricted demographic data (i.e., age in months, zygosity, zygosity source, multiple birth status, birth order), as well as sex, handedness, and family_id, is provided in *participants_restricted.tsv* (variables and properties described in *participants_restricted.json*). Multiple assessments and scales of a common domain are grouped as per the phenotypic domains detailed in Table 3 (i.e., *Puberty, Cognition, Anxiety and/or Depression, Emotional and Social Behaviours, Social Support and Family Functioning, Stress, Sleep and Physical Health, Early Life and Family Demographics, Dietary Behaviour, COVID-19 Pandemic Specific Assessments*), and separated by session. Total/sum scores and individual items (where possible) are stored in a series of .tsv files, with variables and properties described in an accompanying JSON sidecar. Collection statistics (i.e., yes/no) for biological samples (blood, hair, saliva, urine, stool, baby tooth) are additionally provided. An overview of the non-imaging data record is available in Supplementary Table 2, with notations identifying the key measurement domain. Before sharing these phenotypes, we removed all participant identifiable information, including occupation, postcode, and special categories of personable data (i.e., free-response text).

Technical Validation

Imaging Data

3D whole-brain (T1-weighted, T2-weighted, and FLAIR) and TSE scans were visually checked and rated by one author (LTS). Scans were rated using a three-category scale (pass, warn, fail) and scan quality ratings are available in the *derivatives/visual_qc* folder. There was a higher percentage of warn and fail ratings for TSE, T2-weighted and FLAIR scans than T1-weighted scans (Fig. 2a**)**, likely due to increased head movement associated with the acquisition of the TSE, T2-weighted, and FLAIR scans towards the end of the imaging session. The T1-weighted, T2-weighted, and FLAIR images were visually checked to ensure that facial features were successfully removed.

We used the MRtrix3 script *dwipreproc*[111] (which implements the FSL tool *eddy*[112,113]) to calculate volume-to-volume motion estimates (Fig. 2b**)** and the percentage of detected outlier slices (i.e., slices affected by severe signal dropout; Fig. 2c); estimates available in the *derivatives/mrtrix3* folder. The median across participants of average absolute motion was 0.77 mm and 0.76 mm at sessions 1 and 2, respectively. This finding is comparable with the same metric reported in a subset of

the adult HCP[114] (median 0.83 mm). The median percentage of total outlier slices was 1.18% and 1.01% at sessions 1 and 2. The same metric was 1.89% and 0.39% in the developing/neonatal and adult HCP datasets[114], respectively.

Image quality metrics (IQMs) were calculated for task and resting-state fMRI scans using MRIQC[115]. Framewise displacement (FD) head motion IQMs are displayed in Fig. 3d, and all IQMs are available in the *derivatives/mriqc* folder. The median across participants of average FD ranged from 0.16mm (Partly Cloudy) to 0.21mm (session 1 Rest). This finding is comparable with the same metric reported in a dataset of 3-12-year-old children[35] (Partly Cloudy task, median = 0.29 mm) and a dataset of 8-17-year-old children[116] (lexical processing tasks, median across all tasks and sessions = 0.17 mm). The ASL and SWI scans provided have not yet undergone quality checking or pre-processing. However, all scans were checked for apparent artefacts during the scanning session, with affected scans re-acquired (time permitting).

Non-Imaging Phenotypes

Fig. 3 focuses on the design of QTAB and the non-imaging phenotypes. We chose an accelerated longitudinal design (Fig. 3a) so that in the first three years of the study, we worked to recruit every willing participant who was a twin between the ages of 9-14 years and who lived close to the study centre at the University of Queensland, Brisbane. We were successful in recruiting 422 twins across this age range. The long-term aim was to follow them prospectively and gather data on at least one further time point. For the 304 participants returning for session 2, the inter-session interval ranged from 1.1 to 2.5 years (M=1.7±0.3). The interval was partly influenced by pauses in data collection due to the pandemic, availability due to orthodontic treatment, and 5 years of funding. With an accelerated design[117], as the participants age into new categories, they contribute data in every cell, i.e., the sample sizes get larger and larger as the adolescents age into these categories.

The advance of pubertal development can be seen in Fig. 3b. At session 1, more than half of the sample were classified as pre- or early-pubertal. In contrast, at session 2, approximately three-quarters of the sample had progressed to mid- or late-pubertal status, with a small number classified as post-pubertal. Developmental advances in processing speed (Fig. 3c) were also evident throughout the inter-session interval. The NIHTB-CB processing speed task records the number of correctly answered items in 85 seconds. On average, participants correctly answered an additional 8 items at session 2 compared to session 1. A trend for cross-sectional age-related increases in processing speed was also found, with older participants, in general, being able to answer more items within the given timeframe correctly.

We calculated a measure of general cognitive ability using the NIHTB-CB Total Composite Score. Age-corrected standard scores for QTAB are shown in Fig. 3d. They are very close to the US national average scores (115 and 85 indicate performance 1 SD above and below the national average of 100), with a QTAB cohort mean score of 103.5 and SD of 17.2. In addition, 7.9% of the QTAB cohort had scores of 130 or greater (top 2% based on normative NIHTB-CB data), indicating that the cohort is oversampled for high cognitive performance, and 1.4% had scores of 70 or below (bottom 2% nationally of US scores), which suggests very low cognitive functioning and may be indicative of difficulties in school or general functioning. We also found that cognitive ability is more similar between QTAB identical co-twins (MZ twins) than non-identical co-twins (DZ twins), as expected for a trait that is influenced by genetic inheritance[118].

Distributions of symptom levels for depression and anxiety in QTAB are shown in Fig. 3f and Fig. 3g, respectively. Depression symptoms were obtained from the SMFQ, for which a cut-off of 8 has been suggested as a screen for depression in children aged 8-16 years[50]. Approximately 17% of the QTAB cohort scored 8 or higher (Fig. 3f). This is consistent with national surveys of child and adolescent health in Australia (e.g., 20% of adolescents aged 11-17 years reported experiencing high levels

