An Observational Study of Race and Gender Homophily in Nursery Children

Shayna R. Weisz & Yvan I. Russell
Middlesex University, U.K.

Homophily (the preference for similar others) is a commonplace feature of social life. In this observational study, we recorded association patterns (based on spatial proximity or verbal or physical interaction) among children aged 3-4 years old during unstructured playtime in a university nursery. A basic social network analysis and a quadratic assignment procedure revealed gender and race to be significant predictive factors of social interaction, with girls seemingly displaying more racial homophily than boys. Age, parent occupation and number of siblings did not predict interaction patterns.

Key Words: Homophily, Social Network, Race, Gender, Children

'Homophily' – which literally means "love of the same" (Laursen, 2017, p. 282) – is defined as preferential social contact towards those who are similar as compared to those who are different (McPherson et al., 2001). Homophily is a markedly powerful predictor of initial peer attraction and the subsequent establishment of friendships (McDonald et al., 2013). During childhood, friends typically share demographic characteristics, interests, behavioral and personality traits (McCormick et al., 2015).

There are benefits to homophily, such as higher academic achievement among girls socializing with other girls (Connolly, 2004) – but it may also lead to negative effects such as outgroup prejudice (McCormick et al., 2015) or sexism (Karpiak et al., 2007). There are numerous social and academic benefits to encouraging cross-group friendship formation (e.g. Crystal et al., 2008, McGill et al., 2012), and while homophily has been well-studied in older children and adolescents, Eivers et al. (2012) noted that there have been relatively few studies of the friendship choices of younger children. This is despite research showing that important friendships are formed and maintained at a very young age (Holder & Coleman, 2015). Studying younger children would afford us better knowledge on how to prevent the aforementioned negative effects of homophily from an early age (Laghi et al., 2013).

Author info: Correspondence should be sent to: Dr. Yvan Russell, Middlesex University, United Kingdom,

North American Journal of Psychology, 2022, Vol. 24, No. 4, 691-700.

Our study aimed to fill this gap by exploring homophily among 3-4 year olds, specifically investigating which attributes are most salient in their friendship choices. Previous studies have shown that the most consistent predictors of friendship include age, gender, and race (e.g. McPherson et al., 2001; Rubin et al., 2015; Wang et al., 2019), but as perceptions of race, gender, and other attributes are constantly evolving over time these investigations bear repeating, to determine whether changes in social attitudes are impacting children's interaction styles. These three attributes (age, gender, and race) were the focus of our study, along with two other attributes: the child's number of siblings (due to purported links between sibling number and social skills, e.g. Bobbitt-Zeher & Downey, 2013), and parents' occupation (for a possible gauge of the impact of socioeconomic differences).

Due to language limitations, it can be difficult to gather valid data from children of such a young age using the typical self-report methodologies previous studies have employed, such as peer nominations (e.g. asking children to list their top three best friends) or peer ratings (asking children to rate peers on a specific dimension on a Likert scale). However, not only may language limitations at this age present a challenge, but a number of other critiques of the validity of these methods have been raised, such as social desirability effects, the potential for teasing and gossiping to ensue, exhibiting of a position preference, or the difference between someone's "likeability" versus their actual position within a social group (see Rubin et al., 2015 and Yugar & Shapiro, 2001).

Therefore, this study employed the method of social network analysis via direct time-based scan sampling, in order to avoid the potential biases or misapprehension that self-report studies may have involved, and any language barriers that may have interfered. Observational research on children has a long history (Maccoby & Jacklin, 1987) and it provides a valid measure of homophily based on how many individual interactions take place between each child, which could then be visualized through network mapping (Scott, 2000), tested using a quadratic assignment procedure (Whitbred, 2011).

In our study, a basic social network analysis (Scott, 2000) was performed to analyze children's interactions based on observations of their free-play time in a nursery classroom. However, we need to be careful in our terminology. The word "friendship" is vague and can refer to a wide variety of forms of relationships (Pahl & Spencer, 2004). Therefore, we avoided the word "friendship" in the study below. Instead, we focused on simple quantifications of time: how many minutes (if any) that a given child spends in close physical proximity with every other child in the nursery. Then, we looked for homophily by comparing these

quantities across children with the five different attributes mentioned earlier (age, gender, race, number of siblings, and parents' occupation). Homophily can be inferred if there are higher rates of association between children with more similarities than those with fewer similarities. Based on the studies mentioned earlier, we predicted that we would observe homophily based on the previously mentioned five attributes (age, gender, race, number of siblings, and parents' occupation).

