NCDS and BCS70

Twins Sub-Study

2008

USER GUIDE
September 2016, Edition 1

Denise Hawkes, Jon Johnson & Tom Murphy

Contents

1   Acknowledgement

2   Context of the data collection

The data on zygosity of the cohort twins was collected as part of an ESRC project which considered the relative importance of nature, nurture and peer effects on education and employment outcomes (Hawkes 2010). This project developed a relatively underused aspect of the National Child Development Study (NCDS) and the British Cohort Study (BCS70) data sets, that of the data on twins. It developed a more accurate series of measures for zygosity (whether the twins were identical or not) than was previously available. The original zygosity indicator was collected at birth as recorded by the medics who, most likely but not definitely, based this decision on the number of placenta found at births. This measure can substantially underestimate the number of identical twins (Bryan 1992). Previous work using the twins of the cohort studies has either not required the data on the zygosity of twins or has been based on the unreliable available data (Annett 1987, Blanchflower and Elias 1999, Emanuel et al 1992)

The aims and objectives of the project were to:

1. To provide a more complete and accurate data set which identifies the zygosity of the twins in the NCDS and the BCS70

2. To use the data held on the twins in the NCDS and BCS70 to consider the relative importance of nature, nurture and peer effects in childhood on adult outcomes such as education, employment and earnings

3. To identify the relative importance of nature, nurture and peer effects on each adult outcome across the two cohorts and within the cohorts across time in order to assess whether and how these relative effects change over the life course and the stability or change in these patterns for people born at different times

This data note provides a guide to the data for future potential users of the data.

3 Data Collection Process

Before the questionnaire on zygosity could be sent into the field a lot of work had to be undertaken to confirm which cohort members were twins and who their pair was. Information was obtained from across the sweeps of the NCDS and BCS70 where there was a mention of a twin. These suspected cases were then passed to the tracing team as potential twins to be contacted. Appendix F lists the original variable names from the NCDS and BCS70 used to locate likely twins on the data set.

In the NCDS it was possible to identify potential twins using some of the early survey data, in particular the NCDS0 Perinatal Mortality Survey (see appendix F for a full list of these variables), however as there was no twin code in the NCDS it was not clear who each twin was paired with. The twin pairs for NCDS cohort members were identified using personal data in the address database such as their names. Although a twin code was available for the BCS70 this process of confirmation of twins and their pairs was repeated to confirm this twin code.

Once their status as a potential twin was identified, the questionnaire and covering letter were developed (Appendices A-C). This questionnaire was developed based on existing instruments used by other twin studies. While self-reported zygosity is used in many studies where twins are found, it is not studied exclusively. The peas in the pod questions were taken from the St. Thomas's Twin Research Unit Questionnaires and the physical similarities questions were taken from the Australian Twin Register Questionnaires. The questionnaire was extensively peer reviewed internally and also externally by three experienced twin researchers from outside of the Institute. Before the questionnaire entered the field, ethical approval was sought from the NCDS/BCDS70 scientific committee and the Institute of Education research ethics committee. The approval from the NCDS/BCS70 scientific committee was essential as the twins surveyed were part of the cohort studies and therefore issues of respondent burden and appropriateness of the survey had to be considered. In addition, as the project involved human participation by questionnaire the university research ethics committee had to review the project and materials to be sent into the field. Both committees provided invaluable feedback on the materials and these comments were incorporated into the documents that were sent out to the twins.

Table 1 below sets out the sample development. As can be seen from the table, many more twins were originally labelled as suspected twins from the survey data (440 NCDS, 398 BCS70) than could be confirmed as twins from the address database check (254 for the NCDS and 224 for the BCS70). This can be explained by the high rates of infant mortality of twins, subsequent attrition, and incomplete address information. Those with confirmed addresses were sent the post questionnaire.

