Environmental Factors Determine Where the Dutch Live: Results From the Netherlands Twin Register

Gonneke Willemsen, Danielle Posthuma, and Dorret I. Boomsma
Department of Biological Psychology, Vrije Universiteit Amsterdam, the Netherlands

The heritability of the degree of residential area urbanization in twins and their siblings in the Dutch population was examined. The postal code was known for 6879 twins and 2724 siblings registered with the Netherlands Twin Register and born between 1940 and 1983. Using data from Statistics Netherlands (Centraal Bureau voor de Statistiek, 2001), these postal codes could be related to residential area characteristics, including urbanization level. The degree of urbanization was assessed on a 5-point scale: *very heavy, heavy, moderate, low* and *not urbanized*. Genetic model-fitting was carried out in three age cohorts: young adults (born 1975 to 1983), adults (born 1965 to 1974) and older adults (born 1940 to 1964). Twin and sibling resemblance in urbanization level was expressed in polychoric correlations. These correlations decreased from the youngest cohort (.66 to .86) to the oldest cohort (.20 to .58). In all 3 age cohorts, genetic factors did not contribute to familial resemblance. The influence of common environment decreased in importance from the young cohort (70% to 83%) to the old cohort (46% to 47%) and was lower in women than in men in all but the oldest age cohort. This study did not replicate Australian findings of a genetic contribution in the older cohorts; common environmental factors and, increasingly with age, unique environmental factors determine where the Dutch live. Future studies in European and other populations will reveal whether these results are specific to the Dutch population.

Recently Whitfield et al. (2005) reported on the heritability of the choice of residential area in an Australian population. They found the choice of residence to be influenced by environmental factors at a young age but as age increased additive genetic factors became more important. The authors suggest that these results may have implications for the genetic variation in populations and the study of disease distributions. However, the unique geographical characteristics of Australia may profoundly impact individuals' residential choices. Australia is a country with just over 20 million residents on an area of 7,617,931 square kilometres, that is, less than three residents per square kilometre (International Data Base, US Bureau of the Census). Although it is not a densely populated country, there is a large conglomeration of the population in a few cities along the coast. The Australian twin study examined whether individuals lived near the city centre (urban), some distance from the city centre (suburban) or far from the city centre (rural) in these cities. It is likely that the large traveling distances will have influenced individuals to move to the area of employment and this may have also affected the choice of living area within the city. In contrast to Australia, the Netherlands is one of the most densely populated countries in the world. Almost 16.5 million people live on an area of 33,939 square kilometres, that is, more than 480 residents per square kilometre (International Data Base, US Bureau of the Census). The distance between cities is short, in fact the entire country can be traversed in half a day, and individuals can commute easily from their residence to the place of their employment. As such, the Dutch population represents an excellent case to test the generalizability of the results found in the Australian sample. We used the postal codes of the addresses of twins and siblings registered with the Netherlands Twin Register (NTR) to obtain the degree of urbanization from the database of Statistics Netherlands (Centraal Bureau voor de Statistiek, 2001). These data were used to determine the heritability of the degree of residential area urbanization in three different age cohorts.

Methods

Participants

In 1991 the NTR started a longitudinal survey study of health, lifestyle and personality. Addresses of adolescent and adult twins were obtained from council