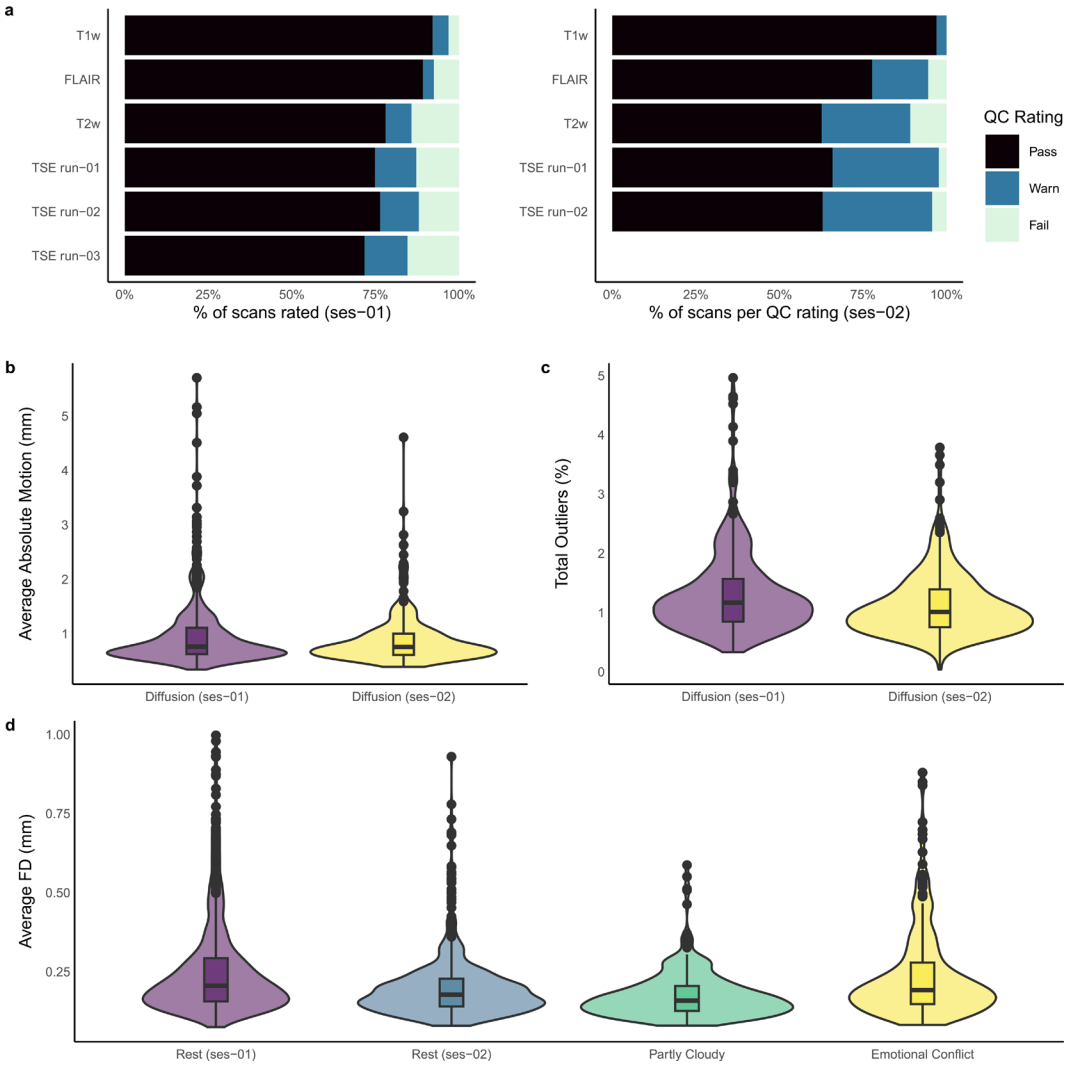

Fig. 2 MRI image quality measures. (**a**) Visual inspection quality control (QC) ratings (pass, warn, fail) for the anatomical scans collected at session 1 and 2 (ses-01, ses-02). A *pass* rating denotes images show clear grey/white matter contrast and are free of motion or ringing artefacts. A *warn* rating indicates images show slight ringing but otherwise good grey/white matter contrast. A *fail* rating denotes images are severely impacted by motion or susceptibility artefacts with cortical and subcortical structures challenging to delineate clearly. (**b**) Violin plot showing the distribution of average absolute motion in mm (i.e., displacement relative to the first volume; lower values are better) for diffusion scans at session 1 and 2. The lower and upper hinges of the box plot represent the first quartile and third quartiles, respectively, and the horizontal line within the box represents the median. Outliers (1.5 times the interquartile range) are represented as individual data points. The median across participants of average absolute motion was similar between sessions 1 and 2 (0.77 mm and 0.76 mm, respectively). (**c**) Violin plot of the percentage of total slices classified as outliers (i.e., slices affected by severe signal dropout; lower values are better) for diffusion scans at sessions 1 and 2. The median percentage of total outlier slices was greater at session 1 (1.18%) than session 2 (1.01%). (**d**) Violin plot showing the distribution of average Framewise Displacement (FD) head motion in mm for resting-state and task fMRI scans (lower values are better); participants with mean FD greater than 1 mm not shown (rest session 1: 15 participants, rest session 2: 1 participant, emotional conflict: 4 participants). The median across participants of average FD was lowest for the Partly Cloudy movie-watching scan (0.16mm) compared to the rest and emotional conflict task scans (0.18-0.21 mm).

****

of psychological distress in the Young Minds Matter 2013-14 Survey[119]). SCAS sum score (Generalized anxiety) was higher in girls than boys and younger compared to older participants, consistent with other studies[49,120]. Of the 54 individuals aged 13-14 years, the mean score (M=13.3) is consistent with that reported for a community sample of 875 adolescents aged 13-14 years (M=13.5) [49] (sum score for males M=18.85±13.07 [versus 18.59±9.89 for QTAB]; females M=25.08±13.37 [versus 24.74±12.76 for QTAB]; combined M=21.72±13.56 [versus 21.67±11.73 for QTAB]). Fewer symptoms were found for those aged 9-12 compared to same-age participants in Spence et al. [120] (M=29.7 versus 31.3 at age 9; M=26.4 versus 31.7 at age 10; M=24.0 versus 31.3 at age 11; M=22.0 versus 28.7 at age 12). Fig. 3h shows the QTAB distribution for daytime sleepiness, as assessed using the PDSS. Excessive daytime sleepiness has been associated with poor stress management and higher levels of depressive mood in adolescents aged 14-19 years[121]. A cut-off above 17 has previously been used to identify excessive sleepiness in 618 children aged 10 to 12 years[122], identifying 18% of the sample. With the same cut-off threshold, 17% of the QTAB cohort would be classified as having excessive daytime sleepiness.

Usage Notes

Analysis of Twin Data

For analyses of genetic (co)variance in the QTAB dataset, we suggest investigators familiarise themselves with the classic twin model[123]. Investigators not interested in genetic (co)variance should nevertheless consider the correlated nature of twin data (i.e., the non-independence of participants) as it may violate statistical test assumptions[124]. Mixed models, which use random effects to model the correlation among twin pairs[125], and structural equation modelling using the classic twin design[18], are widely used approaches in controlling familial relatedness. Example code is provided online (https://github.com/QTAB-STUDY/twin-data-models).

MP2RAGE Acquisitions

The amplified background noise in the T1w MP2RAGE uniform image (*_UNIT1) can cause registration and segmentation issues. One approach is to remove background noise from the MP2RAGE uniform image using the inversion time images[126,127] (see *derivatives/UNIT1_denoised* for an implementation using AFNI[109]). Another technique is to input brain-extracted (i.e., skull-stripped) MP2RAGE uniform images to automated processing pipelines (e.g., FreeSurfer, fMRIPrep). We found that creating a brain mask based on the second inversion time image (*_inv-2_MP2RAGE) and applying this mask to the MP2RAGE uniform image resulted in successful brain extractions (https://github.com/QTAB-STUDY/pre-processing).

Geometric Distortion Correction

Diffusion and rs-fMRI scans were acquired using reversed-phase encoding directions to correct geometric distortions in the images. We recommend using the FSL tools *topup* and *eddy* to correct for distortions and movement in the diffusion scans and *topup* and *applytop* to correct the rs-fMRI scans. Similarly, we recommend using the reversed-phase encoding field maps and the FSL tools *topup* and FEAT to correct the task fMRI scans. Processing pipelines implementing these tools are available online (https://github.com/QTAB-STUDY/pre-processing). Preprocessing pipelines such as fMRIPrep[128] may also be useful for distortion correction.