Table 1: Sample Returns for the NCDS/BCS70 Twin Survey

	Twin Sample	
	NCDS	**BCS70**
Total number of possible twins found on survey data	440	398
Contacted Addresses (contacted sample)	**254**	**224**
Total number of response	185	144
• Responses before reminders	*143*	*110*
• Additional responses after 1st reminder	*34*	*27*
• Additional responses after 2nd reminder	*8*	*7*
Number of twins with useable zygosity (after imputation)	**244**	**202**

Table 1 also sets out the response to the questionnaire from the twins of the NCDS and BCS70 that had complete address data. The questionnaires for both the NCDS and BCS70 twins were sent into the field in July 2008. During the data collection period two reminders were sent out, in October 2008 and December 2008, to those who had not responded in an attempt to optimise the response to this postal questionnaire. The response rates were good for a postal questionnaire with 72% for the NCDS and 65% for the BCS70. The lower BCS70 response rate is as expected given the age of the respondents as they are likely to be more mobile than the NCDS respondents.

Since the data collected was from twin pairs, as long as one twin replies it is possible to gain a zygosity measure for the non-responding twin. Therefore from the 185 responses in the NCDS and the 144 responses in the BCS70, useable data is obtained from 244 of the NCDS twins (leaving only 10 twins, or 5 pairs, from whom no data was collected) and 202 of the BCS70 twins (leaving only 22 twins, or 11 pairs, from whom no data was collected). A flag variable "IMPUTED" was created on the dataset to identify those cases that have zygosity data assigned according to the data of that cohort member's twin. The deposited datasets include only cases with usable zygosity, i.e. where the cohort member either responded to the survey or where they were ineligible (due to refusal or having gone away from their address for example) but had zygosity data imputed from their twin.

4   Variables Included on the Datasets

|	
BCSID/NCDSID	Research case identifier
OUTCOME	Sample Selection and Response
MULTCODE	Corresponding multiple identifier
MULTIPNO	Number of twin
IMPUTED	Zygosity data imputed from other twin
Q1	Are you one of a twin pair?
Q2	Are you a same-sex twin pair?
Q3	Are you an identical twin pair?
Q4	At school, did teachers have trouble telling you apart?
Q5	Were your parents able to tell you apart when you were of school age?
Q6	Were brothers or sisters able to tell you apart when you were of school age?
Q7	Were your close friends able to tell you apart when you were of school age?
Q8	Were strangers able tell you apart when you were of school age?
Q9	In childhood, which of the following would best describe you and your twin?
Q10	Do you have the same eye colour?
Q11A	Do you have similar height?
Q11B	Do you have similar weight?
Q11C	Do you have similar natural hair colour?
Q11D	Do you have similar hair texture?
Q12	Did you and your twin get on well with each other in childhood?
Q13	Are you still in contact with your twin? If so how frequent is this contact?
MIXEDPAIR	Same sex or mixed pair
SELFREPORT-EDZ	Self-reported zygosity
PEASZ	Peas in a pod measure
PHYSICALZ	Physical similarities scale

5 Zygosity Score Measures

The questionnaires were keyed by double entry using the coding frame in appendix D. The deposited dataset contains three distinct measures of zygosity based on the three distinct sets of questions in the questionnaire:

- **SELFREPORTEDZ** - Based on question 2 in the questionnaire, which asked directly whether the twins were identical or not.

- **PEASZ** - Based on a series of five standardised questions (questions 4, 5, 7, 8 and 9 in the questionnaire) including the famous "peas in a pod" question. This measure is considered to be the most reliable of the three. It is used by studies focused on twins such as the St Thomas's Twin Research Unit Register, and should be used as the main measure where indicators differ.

- **PHYSICALZ** - Based on a series of standardised questions (questions 10-11D in the questionnaire) about the twins' physical similarity.

These variables also contain data for non-responding twins, where it could be imputed from their responding twin. The variable "IMPUTED" indicates whether or not data for these three variables (and for MIXEDPAIR) has been imputed from another twin. Appendix E provides the stata syntax used to translate the original keyed data into these derived variables.

Figure 1 below displays the proportion of twins that are classified as identical according to the three available zygosity measures. As can be seen from Figure 1, the BCS70 cohort has a larger proportion of identical twins than the NCDS cohort (29.2% compared to 20.7%, according to the self-reported measure), which could be explained by the early focus in the BCS70 on low birth weight babies with the 22 month survey, keeping more twins engaged with the survey.