registers. In addition, since 1995 twins have been able to voluntarily register themselves, at which point they are included in the survey study. Every 2 to 3 years a questionnaire booklet is sent to twins and their family members (Boomsma et al., 2002; Vink et al., 2004). In addition, yearly information leaflets on the research of the NTR are sent to those registered. As a result of these mailings, the address database is continuously updated. Postal code data were available for 15678 twins and 4116 singleton siblings of twins. For this article, we selected twins and the siblings of twins who completed a survey when they were between 20 and 65 years of age and who were born between 1940 and 1983 (6678 twins and 2336 siblings). In addition, we selected twins and siblings who did not participate in the survey studies and who were born between 1940 and 1965. Twins were included if they were from an opposite-sex pair (otherwise zygosity was unknown) and siblings were included if sex was known. This resulted in 201 additional twins and 388 siblings, bringing the total number to 9603 persons (6879 twins and 2724 siblings). These twins and siblings came from 4158 families, with an average family size of 2.31 (*SD* = 1.11). In 3318 families, data were available for two or more twins and/or siblings. Complete twin pair data were available for 2986 families, and for 85 families, data were only available for siblings not twins. Based on birth year, three cohorts were formed: young adults born between 1975 and 1983, adults born between 1965 and 1975, and older adults born between 1940 and 1965. For the analyses within an age cohort, only the data of twins and siblings within that cohort were taken into account. To avoid numerical problems in estimation, data for a maximum of two male siblings and two female siblings were included in the analyses. Additional siblings (if they fell into the same birth cohort) within a family were not included in the analyses (*N* = 115). This final selection resulted in 2921 persons (2423 twins, 498 siblings) in the young adult cohort, 4359 persons (3095 twins, 1264 siblings) in the adult cohort and 2208 persons (1361 twins, 847 siblings) in the older adult cohort.

Zygosity

Twin zygosity was determined from DNA polymorphisms, or, when DNA was not available, from survey questions. Every survey asked each of the twins whether they were alike in eye color, hair color, face color and face form. Twins also indicated whether they were sometimes mistaken for each other as a child by their parents, other family members and strangers. Parents and siblings were also asked to answer these questions about the twins. Based on the answers to these questions, twin zygosity was determined for every occasion and person separately. Next, all individual judgments were combined to determine one measure of twin zygosity. In 4.7% of the cases the judgment was inconsistent over time and persons. When there was inconsistency over time and persons, the majority of judgments determined the final

outcome. For 869 same-sex twin pairs, zygosity was available from both DNA polymorphisms and questionnaires. If no inconsistencies existed for the questionnaire reports, the correspondence with DNA zygosity was 98% (704 of 715 same-sex twin pairs). When also including zygosity outcomes based on inconsistent questionnaire reports, the correspondence with DNA zygosity was 97% (845 of 869 same-sex pairs). In case of discrepancies between questionnaire and DNA polymorphism, the DNA zygosity was used.

Urbanization

Information on regional characteristics was obtained from the 1992 postal code information provided by Statistics Netherlands (Centraal Bureau voor de Statistiek, 2001). Postal codes are four-digit numbers which indicate a village or a specific region within a city. For each postal code, the Statistics Netherlands data provided an urbanization level on a scale of 1 to 5 (*very heavy, heavy, moderate, low, none*). We used our most recently updated address database (week 4, 2005) to obtain the postal codes and determine the urbanization level for persons registered with the NTR.

Analyses

A threshold model (Falconer & Mackay, 1996; Whitfield et al., 2005) was used to obtain parameter estimates for thresholds, familial correlations and the genetic and environmental variance components. We first tested if the thresholds were similar across sex and across twins and siblings. Next, the extent to which the variation in urbanization was explained by additive genetic factors (A), common or shared environmental factors (C) and unique environmental factors (E) was examined. To determine the importance of the A and C component, the full ACE model was tested against reduced models (AE, CE), using likelihood ratio tests and Akaike's Information Criterion (AIC; Akaike, 1987) to determine the fit of the individual models. Sex differences were tested for in the proportion of variance explained by the A, C and E components. Correlations between the latent A factors were 1 for monozygotic (MZ) and .5 for dizygotic (DZ) and sibling pairs. Correlations between the latent C factors (R_c) were 1 in MZ and same-sex DZ twin pairs. For opposite-sex DZ twin pairs, for same-sex sibling pairs and for opposite-sex sibling pairs separate estimates of R_c were obtained. For the best-fitting model (ACE, AE, CE or E) we tested whether R_c for the opposite-sex DZ twins (DOS) group could be set to 1, if R_c was equal in the same-sex and opposite-sex siblings, and whether R_c in these groups could be constrained at 1. If R_c in opposite-sex pairs does not differ from 1, it indicates the same familial environmental factors are of importance in males and females. The test if R_c is less than 1 in same-sex sibling pairs assesses whether there is a special twin environment that makes twins more similar than siblings. Model-fitting was carried out in Mx (Neale et al., 2003).