Movie-Watching Task

The video stimulus for the movie-watching task was transcoded from a PAL format DVD, resulting in a slightly faster running video than used in prior studies[34,35]. Specifically, the Partly Cloudy video used in the QTAB study was 5:31 minutes:seconds long (5:01 without end credits), compared to 5:45 minutes (5:14 without end credits) in non-PAL format. Event time codes provided in the QTAB dataset

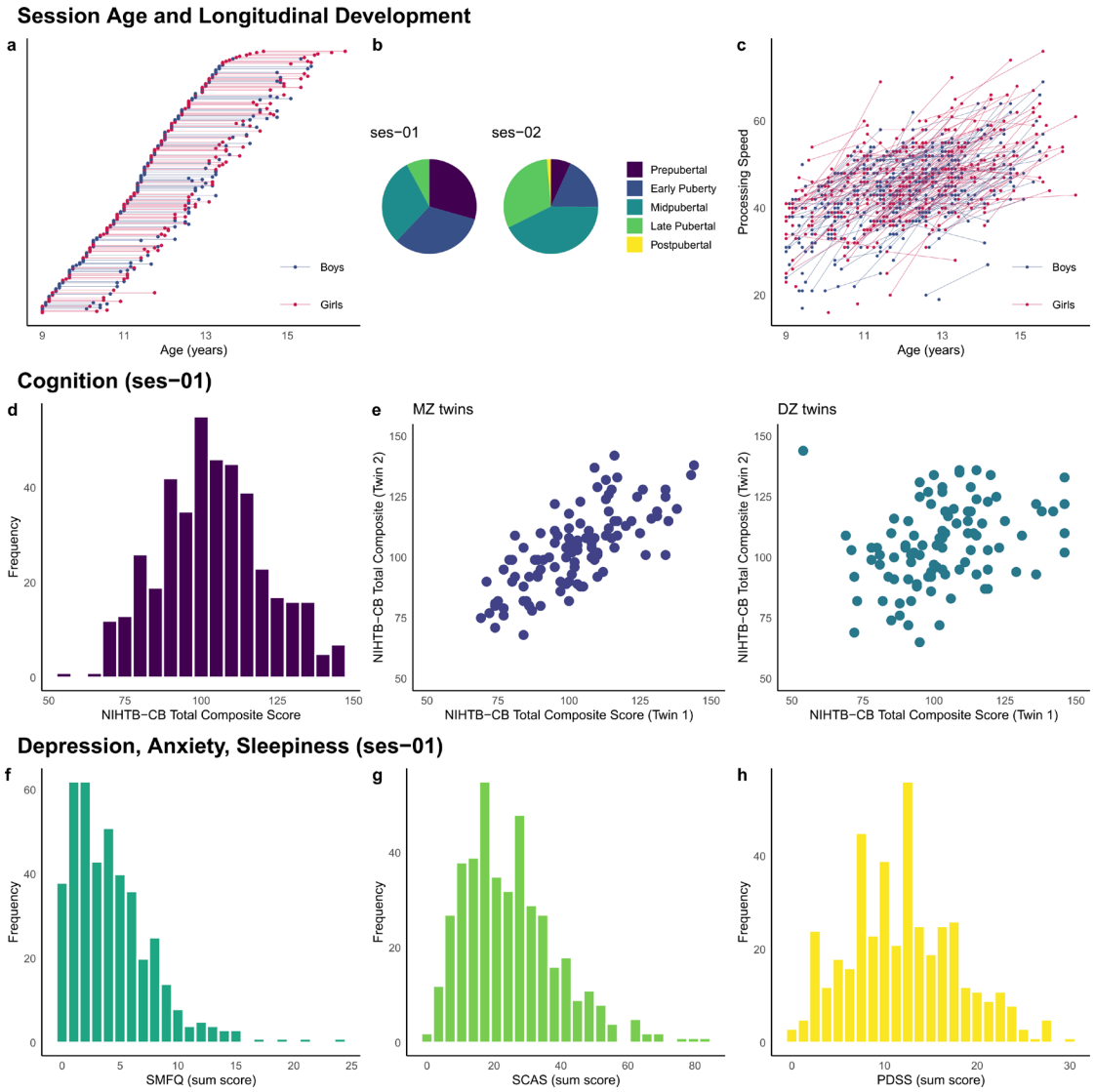

Fig. 3 Examples supporting the technical validation of the QTAB non-imaging phenotypes. Within the 5-year funding period, the accelerated longitudinal design maximised the benefits of cross-sectional and longitudinal data collection (**a**), and developmental changes in pubertal status (**b**) and processing speed (**c**) occurred between sessions 1 and 2. The QTAB cohort was highly representative of community norms in cognitive ability (**d**), with the expected genetic relationship, i.e., identical twins were more alike than non-identical twins (**e**). In addition, the tools chosen to measure depression (**f**) and anxiety (**g**) were sensitive enough to capture individual differences and identify at-risk individuals. Variability in a depression risk factor, i.e., daytime sleepiness (**h**), was consistent with other community samples.

****

(i.e., onset and duration for mental, pain, social, and control conditions; *_*task-emotionalconflict_events.tsv*) correspond to the timings of a prior study[34,129], converted to match the QTAB video stimulus. However, we recommend researchers obtain a PAL format DVD version of the film (or equivalent) to verify these condition timings (Partly Cloudy is included on physical media versions of the Disney/Pixar film Up [2009]).

Miscellaneous Imaging

Multiple TSE scans were collected to average the TSE runs to improve image quality; see Shaw et al. [130] for implementation of TSE alignment and averaging using the QTAB dataset. Two participants (twin pair sub-0109 and sub-0113) required an increased

number of slices for their session 1 structural scans (T1w, T2w, FLAIR) for adequate brain coverage. Three participants (twin pair sub-0200 and sub-0419, sub-0373) have a reduced number of slices in their session 1 structural scans (T1w, T2w, FLAIR) due to operator error. Four participants have task fMRI but not corresponding field map scans (sub-1207, sub-7877, sub-8742, sub-9549). One participant has an incomplete T1w acquisition at session 2 (sub-0271, missing UNIT1, inv-1 images, but has UNIT1_denoised, inv-2 images). During session 1, the T1w MP2RAGE sequence name changed (MP2RAGE_wip900C_VE11C to MP2RAGE_wip900D_VE11C); however, there was no change to the sequence parameters.

Non-Imaging Phenotypes

Reversed scored questionnaire items have been re-coded in the non-imaging phenotypes dataset. Parent responses for several scales (i.e., APQ, FAD, LTE, PaSS, PSES, PSS, BRS, and COVID-19 Experiences and Worries) are for the family (i.e., scores are the same for co-twins within a family). Australia is a multicultural country, as reflected in the QTAB ancestry measure. This measure reflects self-perceived group identification and may differ from a person's genetic ancestry, as obtained from genotyping. We also note that 29 participants have higher puberty scale scores at session 1 than session 2 (see the *PDS_scores_ses01_greater_ses02* variable in *01_puberty-ses01.tsv*). We believe this disparity reflects improved self-report puberty measurement with increased age. We suggest replacing session 01 *PDS_scale_score*, *PDS_category_score*, *Gondal_score*, and *Adrenal_score* variables with the corresponding session 02 variables. Due to copyright restrictions[131], scale items from questionnaires are not included in the non-imaging phenotypes dataset. However, we provide detailed instructions for linking item variables to the published questionnaire items (see *00_non_imaging_phenotypes_overview.pdf* in the non-imaging phenotypes dataset; also provided in Supplementary Table 3). Where necessary, we contacted scale/questionnaire authors to obtain permission to use their measure and share the data collected.