Figure 1 also shows a similar pattern with the zygosity measurements in both cohorts. As expected , the peas in a pod questions appear to correlate more closely with the self-reported identification than the physical similarities questions, indicating that the physical similarity questions alone have a tendency to overestimate the prevalence of identical twins.

Figure 1: Proportion of twins classified as identical
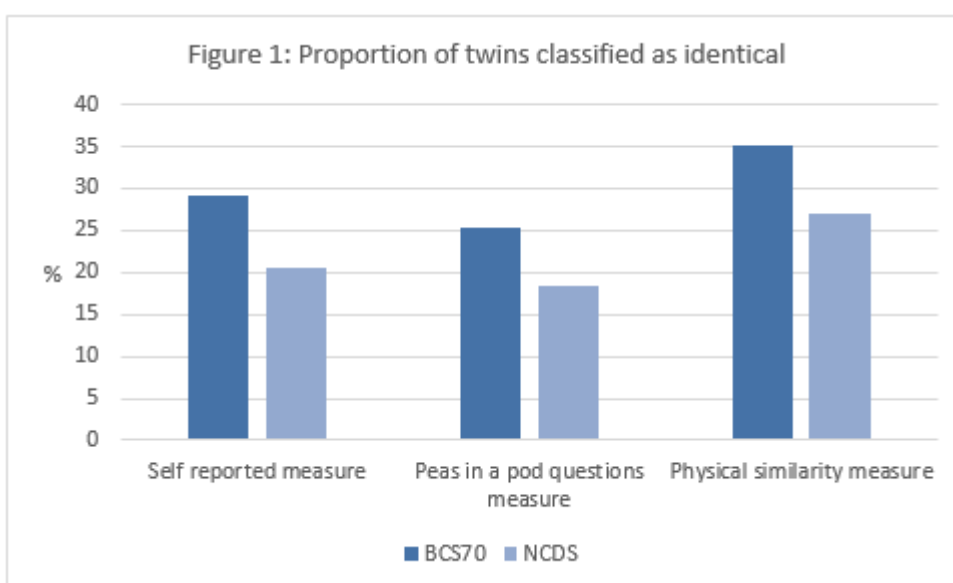
6 References

Annett, M. (1987) 'Handedness in Twins: the right shift theory.' NCDS User Support Group Working Paper 22 , Centre for Longitudinal Studies, London.

Blanchflower, D. and Elias, P. (1999) '*Ability, Schooling and Earnings: Are twins different?*', Mimeo, University of Warwick.

Bryan, E. (1992) Twins and Higher Multiple Births - A Guide to their Nature and Nurture. London/ Melbourne/Auckland: Edward Arnold.

Emanuel, I., Filakti, H., Alberman, E. and Evans, S. J. W. (1992) 'Intergenerational studies of human birthweight from the 1958 birth cohort. Do parents who were twins have babies as heavy as those born to singletons?' *, British Journal of Obstetrics and Gynaecology* , Volume 99, pp 836-840.

Hawkes, Denise (2010) "The Relative Importance of Nature, Nurture and Peer Effects on Adult Outcomes: Full Research Report ESRC End of Award Report, RES-000-22-1545". Swindon: ESRC http://www.esrc.ac.uk/my-esrc/grants/RES-000-22-1545/read

7 Appendix A:NCDS/BCS70 Twins Questionnaire

1. Are you one of a twin pair?

☐ Yes	➜	Please answer question 2
☐ Yes, but twin died at birth	➜	Please skip to the end of the questionnaire
☐ No	➜	Please skip to the end of the questionnaire

2. Are you a same-sex twin pair?

☐ Yes, both female	➜	Please answer question 3
☐ Yes, both male	➜	Please answer question 3
☐ No, one female, one male	➜	Please skip to question 12

3. Are you an identical twin pair?

☐ Yes
☐ No
☐ Don't know

4. At school, did teachers have trouble telling you apart?

☐ Yes
☐ No
☐ Don't know

5. Were your parents able to tell you apart when you were of school age?

☐ Yes
☐ No
☐ Don't know

6. Were your brothers/sisters able to tell you apart when you were of school age?

☐ Yes
☐ No
☐ Don't know
☐ Did not have any brothers/sisters

7. Were your close friends able to tell you apart when you were of school age?

☐ Yes
☐ No
☐ Don't know

8. Were strangers able tell you apart when you were of school age?

- [ ] Yes
- [ ] No
- [ ] Don't know

9. In childhood, which of the following would best describe you and your twin (please tick one only)?

- [ ] As alike as peas in a pod
- [ ] Of ordinary sibling likeness
- [ ] Don't know