Results

Table 1 shows the prevalence of each of the urbanization levels separately for the three age cohorts. As can be seen, younger twins and siblings lived more often in heavy urbanized areas than older twins and siblings (twins χ^2 = 148.076, p < .001; siblings χ^2 = 51.977, p < .001). There were no sex differences in the thresholds within any of the age cohorts. Twins and siblings differed in the thresholds for the two younger age cohorts; twins lived more often in heavy urbanized areas than siblings in these cohorts.

Table 2 shows the estimates for the polychoric correlations for the three age groups as obtained by maximum likelihood estimation. These correlations suggest that, in all three age groups, the urbanization level of the residential area is for a large part determined by common environmental influences.

Table 3 provides the results of the modeling of the contribution of genetic, common environmental and unique environmental factors to the variance in urbanization level. Similar results were found for all three age groups; dropping A from the model did not significantly reduce the fit of the model, but removing C from the model was not permitted. In all three age groups, the CE model provided the best description of the data, which is in line with the pattern already suggested by the twin and sibling correlations. The contribution of C decreased from the youngest (83% in men, 70% in women) to the oldest cohort (46% in men, 47% in women). In the youngest two cohorts the proportion of the variance explained by common environmental factors was different for men and women. Common environment contributed more to urbanization level in men than in women. For the oldest cohort, the contribution of C was equal in men and women. Table 3 gives the estimates for R_c. In the youngest cohort, the common environmental correlation (R_c) could be restricted to 1 in the DOS twins, the same-sex and opposite siblings groups. In the second oldest cohort the DOS R_c is less than 1, but at .84 it is higher than the same-sex and opposite-sex siblings for which R_c could be set to be equal at .75. In the oldest cohort, neither DOS R_c or same-sex and opposite-sex sibling R_c could be restricted to 1; the environmental correlation could be set to be equal in these three groups at .61. It is clear that especially in the older adults, the common environmental factors that contribute to resemblance in urbanization level are different in the two sexes and in twins and siblings.

Discussion

For our Dutch twins and siblings, the urbanization level of the residential area was determined by common and unique environmental factors. Genetic factors did not exert an influence on the urbanization level in any of the age groups. These results only partly confirm findings in an Australian sample (Whitfield et al., 2005). Whitfield et al. (2005) found the distance to the city centre to be determined by common and unique environmental factors in younger cohorts, but in older populations (born before 1965), genetic factors also played a role. Considering the differences between the two studies, in particular between the two study populations, the discrepancy in the findings for the older participants may be attributed to a number of factors.

The two countries differ strongly in population density, with the Netherlands being one of the most densely populated countries in the world while Australia is vastly larger and has a population density which is not even one hundredth of the Dutch population density. As a result, the operationalization of urbanization level as the distance to the city centre which was used in the Australian sample is almost meaningless in the Netherlands as Dutch cities are in close proximity, in particular in the midwestern part of the Netherlands. The discrepancy in results may therefore be caused in part by differences in the definition of urbanization level.

Being so densely populated, the Netherlands has a scarcity of suitable housing, in particular in certain popular areas and for persons with lower incomes. This has led to the establishment of a housing allocation system based on points. These points are in part awarded based on the duration of the time a person is registered, and registration is often limited to those already living in the area or with economic ties to the area. This system is not only used for rental housing but also for determining rank order for the allocation

---

Table 1

Degree of Urbanization of the Residence Areas of Twins and Their Siblings for the Three Age Cohorts Expressed as the Percentage (*N*) of Individuals Within Each of the Urbanization Levels