Code Availability

DICOM format MRI data was converted to a BIDS compatible dataset using HeuDiConv[132] (v0.9.0; https://github.com/nipy/heudiconv). Facial features were removed from structural scans using BIDSonym[133] (v0.0.5; https://github.com/peerherholz/bidsonym) and the tools *flirt* and *fslmaths* in FSL[134] (v5.0.1; https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/). Code used in data organisation and defacing is available online (https://github.com/QTAB-STUDY/dicom-to-bids). Head movement and outlier metrics for diffusion scans were calculated using the tool *dwipreproc* in MRtrix3[111] (v3.0_RC3; https://www.mrtrix.org). Task and resting-state fMRI quality checking metrics were calculated using MRIQC[115] (v0.16.1; https://github.com/nipreps/mriqc).

Acknowledgements

We are forever grateful to the twins and their families for their willingness to participate in our studies. We thank Liza van Eijk, Victoria O'Callaghan, Islay Davies, Ethan Campi, Kimberley Huang, Eleanor Roga, Michael Day, Aiman Al-Najjar, Zoie Nott, Tom Shaw, Nicole Atcheson, and Sarah Daniel for data acquisition. We thank Naomi Wray for funding the collection of metabolic samples, including detailed dietary data, Ian Hickie and Kathleen Merikangas for funding and support of actigraphy data, and Sarah Medland and ENIGMA GWAS for funding genotyping. Special thanks to Anjali Henders, Leanne Wallace, Lorelle Nunn and the many laboratory assistants at the Human Studies Unit (part of the Program in Complex Trait Genomics based at the Institute of Molecular Bioscience, University of Queensland) for the processing and storage of biological samples. Thanks also to Julie Henry for helpful discussion of social cognition measures. The QTAB project was funded by the National Health and

Medical Research Council (NHMRC), Australia (Project Grant ID: 1078756 to MJW), the Queensland Brain Institute, University of Queensland, and with the assistance of resources from the Centre for Advanced Imaging and the Queensland Cyber Infrastructure Foundation, University of Queensland. We acknowledge the Queensland Twin Registry (QTwin) (https://www.qimrberghofer.edu.au/study/queensland-twin-registry-study) for generously sharing database information for recruitment. Recruitment was further facilitated through access to Twins Research Australia, a national resource supported by a Centre of Research Excellence Grant (ID: 1079102) from the NHMRC. Lastly, we thank the many researchers worldwide for providing access to their assessments.

All publications of findings using data from the QTAB project should consider including the following credit and references, which should be incorporated within the "Acknowledgements" of such publication: "This research has been conducted using the QTAB project resource[107,109]".

Author contributions

Lachlan T. Strike – data curation, visual quality checking, pre-processing and analysis, writing – original draft.
Narelle K. Hansell - designed the data acquisition protocol, data curation, writing – original draft.
Kai-Hsiang Chuang – designed the data acquisition protocol.
Jessica L. Miller – data collection, study management.
Greig I. de Zubicaray - designed the data acquisition protocol, writing - review & editing, funding acquisition.
Paul M. Thompson - designed the data acquisition protocol, writing - review & editing, funding acquisition.
Katie L. McMahon - designed the data acquisition protocol, writing - review & editing, funding acquisition.
Margaret J. Wright - designed the data acquisition protocol, writing - review & editing, funding acquisition.

Competing interests

The authors declare no competing interests. PMT received a research grant from Biogen, Inc., for work unrelated to the current manuscript.

References

1 Lichenstein, S. D., Verstynen, T. & Forbes, E. E. Adolescent brain development and depression: A case for the importance of connectivity of the anterior cingulate cortex. *Neurosci Biobehav Rev* **70**, 271-287, https://doi.org/10.1016/j.neubiorev.2016.07.024 (2016).

2 Jones, S. A., Morales, A. M., Lavine, J. B. & Nagel, B. J. Convergent neurobiological predictors of emergent psychopathology during adolescence. *Birth Defects Res* **109**, 1613-1622, https://doi.org/10.1002/bdr2.1176 (2017).

3 Romeo, R. D. The impact of stress on the structure of the adolescent brain: Implications for adolescent mental health. *Brain Res* **1654**, 185-191, https://doi.org/10.1016/j.brainres.2016.03.021 (2017).

4 Kessler, R. C. *et al.* Lifetime prevalence of age-of-onset distributions of DMS-IF in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* **62**, 593-602 (2005).

5 Alegria, M., NeMoyer, A., Falgas Bague, I., Wang, Y. & Alvarez, K. Social Determinants of Mental Health: Where We Are and Where We Need to Go. *Curr Psychiatry Rep* **20**, 95, https://doi.org/10.1007/s11920-018-0969-9 (2018).

6 Johnson, D., Dupuis, G., Piche, J., Clayborne, Z. & Colman, I. Adult mental health outcomes of adolescent depression: A systematic review. *Depress Anxiety* **35**, 700-716, https://doi.org/10.1002/da.22777 (2018).

7 Clayborne, Z. M., Varin, M. & Colman, I. Systematic Review and Meta-Analysis: Adolescent Depression and Long-Term Psychosocial Outcomes. *J Am Acad Child Adolesc Psychiatry* **58**, 72-79, https://doi.org/10.1016/j.jaac.2018.07.896 (2019).

8 Casey, B. J., Heller, A. S., Gee, D. G. & Cohen, A. O. Development of the emotional brain. *Neurosci Lett* **693**, 29-34, https://doi.org/10.1016/j.neulet.2017.11.055 (2019).

9 Yurgelun-Todd, D. Emotional and cognitive changes during adolescence. *Curr Opin Neurobiol* **17**, 251-257, https://doi.org/10.1016/j.conb.2007.03.009 (2007).

10 Short, M. A. & Chee, M. W. L. Adolescent sleep restriction effects on cognition and mood. *Prog Brain Res* **246**, 55-71, https://doi.org/10.1016/bs.pbr.2019.02.008 (2019).

11 Pfeifer, J. H. & Allen, N. B. Puberty Initiates Cascading Relationships Between Neurodevelopmental, Social, and Internalizing Processes Across Adolescence. *Biol Psychiatry* **89**, 99-108, https://doi.org/10.1016/j.biopsych.2020.09.002 (2021).

12 Naninck, E. F., Lucassen, P. J. & Bakker, J. Sex differences in adolescent depression: do sex hormones determine vulnerability? *J Neuroendocrinol* **23**, 383-392, https://doi.org/10.1111/j.1365-2826.2011.02125.x (2011).

13 Deardorff, J. *et al.* Girls' Pubertal Timing and Tempo and Mental Health: A Longitudinal Examination in an Ethnically Diverse Sample. *J Adolesc Health* **68**, 1197-1203, https://doi.org/10.1016/j.jadohealth.2021.01.020 (2021).