10. Do you have the same eye colour?

- [ ] Yes
- [ ] No
- [ ] Don't know

11. Do you have similar height, weight and natural hair colour and texture?

Height	[ ]	Yes	[ ] No	[ ]	Don't Know
Weight	[ ]	Yes	[ ] No	[ ]	Don't Know
Natural Hair Colour	[ ]	Yes	[ ] No	[ ]	Don't Know
Hair Texture	[ ]	Yes	[ ] No	[ ]	Don't Know

12. Did you and your twin get on well with each other in childhood?

- [ ] Yes
- [ ] No
- [ ] Don't know

13. Are you still in contact with your twin? If so how frequent is this contact?

- [ ] No
- [ ] Yes – daily
- [ ] Yes – a few times a week
- [ ] Yes – a few times a month
- [ ] Yes – less than a few times a month

Thank you for completing this questionnaire.

Please return it in the pre-paid envelope enclosed together with the updated address information for you and your twin.

8 Appendix B: Covering Letter for BCS70

Our Address

Dear XXXXX

You may remember that in the past you have helped with the surveys of the British Cohort Study (BCS70) - an important ongoing research study of all those in Great Britain who were born in a single week in 1970. Your help with earlier BCS70 surveys was very important and much appreciated - it helped make the study a success.

We are writing because our records show that you are one of a twin pair and we would like to invite you to take part in a special survey of twins. As one of a twin pair you are a very special member of the study and we hope you will agree to help make our current survey of twins a success. Indeed, as there are only 200 sets of twins in the study, you are crucial to the project as we cannot replace or substitute another person's information for yours.

All we would like you to do is to complete the enclosed questionnaire and return it in the envelope provided together with the updated address information for both you and your twin. The questionnaire should only take you around 15 minutes to complete. Your answers, when combined with information you have given during earlier surveys, will help us learn more about the impact of childhood experiences on employment and wages in adult life.

As a twin, you may have already taken part been in other projects which seek to gather information about your experiences but, as a very special part of the BCS70, I hope you will agree to help with this survey.

As always, the information that you give will be treated as confidential, and will not be reported in any way which enables you to be identified. We will provide a summary of what we find when the project is completed.

If you have any queries about the survey, please contact my colleague, Denise Hawkes - you can telephone 020 7612 6881, or email d.hawkes@ioe.ac.uk. Alternatively, you can call the usual freephone number - 0500 600 616 - or email bcs70@ioe.ac.uk.

I would like to thank you in advance for your time.

Yours sincerely,

Professor Heather Joshi

9   Appendix C: Covering Letter for NCDS

Our Address

Dear XXXXX

You may remember that in the past you have helped with the surveys of the National Child Development Study (NCDS) - an important ongoing research study of all those in Great Britain who were born in a single week in 1958. Your help with earlier NCDS surveys was very important and much appreciated - it helped make the study a success.

We are writing because our records show that you are one of a twin pair and we would like to invite you to take part in a special survey of twins. As one of a twin pair you are a very special member of the study and we hope you will agree to help make our current survey of twins a success. Indeed, as there are only 250 sets of twins in the study, you are crucial to the project as we cannot replace or substitute another person's information for yours.

All we would like you to do is to complete the enclosed questionnaire and return it in the envelope provided together with the updated address information for both you and your twin. The questionnaire should only take you around 15 minutes to complete. Your answers, when combined with information you have given during earlier surveys, will help us learn more about the impact of childhood experiences on employment and wages in adult life.

As a twin, you may have already taken part been in other projects which seek to gather information about your experiences but, as a very special part of the NCDS, I hope you will agree to help with this survey.

s always, the information that you give will be treated as confidential, and will not be reported in any way which enables you to be identified. We will provide a summary of what we find when the project is completed.

If you have any queries about the survey, please contact my colleague, Denise Hawkes - you can telephone 020 7612 6881, or email d.hawkes@ioe.ac.uk. Alternatively, you can call the usual freephone number - 0500 600 616 - or email ncds@ioe.ac.uk.