	1975–1983		1965–1974		1940–1964	
Urbanization level	Twins	Siblings	Twins	Siblings	Twins	Siblings
Very heavy	26.1 (632)	20.3 (101)	21.4 (662)	18.1 (229)	11.2 (152)	9.7 (82)
Heavy	20.9 (506)	22.1 (110)	18.1 (561)	19.2 (243)	19.9 (271)	18.7 (158)
Moderate	17.7 (429)	17.5 (87)	20.4 (631)	18.8 (237)	24.3 (331)	24.4 (207)
Low	18.0 (436)	21.7 (108)	21.0 (651)	21.8 (276)	23.4 (319)	26.2 (222)
None	17.3 (420)	18.5 (92)	19.1 (590)	22.1 (279)	21.2 (288)	21.0 (178)

Table 2

Polychoric Correlations (95%CI) for Twin and Siblings Pairs in the Three Age Cohorts

	1975–1983	1965–1974	1940–1964
MZM twins	.82 (.76–.87)	.74 (.67–.81)	.41 (.22–.55)
DZM twins	.86 (.78–.90)	.74 (.65–.80)	.58 (.36–.72)
MZF twins	.71 (.64–.76)	.62 (.54–.68)	.48 (37.–.72)
DZF twins	.66 (.56–.73)	.63 (.54–.70)	.44 (.28–.57)
DOS twins	.76 (.69–.81)	.57 (.49–.64)	.20 (.02–.37)
Male siblings	.74 (.62–.82)	.55 (.45–.63)	.25 (.10–.39)
Female siblings	.68 (.58–.76)	.52 (.43–.60)	.28 (.16–.39)
Opposite sex siblings	.69 (.61–.75)	.49 (.42–.55)	.34 (.25–.42)

Note: Sibling pairs may consist of two singleton siblings or a twin and its singleton sibling.

of subsidized newly built housing. As such, persons may often get housing near the area of their parental home, as they are more likely to be eligible for housing here, which would be seen in the present results as an increased influence of common environmental factors. Persons with a higher income though are not eligible for governmental housing and their choice of residential area is therefore less restricted and more often based on their economic situation.

As the Australian sample included more old-aged twins, it is also possible that the age range in our sample was not wide enough to show the influence of genes in the older part of the population. However, our oldest cohort did correspond to the middle-aged cohort in the Australian sample, for which genetic effects were also evident. Still, it is possible that in the

Dutch population genes will exert an influence only when individuals retire and commuting to work is no longer necessary. Indeed, the largest part of the migration from urban to rural areas in the Netherlands has been found to be contributable to middle-aged and elderly persons, in particular 55- to 64-year-olds (van Dam et al., 2002).

The discrepancies in results may also be caused by differences in study design. Singleton siblings of twins were included in our design and a measure of urbanization was used that had five categories instead of three. Using the extended twin-sibling design increases the power to detect the influence of common environmental factors (Posthuma & Boomsma, 2000). To determine whether this difference in design may explain the difference in results, genetic modeling in twins only was carried out. These results were similar to those found with the twin-sibling design; in all three age cohorts, individual differences in the urbanization level of the residential area were explained by common and unique environmental influences, not by genes.

In the youngest age cohorts, twins lived more often in heavily urbanized areas than siblings. This may reflect the fact that the singleton siblings were often older than their twin siblings (siblings were older than twins in more than 65% of the families within each age cohort) and therefore more likely to be working and have accommodation other than student or starter housing. In general, younger persons (18 to 25 years) tend to move into the cities for employment, only to move out towards quieter areas a couple years later (van Dam et al., 2002). The fact that no twin singleton

Table 3.