14 Abraham, Z. K. & Sher, L. Adolescent suicide as a global public health issue. *Int J Adolesc Med Health* **31**, https://doi.org/10.1515/ijamh-2017-0036 (2017).

15 Sinclair, D., Purves-Tyson, T. D., Allen, K. M. & Weickert, C. S. Impacts of stress and sex hormones on dopamine neurotransmission in the adolescent brain. *Psychopharmacology (Berl)* **231**, 1581-1599, https://doi.org/10.1007/s00213-013-3415-z (2014).

16 Lamblin, M., Murawski, C., Whittle, S. & Fornito, A. Social connectedness, mental health and the adolescent brain. *Neurosci Biobehav Rev* **80**, 57-68, https://doi.org/10.1016/j.neubiorev.2017.05.010 (2017).

17 Volkow, N. D. *et al.* The conception of the ABCD study: From substance use to a broad NIH collaboration. *Dev Cogn Neurosci* **32**, 4-7, https://doi.org/10.1016/j.dcn.2017.10.002 (2018).
18 Iacono, W. G. *et al.* The utility of twins in developmental cognitive neuroscience research: How twins strengthen the ABCD research design. *Dev Cogn Neurosci* **32**, 30-42, https://doi.org/10.1016/j.dcn.2017.09.001 (2018).
19 Somerville, L. H. *et al.* The Lifespan Human Connectome Project in Development: A large-scale study of brain connectivity development in 5-21 year olds. *Neuroimage* **183**, 456-468, https://doi.org/10.1016/j.neuroimage.2018.08.050 (2018).
20 Tobe, R. H. *et al.* A longitudinal resource for studying connectome development and its psychiatric associations during childhood. *Sci Data* **9**, 300, https://doi.org/10.1038/s41597-022-01329-y (2022).
21 Wierenga, L. M. *et al.* Greater male than female variability in regional brain structure across the lifespan. *Hum Brain Mapp* **43**, 470-499, https://doi.org/10.1002/hbm.25204 (2022).
22 Frangou, S. *et al.* Cortical thickness across the lifespan: Data from 17,075 healthy individuals aged 3-90 years. *Hum Brain Mapp* **43**, 431-451, https://doi.org/10.1002/hbm.25364 (2022).
23 Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7, https://doi.org/10.1186/s13742-015-0047-8 (2015).
24 Marees, A. T. *et al.* A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *Int J Methods Psychiatr Res* **27**, e1608, https://doi.org/10.1002/mpr.1608 (2018).
25 McMillan, J., Beavis, A. & Jones, F. W. The AUSE106: A new socioeconomic index for Australia. *Journal of Sociology* **45**, 123-149 (2009).
26 Li, Z., Umstad, M. P., Hilder, L., Xu, F. & Sullivan, E. A. Australian national birthweight percentiles by sex and gestational age for twins, 2001-2010. *BMC Pediatr* **15**, 148, https://doi.org/10.1186/s12887-015-0464-y (2015).
27 Wang, A. Y. *et al.* Neonatal outcomes among twins following assisted reproductive technology: an Australian population-based retrospective cohort study. *BMC Pregnancy Childbirth* **18**, 320, https://doi.org/10.1186/s12884-018-1949-0 (2018).
28 Marques, J. P. *et al.* MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *Neuroimage* **49**, 1271-1281, https://doi.org/10.1016/j.neuroimage.2009.10.002 (2010).
29 Van de Moortele, P. F. *et al.* T1 weighted brain images at 7 Tesla unbiased for Proton Density, T2* contrast and RF coil receive B1 sensitivity with simultaneous vessel visualization. *Neuroimage* **46**, 432-446, https://doi.org/10.1016/j.neuroimage.2009.02.009 (2009).
30 Vanderwal, T., Kelly, C., Eilbott, J., Mayes, L. C. & Castellanos, F. X. Inscapes: A movie paradigm to improve compliance in functional magnetic resonance imaging. *Neuroimage* **122**, 222-232, https://doi.org/10.1016/j.neuroimage.2015.07.069 (2015).
31 Marusak, H. A., Martin, K. R., Etkin, A. & Thomason, M. E. Childhood trauma exposure disrupts the automatic regulation of emotional processing. *Neuropsychopharmacology* **40**, 1250-1258, https://doi.org/10.1038/npp.2014.311 (2015).
32 Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R. & Hirsch, J. Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron* **51**, 871-882, https://doi.org/10.1016/j.neuron.2006.07.029 (2006).
33 Reher, K. & Sohn, P. Partly Cloudy [Motion Picture]. *USA: Pixar Animation Studios and Walt Disney Pictures* (2009).

34 Jacoby, N., Bruneau, E., Koster-Hale, J. & Saxe, R. Localizing Pain Matrix and Theory of Mind networks with both verbal and non-verbal stimuli. *Neuroimage* **126**, 39-48, https://doi.org/10.1016/j.neuroimage.2015.11.025 (2016).
35 Richardson, H., Lisandrelli, G., Riobueno-Naylor, A. & Saxe, R. Development of the social brain from age three to twelve years. *Nat Commun* **9**, 1027, https://doi.org/10.1038/s41467-018-03399-2 (2018).
36 Weintraub, S. *et al.* Cognition assessment using the NIH Toolbox. *Neurology* **80**, S54-64, https://doi.org/10.1212/WNL.0b013e3182872ded (2013).
37 Faria, E. R., Franceschini Sdo, C., Peluzio Mdo, C., Sant'Ana, L. F. & Priore, S. E. Methodological and ethical aspects of the sexual maturation assessment in adolescents. *Rev Paul Pediatr* **31**, 398-405, https://doi.org/10.1590/S0103-05822013000300019 (2013).
38 Morris, N. M. & Udry, J. R. Validation of a self-administered instrument to assess stage of adolescent development. *J Youth Adolesc* **9**, 271-280, https://doi.org/10.1007/BF02088471 (1980).
39 Marshall, W. A. & Tanner, J. M. Variations in pattern of pubertal changes in girls. *Arch Dis Child* **44**, 291-303, https://doi.org/10.1136/adc.44.235.291 (1969).
40 Marshall, W. A. & Tanner, J. M. Variations in the pattern of pubertal changes in boys. *Arch Dis Child* **45**, 13-23, https://doi.org/10.1136/adc.45.239.13 (1970).
41 Petersen, A. C., Crockett, L., Richards, M. & Boxer, A. A self-report measure of pubertal status: Reliability, validity, and initial norms. *J Youth Adolesc* **17**, 117-133, https://doi.org/10.1007/BF01537962 (1988).
42 Gur, R. C. *et al.* A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. *J Neurosci Methods* **187**, 254-262, https://doi.org/10.1016/j.jneumeth.2009.11.017 (2010).
43 Benton, A., Hannay, H. J. & Varney, N. R. Visual perception of line direction in patients with unilateral brain disease. *Neurology* **25**, 907-910, https://doi.org/10.1212/wnl.25.10.907 (1975).
44 Wechsler, D. *WISC-V: Wechsler Intelligence Scale for Children (5th ed.)*. (Psychological Corporation, 2014).
45 Delis, D. C., Kaplan, E. & Kramer, J. H. *Delis Kaplan Executive Function System: Technical Manual*. (TX: Psychological Corporation, 2001).
46 Talbot, K. D. & Kerns, K. A. Event- and time-triggered remembering: the impact of attention deficit hyperactivity disorder on prospective memory performance in children. *J Exp Child Psychol* **127**, 126-143, https://doi.org/10.1016/j.jecp.2014.02.011 (2014).
47 van der Meulen, A., Roerig, S., de Ruyter, D., van Lier, P. & Krabbendam, L. A Comparison of Children's Ability to Read Children's and Adults' Mental States in an Adaptation of the Reading the Mind in the Eyes Task. *Front Psychol* **8**, 594, https://doi.org/10.3389/fpsyg.2017.00594 (2017).
48 Overgaauw, S., Rieffe, C., Broekhof, E., Crone, E. A. & Guroglu, B. Assessing Empathy across Childhood and Adolescence: Validation of the Empathy Questionnaire for Children and Adolescents (EmQue-CA). *Front Psychol* **8**, 870, https://doi.org/10.3389/fpsyg.2017.00870 (2017).
49 Spence, S. H., Barrett, P. M. & Turner, C. M. Psychometric properties of the Spence Children's Anxiety Scale with young adolescents. *J Anxiety Disord* **17**, 605-625, https://doi.org/10.1016/s0887-6185(02)00236-0 (2003).
50 Angold, A., Costello, E. J. & Messer, S. C. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *Int J Methods Psychiatr Res* **5**, 237-249 (1995).
51 Hickie, I. B. *et al.* Development of a simple screening tool for common mental disorders in general practice. *Med J Aust* **175**, S10-17, https://doi.org/10.5694/j.1326-5377.2001.tb143784.x (2001).