I would like to thank you in advance for your time.

Yours sincerely,

Professor Heather Joshi

10   Appendix D: Keying Code for Data Entry

1. Are you one of a twin pair?

☐ Yes   ➜   Please answer question 2
☐ Yes, but twin died at birth   ➜   Please skip to the end of the questionnaire
☐ No   ➜   Please skip to the end of the questionnaire

2. Are you a same-sex twin pair?

☐ Yes, both female   ➜   Please answer question 3
☐ Yes, both male   ➜   Please answer question 3
☐ No, one female, one male   ➜   Please skip to question 12

3. Are you an identical twin pair?

☐ Yes
☐ No
☐ Don't know

4. At school, did teachers have trouble telling you apart?

☐ Yes
☐ No
☐ Don't know

5. Were your parents able to tell you apart when you were of school age?

☐ Yes
☐ No
☐ Don't know

6. Were your brothers/sisters able to tell you apart when you were of school age?

☐ Yes
☐ No
☐ Don't know
☐ Did not have any brothers/sisters

7. Were your close friends able to tell you apart when you were of school age?

☐ Yes
☐ No
☐ Don't know

8. Were strangers able tell you apart when you were of school age?

☐ Yes
☐ No
☐ Don't know

9. In childhood, which of the following would best describe you and your twin (please tick one only)?

☐ As alike as peas in a pod
☐ Of ordinary sibling likeness
☐ Don't know

10. Do you have the same eye colour?

☐ Yes
☐ No
☐ Don't know

11. Do you have similar height, weight and natural hair colour and texture?

Height	☐ Yes	☐ No	☐ Don't Know
Weight	☐ Yes	☐ No	☐ Don't Know
Natural Hair Colour	☐ Yes	☐ No	☐ Don't Know
Hair Texture	☐ Yes	☐ No	☐ Don't Know

12. Did you and your twin get on well with each other in childhood?

☐ Yes
☐ No
☐ Don't know

13. Are you still in contact with your twin? If so how frequent is this contact?

☐ No
☐ Yes – daily
☐ Yes – a few times a week
☐ Yes – a few times a month
☐ Yes – less than a few times a month

11 Appendix E: STATA Do-file used to construct the derived variables from the original data

generate mixedpair = 1 if Q2==3

replace mixedpair=0 if Q2==1|Q2==2

*mixed sex pair 1, same sex pair 0

*SELF REPORTED ZYGOSITY

*1 if identical, 2 if not, 3 if not known (treated as fraternal)

*also include fraternal mixed sex pairs

generate selfreportedz = 1 if Q3==1

replace selfreportedz=0 if Q3==2|Q3==3|Q2==3

*1 identical 0 fraternal

*Peas in a pod identity scale

*include as fraternal mixed self pairs

*Initially set to 0, so can pick up partial responders that have enough info to be identified as identical

gen zyg1=0

replace zyg1=1 if Q4==1

gen zyg2=0

replace zyg2=1 if Q5==2

gen zyg3=0

replace zyg3=1 if Q7==2

gen zyg4=0

replace zyg4=1 if Q8==2

gen zyg5=0

replace zyg5=1 if Q9==1

*generate fra1-5 to pick up partial responders that can be identified as fraternal