Model-Fitting Results and Estimates of the Proportional Contribution of A, C and E

	1975–1983				Men			Women			R_c		
Model	−2LL	df	Δχ²	p	A	C	E	A	C	E	DOS	sib-ss	sib-os
ACE	8545.101	2906			.00	.83	.17	.02	.68	.30	1	.92	.93
AE	8703.309	2908	158.199	.00	.85	—	.15	.75	—	25	—	—	—
CE	8545.212	2908	.110	.94	—	.83	.17	—	.70	.30	1	.92*	.92*
	1965–1974				Men			Women			R_c		
Model	−2LL	df	Δχ²	p	A	C	E	A	C	E	DOS	sib-ss	sib-os
ACE	13217.850	4344			.00	.73	.27	.00	.63	.37	.84	.78	.72
AE	13325.387	4346	107.537	.00	.80	—	.20	.69	—	.31	—	—	—
CE	13217.853	4346	.003	>.99	—	.73	.27	—	.63	.37	.84	.78	.72
	1940–1964				Men			Women			R_c		
Model	−2LL	df	Δχ²	p	A	C	E	A	C	E	DOS	sib-ss	sib-os
ACE	6763.773	2197			.00	.46	.54	.00	.47	.53	.42	.57	.71
AE	6780.270	2199	16.497	.00	.49	—	.51	.52	—	.48	—	—	—
CE	6763.773	2199	0.000	>.99	—	.46	.54	—	.47	.53	.42	.57	.72

Note: The last three columns give the estimates for the common environmental correlation for opposite-sex twin pairs, same-sex sibling pairs and opposite-sex sibling pairs.

R_c = environmental correlation, DOS = opposite-sex DZ twins, sibs-ss = same-sex siblings, sib-os = opposite-sex siblings, * = R_c equals 1.

In the AE model, estimates for R_c are not obtained as C is dropped from the model.

In the 1965–1974 cohort, R_c can be set to be equal in sib-ss and sib-os at .75.

In the 1940–1964 cohort, R_c can be set to be equal in DOS, sib-ss en sib-os at .61

differences were found in urbanization level in the older cohort suggests that this is a likely explanation.

Age cohort also influenced the effect of the common environment; C estimates decreased from the younger cohort to the older cohort. As people grow older, the influence of their family background decreases while they will have to take into account their changing individual situation due to, for instance, the demands of employment, raising children or health problems. Also, in the youngest cohort, a number of twins may still have been living at home or may have used the parental home address for correspondence while living in student accommodation, and this will have added to the estimated influence of C. The decreased R_c for same-sex siblings in the older cohorts seems to suggest a special twin environment that makes twins more similar than siblings at an older age, though this may reflect the age similarity in twins. Still, it is possible that twins are more inclined to move closer to each other at a later age, when demands of work and children lessen, than are singleton siblings. To determine this, the actual distance between the two residences, rather than the residential characteristics itself will have to be examined.

Sex also had an effect on the influence of the common environment: the influence of C was larger for men than for women, though not in the oldest cohort. The decreased opposite-sex twin and sibling correlations for the oldest two cohorts further indicate that different environmental factors operate in men and women. These sex differences may be related to the fact that, particularly in older generations, the man is more often the main provider (Fokkema, 2002) and the man's place of employment may determine the residential area. As stated previously, when persons get older, other nonwork-related factors, such as health and closeness to family members, become more important. This is reflected in the increased influence of unique environmental factors in both men and women in the oldest cohort. As a result, the estimate of the influence of common environment becomes similar for men and women, although the lower R_c in this cohort indicates that these environmental factors differ for the two sexes.

Determining the reasons for individual differences in the urbanization level of one's residential area is important, as high urbanization levels are generally associated with decreased mental and physical health. For instance, the prevalence of cardiovascular disease, depression and schizophrenia are all increased in areas with a high urbanization level (e.g., Pedersen & Mortensen, 2001; Peen & Dekker, 2003; Sundquist et al., 2004). The development of these disorders has been shown to be partly influenced by genetic factors (e.g., Boomsma et al., 2000; Haukka et al., 2004; O'Donovan et al., 2003; Scheuner, 2001). The present study showed that individual differences in urbanization level in the Netherlands are not genetic in origin, providing further support for the hypothesis that a complex gene–environment interaction determines the impact of the urbanization level on health, probably through effects of lifestyle and sociodemographic factors (Haukka et al., 2004; Peen & Dekker, 2003).