52 Couvy-Duchesne, B., Davenport, T. A., Martin, N. G., Wright, M. J. & Hickie, I. B. Validation and psychometric properties of the Somatic and Psychological HEalth REport (SPHERE) in a young Australian-based population sample using non-parametric item response theory. *BMC Psychiatry* **17**, 279, https://doi.org/10.1186/s12888-017-1420-1 (2017).
53 Geurten, M., Catale, C., Gay, P., Deplus, S. & Billieux, J. Measuring Impulsivity in Children: Adaptation and Validation of a Short Version of the UPPS-P Impulsive Behaviors Scale in Children and Investigation of its Links With ADHD. *J Atten Disord* **25**, 105-114, https://doi.org/10.1177/1087054718775831 (2021).
54 Goodman, A. & Goodman, R. Strengths and difficulties questionnaire as a dimensional measure of child mental health. *J Am Acad Child Adolesc Psychiatry* **48**, 400-403, https://doi.org/10.1097/CHI.0b013e3181985068 (2009).
55 Watt, D., Hopkinson, L., Costello, S. & Roodenburg, J. Initial Validation and Refinement of the Hierarchical Inventory of Personality for Children in the Australian Context. *Australian Psychologist* **52**, 61-71, https://doi.org/10.1111/ap.12213 (2020).
56 Allison, C., Auyeung, B. & Baron-Cohen, S. Toward brief "Red Flags" for autism screening: The Short Autism Spectrum Quotient and the Short Quantitative Checklist for Autism in toddlers in 1,000 cases and 3,000 controls [corrected]. *J Am Acad Child Adolesc Psychiatry* **51**, 202-212 e207, https://doi.org/10.1016/j.jaac.2011.11.003 (2012).
57 Abela, J. R., Brozina, K. & Haigh, E. P. An examination of the response styles theory of depression in third- and seventh-grade children: a short-term longitudinal study. *J Abnorm Child Psychol* **30**, 515-527, https://doi.org/10.1023/a:1019873015594 (2002).
58 Thompson, M., Kaslow, N. J., Weiss, B. & Nolen-Hoeksema, S. Children's Attributional Style Questionnaire—Revised: Psychometric examination. *Psychological Assessment* **10**, 166-170, https://doi.org/10.1037/1040-3590.10.2.166 (1998).
59 Oldehinkel, A. J., Hartman, C. A., De Winter, A. F., Veenstra, R. & Ormel, J. Temperament profiles associated with internalizing and externalizing problems in preadolescence. *Dev Psychopathol* **16**, 421-440, https://doi.org/10.1017/s0954579404044591 (2004).
60 Zimet, G. D., Powell, S. S., Farley, G. K., Werkman, S. & Berkoff, K. A. Psychometric characteristics of the Multidimensional Scale of Perceived Social Support. *J Pers Assess* **55**, 610-617, https://doi.org/10.1080/00223891.1990.9674095 (1990).
61 Shelton, K. K., Frick, P. J. & Wootton, J. Assessment of parenting practices in families of elementary school-age children. *Journal of Clinical Child Psychology* **25**, 317-329, https://doi.org/10.1207/s15374424jccp2503_8 (1996).
62 Kabacoff, R. I., Miller, I. W., Bishop, D. S., Epstein, N. B. & Keitner, G. I. A psychometric study of the McMaster Family Assessment Device in psychiatric, medical ,and nonclinical samples. *J Fam Psychol* **3**, 431-439 (1990).
63 Kearney, C. A., Drabman, R. S. & Beasley, J. F. The trials of childhood: the development, reliability, and validity of the Daily Life Stressors Scale. *J Child Fam Stud* **2**, 371-388 (1993).
64 Bond, L., Carlin, J. B., Thomas, L., Rubin, K. & Patton, G. Does bullying cause emotional problems? A prospective study of young teenagers. *BMJ* **323**, 480-484, https://doi.org/10.1136/bmj.323.7311.480 (2001).
65 Bond, L., Wolfe, S., Tollit, M., Butler, H. & Patton, G. A comparison of the Gatehouse Bullying Scale and the peer relations questionnaire for students in secondary school. *J Sch Health* **77**, 75-79, https://doi.org/10.1111/j.1746-1561.2007.00170.x (2007).