gen fra1=0

replace fra1=1 if Q4==2

gen fra2=0

replace fra2=1 if Q5==1

gen fra3=0

replace fra3=1 if Q7==1

gen fra4=0

replace fra4=1 if Q8==1

```
gen fra5=0

replace fra5=1 if Q9==2

gen zygosityscore = zyg1+zyg2+zyg3+zyg4+zyg5

gen frascore = fra1+fra2+fra3+fra4+fra5

gen peasz = 1 if zygosityscore>=3 & zygosityscore~=.

replace peasz = 0 if mixedpair==1

replace peasz = 0 if peasz==. & frascore>=3

*physical charateristics measure

*again mixed sex pairs included as fraternal

gen phy1=1 if Q10==1

replace phy1=0 if Q10==2|Q10==3

gen phy2=1 if Q11A==1

replace phy2=0 if Q11A==2|Q11A==3

gen phy3=1 if Q11B==1

replace phy3=0 if Q11B==2|Q11B==3

gen phy4=1 if Q11C==1

replace phy4=0 if Q11C==2|Q11C==3

gen phy5=1 if Q11D==1

replace phy5=0 if Q11D==2|Q11D==3

gen physicalscore = phy1+phy2+phy3+phy4+phy5

gen physicalz = 1 if physicalscore>=4 & physicalscore~=.

replace physicalz = 0 if physicalscore<4

replace physicalz = 0 if mixedpair==1

*rename Q1 and reorder file for reshape

ren Q1 Q1_

order MULTCODE MULTIPNO

*reshape to genereate derived vars with data imputed from other twin

reshape wide SERIAL-physicalz, i(MULTCODE) j(MULTIPNO)

*using one twin to report for two

gen MIXEDPAIR = 1 if (mixedpair1==1 & mixedpair2==1)|(mixedpair1==1 & mixed-
pair2==.)|(mixedpair1==. & mixedpair2==1)

replace MIXEDPAIR = 0 if (mixedpair1==0 & mixedpair2==0)|(mixedpair1==0 & mixed-
pair2==.)|(mixedpair1==. & mixedpair2==0)

gen SELFREPORTEDZ = 1 if (selfreportedz1==1 & selfreportedz2==1)|(selfreportedz1==1 &
selfreportedz2==.)|(selfreportedz1==. & selfreportedz2==1)
```

replace SELFREPORTEDZ = 0 if (selfreportedz1==0 & selfreportedz2==0)|(selfreportedz1==0 & selfreportedz2==.)|(selfreportedz1==. & selfreportedz2==0)

gen PEASZ = 1 if (peasz1==1 & peasz2==1)|(peasz1==1 & peasz2==.)|(peasz1==. & peasz2==1)

replace PEASZ = 0 if (peasz1==0 & peasz2==0)|(peasz1==0 & peasz2==.)|(peasz1==. & peasz2==0)

gen PHYSICALZ = 1 if (physicalz1==1 & physicalz2==1)|(physicalz1==1 & physicalz2==.)|(physicalz1==. & physicalz2==1)

replace PHYSICALZ = 0 if (physicalz1==0 & physicalz2==0)|(physicalz1==0 & physicalz2==.)|(physicalz1==. & physicalz2==0)

reshape long

replace MIXEDPAIR = mixedpair if MIXEDPAIR==.

replace SELFREPORTEDZ = selfreportedz if SELFREPORTEDZ==.

replace PEASZ = peasz if PEASZ==.

replace PHYSICALZ = physicalz if PHYSICALZ==.

*generate flag for whether data has been imputed from other twin

gen IMPUTED=0 if MIXEDPAIR~=.

replace IMPUTED=1 if (MIXEDPAIR~=. & mixedpair==.) | (SELFREPORTEDZ~=. & selfreportedz==.) | (PEASZ~=. & peasz==.) | (PHYSICALZ~=. & physicalz==.)

12   Appendix F: Survey data variables used to identify twin pairs in NCDS and BCS70

NCDS	
Variable	**Description**
serial	Identifier
n622	Sex of child
n0region	Region at PMS (1958) - Birth
n504	Has the patient had any previous pregnancies (including miscarriages)?
n1811	Multiple birth identification (e.g. first twin, second twin, first triplet etc.)
n1846	Twin identification (i.e. whether twin is identical, based on number of placentas present at birth)
n1847	Twin identification by sex (i.e. whether child is a same-sex twin)
n1848	Twin identification, sole survivor (i.e. whether child is the sole survivor of a twin pair)
n553	Mother's age last birthday, in years
n494	Husband's age in years, 1958

BCS70	
Variable	**Description**
chesno/chesno2	Identifier
ctc	CHES Number Twin code
ccd	CHES Number Check digit
key	Identifier
serial	Identifier
tc2	Identifier
tc10	PI -Singleton, twin code (i.e. whether singleton, first twin, or second twin)
d002	Singleton or Twin
e001	Singleton or Twin
a0002	Multiplicity Code
a0248	Birth Multiple Single, Twin, Triple
twin	Number of other matching twin