In conclusion, environmental factors, not genes, determine where the Dutch live. As the Dutch grow older, the influence of shared environment decreases. These results differ in part from Australian findings. Future studies in European and other populations are needed to determine whether these results are specific to the Dutch population.

References

Akaike, H. (1987). Factor analysis and AIC. *Psychometrica, 52*, 317–332.

Boomsma, D. I., Beem, A. L., van den Berg, M., Dolan, C. V., Koopmans, J. R., Vink, J. M., de Geus, E. J., & Slagboom, P. E. (2000). Netherlands twin family study of anxious depression (NETSAD). *Twin Research, 3*, 323–234.

Boomsma, D. I., Vink, J. M., van Beijsterveldt, T. C., de Geus, E. J., Beem, A. L., Mulder, E. J., Derks, E. M., Riese, H., Willemsen, G. A., Bartels, M., van den Berg, M., Kupper, N. H., Polderman, T. J., Posthuma, D., Rietveld, M. J., Stubbe, J. H., Knol, L. I., Stroet, T., & van Baal G. C. (2002). Netherlands Twin Register: A focus on longitudinal research. *Twin Research, 5*, 401–406.

Centraal Bureau voor de Statistiek. (2001). *Kerncijfers Viercijferige Postal codegebieden 1999* (Indices of four-digit postal code areas 1999). Voorburg, the Netherlands: Heerlen.

Fokkema, T. (2002). Combining a job and children: Contrasting the health of married and unmarried women in the Netherlands? *Social Science & Medicine, 54*, 741–752.

Falconer, D. S., & Mackay, T. F. C. (1996). Quantitative genetics. Essex, UK: Longman.

Haukka, J. K., Suvisaari, J., & Lonnqvist, J. (2004). Family structure and risk factors for schizophrenia: Case-sibling study. *BMC Psychiatry, 27*, 41.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). Mx: Statistical modeling (6th ed.). Richmond, VA: Department of Psychiatry, Medical College of Virginia.

O'Donovan, M. C., Williams, N. M., & Owen, M. J. (2003). Recent advances in the genetics of schizophrenia. *Human Molecular Genetics, 15*, R125–133.

Pedersen, C. B., & Mortensen, P. B. (2001). Evidence of a dose-response relationship between urbanicity during upbringing and schizophrenia risk. *Archives of General Psychiatry, 58*, 1039–1046.

Peen, J., & Dekker, J. (2003). Urbanization as a risk factor for psychiatric admission. *Social Psychiatry & Psychiatric Epidemiology, 38*, 535–538.

Posthuma, D., & Boomsma, D. I. (2000). A note on the statistical power in extended twin designs. *Behavioral Genetics, 30*, 147–158.

Scheuner, M. T. (2001). Genetic predisposition to coronary artery disease. *Current Opinions in Cardiology, 16*, 251–260.

Sundquist, S., Frank, G., & Sundquist, J. (2004). Urbanization and incidence of psychosis and depression: Follow-up study of 4.4 million women and men in Sweden. *British Journal of Psychiatry, 184*, 293–298.

US Bureau of the Census. (2004, September). International data base. Retrieved March 22, 2005, from http://www.census.gov/ipc/www

Van Dam, F., Heins, S., & Elbersen, B. S. (2002). Lay discourses of the rural and stated and revealed preferences for rural living: Some evidence of the existence of rural idyll in the Netherlands. *Journal of Rural Studies, 18*, 461–476.

Vink J. M., Willemsen G., Stubbe J. H., Middeldorp C. M., Ligthart R. S., Baas K. D., Dirkzwager, H. J., de Geus, E. J., & Boomsma, D. I. (2004). Estimating non-response bias in family studies: Application to mental health and lifestyle. *European Journal of Epidemiology, 19*, 623–630.

Whitfield J. B., Zhu G., Heath, A. C., & Martin, N. G. (2005). Choice of residential location: Chance, family influence, or genes? *Twin Research and Human Genetics, 8*, 22–26.