66 Upthegrove, R. *et al.* Adverse childhood events and psychosis in bipolar affective disorder. *Br J Psychiatry* **206**, 191-197, https://doi.org/10.1192/bjp.bp.114.152611 (2015).
67 Favaro, A., Tenconi, E., Degortes, D., Manara, R. & Santonastaso, P. Neural correlates of prenatal stress in young women. *Psychol Med* **45**, 2533-2543, https://doi.org/10.1017/S003329171500046X (2015).
68 Berry, J. O. & Jones, W. H. The Parental Stress Scale: Initial Psychometric Evidence. *Journal of Social and Personal Relationships* **12**, 463-472, https://doi.org/10.1177/0265407595123009 (2016).
69 Motrico, E. *et al.* Psychometric properties of the List of Threatening Experiences--LTE and its association with psychosocial factors and mental disorders according to different scoring methods. *J Affect Disord* **150**, 931-940, https://doi.org/10.1016/j.jad.2013.05.017 (2013).
70 Drake, C. *et al.* The pediatric daytime sleepiness scale (PDSS): sleep habits and school outcomes in middle-school children. *Sleep* **26**, 455-458, https://doi.org/10.1093/sleep/26.4.455 (2003).
71 Bruni, O. *et al.* The Sleep Disturbance Scale for Children (SDSC). Construction and validation of an instrument to evaluate sleep disturbances in childhood and adolescence. *J Sleep Res* **5**, 251-261, https://doi.org/10.1111/j.1365-2869.1996.00251.x (1996).
72 Kocevska, D. *et al.* The Developmental Course of Sleep Disturbances Across Childhood Relates to Brain Morphology at Age 7: The Generation R Study. *Sleep* **40**, https://doi.org/10.1093/sleep/zsw022 (2017).
73 Achenbach, T. M. & Rescorla, L. A. *Manual for the ASEBA Preschool forms & Profiles*. (Usniversity of Vermont, Research Center for Children, Youth & Families, 2000).
74 ABS. Vol. 2033.0.55.00l (Australian Bureau of Statistics, Canberra, 2016).
75 Adler, N. E., Epel, E. S., Castellazzo, G. & Ickovics, J. R. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. *Health Psychol* **19**, 586-592, https://doi.org/10.1037//0278-6133.19.6.586 (2000).
76 Watson, J. F., Collins, C. E., Sibbritt, D. W., Dibley, M. J. & Garg, M. L. Reproducibility and comparative validity of a food frequency questionnaire for Australian children and adolescents. *Int J Behav Nutr Phys Act* **6**, 62, https://doi.org/10.1186/1479-5868-6-62 (2009).
77 Thorisdottir, I. E., Sigurvinsdottir, R., Asgeirsdottir, B. B., Allegrante, J. P. & Sigfusdottir, I. D. Active and Passive Social Media Use and Symptoms of Anxiety and Depressed Mood Among Icelandic Adolescents. *Cyberpsychol Behav Soc Netw* **22**, 535-542, https://doi.org/10.1089/cyber.2019.0079 (2019).
78 Frison, E. & Eggermont, S. Toward an Integrated and Differential Approach to the Relationships Between Loneliness, Different Types of Facebook Use, and Adolescents' Depressed Mood. *Communication Research* **47**, 701-728, https://doi.org/10.1177/0093650215617506 (2015).
79 *UCLA Brief COVID-19 Screen for Child/Adolescent PTSD*, https://istss.org/getattachment/Clinical-Resources/Assessing-Trauma/UCLA-Posttraumatic-Stress-Disorder-Reaction-Index/UCLA-Brief-COVID-19-Screening-Form-English-4-13-20.pdf (2020).
80 Cohen, S., Kamarck, T. & Mermelstein, R. A global measure of perceived stress. *J Health Soc Behav* **24**, 385-396 (1983).
81 Smith, B. W. *et al.* The brief resilience scale: assessing the ability to bounce back. *Int J Behav Med* **15**, 194-200, https://doi.org/10.1080/10705500802222972 (2008).
82 Cohen, J. R., So, F. K., Hankin, B. L. & Young, J. F. Translating Cognitive Vulnerability Theory Into Improved Adolescent Depression Screening: A

Receiver Operating Characteristic Approach. *J Clin Child Adolesc Psychol* **48**, 582-595, https://doi.org/10.1080/15374416.2017.1416617 (2019).
83 Zimet, G. D., Dahlem, N. W., Zimet, S. G. & Farley, G. K. The Multidimensional Scale of Perceived Social Support. *Journal of Personality Assessment* **52**, 30-41, https://doi.org/10.1207/s15327752jpa5201_2 (1988).
84 Miller, I. W., Ryan, C. E., Keitner, G. I., Bishop, D. S. & Epstein, N. B. The McMaster Approach to Families: theory, assessment, treatment and research. *Journal of Family Therapy* **22**, 168-189, https://doi.org/10.1111/1467-6427.00145 (2008).
85 Whittle, S. *et al.* Positive parenting predicts the development of adolescent brain structure: a longitudinal study. *Dev Cogn Neurosci* **8**, 7-17, https://doi.org/10.1016/j.dcn.2013.10.006 (2014).
86 Whittle, S. *et al.* Observed Measures of Negative Parenting Predict Brain Development during Adolescence. *PLoS One* **11**, e0147774, https://doi.org/10.1371/journal.pone.0147774 (2016).
87 Foulkes, L. & Blakemore, S. J. Studying individual differences in human adolescent brain development. *Nat Neurosci* **21**, 315-323, https://doi.org/10.1038/s41593-018-0078-4 (2018).
88 Schwartz, O. S. *et al.* Parenting During Early Adolescence and Adolescent-Onset Major Depression. *Clinical Psychological Science* **2**, 272-286, https://doi.org/10.1177/2167702613505531 (2013).
89 Eun, J. D., Paksarian, D., He, J. P. & Merikangas, K. R. Parenting style and mental disorders in a nationally representative sample of US adolescents. *Soc Psychiatry Psychiatr Epidemiol* **53**, 11-20, https://doi.org/10.1007/s00127-017-1435-4 (2018).
90 Denham, R. *et al.* Frequent peer problems in Australian children and adolescents. *Journal of Aggression, Conflict and Peace Research* **8**, 162-173, https://doi.org/10.1108/jacpr-11-2015-0196 (2016).
91 Thomas, H. J. *et al.* Association of different forms of bullying victimisation with adolescents' psychological distress and reduced emotional wellbeing. *Aust N Z J Psychiatry* **50**, 371-379, https://doi.org/10.1177/0004867415600076 (2016).
92 Eiland, L. & Romeo, R. D. Stress and the developing adolescent brain. *Neuroscience* **249**, 162-171, https://doi.org/10.1016/j.neuroscience.2012.10.048 (2013).
93 Prochaska, J. J., Sallis, J. F. & Long, B. A physical activity screening measure for use with adolescents in primary care. *Arch Pediatr Adolesc Med* **155**, 554-559, https://doi.org/10.1001/archpedi.155.5.554 (2001).
94 Calamaro, C. J., Mason, T. B. & Ratcliffe, S. J. Adolescents living the 24/7 lifestyle: effects of caffeine and technology on sleep duration and daytime functioning. *Pediatrics* **123**, e1005-1010, https://doi.org/10.1542/peds.2008-3641 (2009).
95 Gamble, A. L. *et al.* Adolescent sleep patterns and night-time technology use: results of the Australian Broadcasting Corporation's Big Sleep Survey. *PLoS One* **9**, e111700, https://doi.org/10.1371/journal.pone.0111700 (2014).
96 Sletten, T. L. *et al.* Genetic and environmental contributions to sleep-wake behavior in 12-year-old twins. *Sleep* **36**, 1715-1722, https://doi.org/10.5665/sleep.3136 (2013).
97 Wright, M. J. & Martin, N. G. Brisbane Adolescent Twin Study: Outline of study methods and research projects. *Australian Journal of Psychology* **56**, 65-78, https://doi.org/10.1080/00049530410001734865 (2007).
98 Hansell, N. K. *et al.* Long-term stability and heritability of telephone interview measures of alcohol consumption and dependence. *Twin Res Hum Genet* **11**, 287-305, https://doi.org/10.1375/twin.11.3.287 (2008).

99 Yap, C. X. *et al.* Autism-related dietary preferences mediate autism-gut microbiome associations. *Cell* **184**, 5916-5931 e5917, https://doi.org/10.1016/j.cell.2021.10.015 (2021).
100 El-Heneidy, A. *et al.* Respiratory virus detection during the COVID-19 pandemic in Queensland, Australia. *Aust N Z J Public Health* **46**, 10-15, https://doi.org/10.1111/1753-6405.13168 (2022).
101 Foa, E. B., Johnson, K. M., Feeny, N. C. & Treadwell, K. R. The child PTSD Symptom Scale: a preliminary examination of its psychometric properties. *J Clin Child Psychol* **30**, 376-384, https://doi.org/10.1207/S15374424JCCP3003_9 (2001).
102 Tedeschi, R. G., Cann, A., Taku, K., Senol-Durak, E. & Calhoun, L. G. The Posttraumatic Growth Inventory: A Revision Integrating Existential and Spiritual Change. *J Trauma Stress* **30**, 11-18, https://doi.org/10.1002/jts.22155 (2017).
103 Rakesh, G. *et al.* Resilience as a translational endpoint in the treatment of PTSD. *Mol Psychiatry* **24**, 1268-1283, https://doi.org/10.1038/s41380-019-0383-7 (2019).
104 El Mlili, N., Ahabrach, H. & Cauli, O. Hair Cortisol Concentration as a Biomarker of Sleep Quality and Related Disorders. *Life (Basel)* **11**, https://doi.org/10.3390/life11020081 (2021).
105 Ford, J. L., Boch, S. J. & Browning, C. R. Hair cortisol and depressive symptoms in youth: An investigation of curvilinear relationships. *Psychoneuroendocrinology* **109**, 104376, https://doi.org/10.1016/j.psyneuen.2019.104376 (2019).
106 Kische, H. *et al.* Associations of saliva cortisol and hair cortisol with generalized anxiety, social anxiety, and major depressive disorder: An epidemiological cohort study in adolescents and young adults. *Psychoneuroendocrinology* **126**, 105167, https://doi.org/10.1016/j.psyneuen.2021.105167 (2021).
107 Strike, L. T. *et al.* Queensland Twin Adolescent Brain (QTAB). *OpenNeuro*, https://doi.org/10.18112/openneuro.ds004146.v1.0.4 (2022).
108 Gorgolewski, K. J. *et al.* The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci Data* **3**, 160044, https://doi.org/10.1038/sdata.2016.44 (2016).
109 Strike, L. T. *et al.* Queensland Twin Adolescent Brain (QTAB) Non-Imaging Phenotypes. *Zenodo*, https://doi.org/10.5281/zenodo.7353054 (2022).
110 Strike, L. T. QTAB Data Use Agreement. *Zenodo*, https://doi.org/10.5281/zenodo.7430429 (2022).
111 Tournier, J. D. *et al.* MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage* **202**, 116137, https://doi.org/10.1016/j.neuroimage.2019.116137 (2019).
112 Andersson, J. L. R., Graham, M. S., Zsoldos, E. & Sotiropoulos, S. N. Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *Neuroimage* **141**, 556-572, https://doi.org/10.1016/j.neuroimage.2016.06.058 (2016).
113 Andersson, J. L. R. & Sotiropoulos, S. N. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* **125**, 1063-1078, https://doi.org/10.1016/j.neuroimage.2015.10.019 (2016).
114 Bastiani, M. *et al.* Automated processing pipeline for neonatal diffusion MRI in the developing Human Connectome Project. *Neuroimage* **185**, 750-763, https://doi.org/10.1016/j.neuroimage.2018.05.064 (2019).
115 Esteban, O. *et al.* MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLoS One* **12**, e0184661, https://doi.org/10.1371/journal.pone.0184661 (2017).

116 Lytle, M. N., McNorgan, C. & Booth, J. R. A longitudinal neuroimaging dataset on multisensory lexical processing in school-aged children. *Sci Data* **6**, 329, https://doi.org/10.1038/s41597-019-0338-5 (2019).
117 Galbraith, S., Bowden, J. & Mander, A. Accelerated longitudinal designs: An overview of modelling, power, costs and handling missing data. *Stat Methods Med Res* **26**, 374-398, https://doi.org/10.1177/0962280214547150 (2017).
118 Haworth, C. M. *et al.* The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol Psychiatry* **15**, 1112-1120, https://doi.org/10.1038/mp.2009.55 (2010).
119 Australian Institute of Health and Welfare. Australia's Youth: In Brief. (AIHW, Canberra, 2021).
120 Spence, S. H. A measure of anxiety symptoms among children. *Behav Res Ther* **36**, 545-566, https://doi.org/10.1016/s0005-7967(98)00034-5 (1998).
121 Meyer, C. *et al.* Proposal of cutoff points for pediatric daytime sleepiness scale to identify excessive daytime sleepiness. *Chronobiol Int* **35**, 303-311, https://doi.org/10.1080/07420528.2017.1400980 (2018).
122 Sung, M. *et al.* Assessment of vitamin D, exercise, and lipid profile associated with excessive daytime sleepiness in school children. *Sleep Med* **77**, 51-57, https://doi.org/10.1016/j.sleep.2020.11.017 (2021).
123 Neale, M. & Cardon, L. R. *Methodology for genetic studies of twins and families*. Vol. 67 (Springer Science & Business Media, 2013).
124 Carlin, J. B., Gurrin, L. C., Sterne, J. A., Morley, R. & Dwyer, T. Regression models for twin studies: a critical review. *Int J Epidemiol* **34**, 1089-1099, https://doi.org/10.1093/ije/dyi153 (2005).
125 Visscher, P. M., Benyamin, B. & White, I. The use of linear mixed models to estimate variance components from data on twin pairs by maximum likelihood. *Twin Res* **7**, 670-674, https://doi.org/10.1375/1369052042663742 (2004).
126 Kashyap, S. *srikash/3dMPRAGEise: ondu*, https://doi.org/10.5281/zenodo.4626825 (2021).
127 Marques, J. P. & Zwiers, M. *MP2RAGE-related-scripts*, https://github.com/JosePMarques/MP2RAGE-related-scripts (2021).
128 Esteban, O. *et al.* fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat Methods* **16**, 111-116, https://doi.org/10.1038/s41592-018-0235-4 (2019).
129 *Theory of Mind and Pain Matrix Localizer / Movie Viewing Experiment*, http://saxelab.mit.edu/theory-mind-and-pain-matrix-localizer-movie-viewing-experiment (2019).
130 Shaw, T. B. *et al.* Non-linear realignment improves hippocampus subfield segmentation reliability. *Neuroimage* **203**, 116206, https://doi.org/10.1016/j.neuroimage.2019.116206 (2019).
131 Hays, R. D., Weech-Maldonado, R., Teresi, J. A., Wallace, S. P. & Stewart, A. L. Commentary: Copyright Restrictions Versus Open Access to Survey Instruments. *Med Care* **56**, 107-110, https://doi.org/10.1097/MLR.0000000000000857 (2018).
132 Halchenko, Y. *et al.* nipy/heudiconv v0.9.0. https://doi.org/10.5281/zenodo.4390433 (2020).
133 Herholz, P., Ludwig, R. & Poline, J.-B. BIDSonym: a BIDS App for the pseudo-anonymization of neuroimaging datasets. *Journal of Open Source Software* **6**, 3169, https://doi.org/10.21105/joss.03169 (2021).
134 Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. & Smith, S. M. Fsl. *Neuroimage* **62**, 782-790, https://doi.org/10.1016/j.neuroimage.2011.09.015 (2012).