METHOD

Participants

The participants were children who attended the Middlesex University nursery in London, UK (n = 31). They consisted of 16 girls and 15 boys, 16 of whom were three years old and 15 of whom were four. As well as age and sex, the children's race, number of siblings, and their parents' occupation were also ascertained, and the numbers of each are shown in Tables 1-3 below. Nursery staff were not recorded as part of the data, and children's siblings were not present during the study. Prior to the study, consent forms were collected from children's parents/guardians. This study was approved in advance by the research ethics committee in the Psychology Department of Middlesex University.

Table 1: Number of children whose parents' fell into each occupational category

Academics	Non-academic university staff	Students	Other
10	8	4	9

Table 2: Number of siblings belonging to each child

Only child	One sibling	Two siblings	More than two siblings
10	17	2	2

Table 3: Number of children in each racial group

Asian	Asian Mixed	Black	Black Mixed	Other	White British	White Other
4	2	2	3	1	11	8

Materials

The children's interactions were recorded by the observer using physical data collection sheets, each consisting of a set of six boxes

("outer boxes"), which were further divided up into ten "inner boxes'. Each "outer box" represented one minute of observation time, and as each new minute struck (as measured by a stopwatch), the observer would note down each child's assigned two-letter code (for ease of quick writing, as well as for anonymity) within one of the ten boxes, depending on with whom else they were interacting. An example of one "outer box" and its ten "inner boxes" containing the children's interaction clusters is shown in Figure 1 below.

Children typically clustered in no more than ten separate groups, so ten "inner boxes" were sufficient for each minute. "Interactions" were considered to be any time a child was within approximately one meter of another, and facing them. If a child moved around within the space of a minute, it was their location at the top of the minute that was recorded. All observations were conducted in the nursery classroom or the designated nursery playground. Subsequently, each interaction between any two children ("dyad") were tallied, and then analyzed using Microsoft Excel, IBM SPSS Statistics 25, and UCINET (Borgatti et al., 2002). To clarify, if three children were interacting with each other, three "dyads" of interaction were present, as the study aimed to analyze each child's individual interaction time with each of the other children.

Figure 1: An example of one of the "outer boxes" depicting children's interactions within a given minute in the classroom (the two-letter codes shown here are not the real ones used in the study). Here, CE would represent a child playing alone, and the lower-right box would typically represent five children doing a joint activity (dimensions of the actual boxes differed from that shown here).

AB CD EF GH	IJ KL
MN OP QR ST	UV WX YZ
AC BD	CE
FH GI	HJ KM LN
MO NP QS TV WY	XZ AD BE CF DG

Procedure

1319 minutes (~22 hours) of observational data were collected by one observer (SRW) over three months of twice-weekly nursery visits in 2019. The number of appearances of each child was counted to calculate the number of minutes for which they were observed playing alone, or with each of the other children, as well as their total observed time. The use of time-based scan sampling (Altmann, 1974) meant that the classroom was scanned every sixty seconds, and a record was made of who was interacting with whom (defined as children sitting within one

meter of each other, or clearly conversing or interacting with each other, such as chasing each other around the classroom). If sitting within one meter of each other but clearly facing away from each other and not interacting, children were not considered to be in a cluster together. Children's interactions were only recorded during 'free-play' time (as opposed to during organized activities) in order to be able to reflect the children's free choice of where to be in the classroom and with whom to play. During the study, parents were provided with a three-item questionnaire to ascertain their occupation, their child's race, and their child's number of siblings (no siblings were present in the classroom). The nursery staff supplied the children's first names, gender and age.

RESULTS

Not every child was present every day. Ergo, to account for differing total observation times per child, a divisor was created by summing the total observation times for each dyad, and dividing their interaction time by this figure. In our social network analysis of 31 children, every child had an individual score. This was based on the child's relationship with everyone else. Thus, if child A and child B were observed in the same room for 100 minutes (the divisor) but were recorded within the same box on the datasheet for 50 minutes (the numerator), then child A and child B would have a dyadic score of 0.5 for each other. Then, every child was assigned a score from 0-1 to indicate that child's score averaged out across the 30 other children in the sample (incidentally, there were no significant gender differences in that 0-1 score).

We found significant differences when we analyzed dyadic scores according to the child's attributes. A matrix of all dyadic scores was collated in a UCINET matrix to facilitate the Quadratic Assignment Procedure (QAP) (Whitbred, 2011). The responses to the questionnaires were collated to obtain five attributes for each child (age, gender, race, number of siblings, and parents' occupation). These were then converted into binary format (0 = a shared attribute; 1 = a difference) and individual matrices were formed for each attribute. As shown in Table 4, there were only two attributes – gender and race – that reached statistical significance. These two attributes also had the strongest correlations of all five attributes. Of these two attributes, gender was the stronger effect. To compare gender differences in racial homophily, two further QAP correlations were performed, resulting in a higher correlation for girls compared to boys, and a significant result for girls ($r = 0.191$, $p = .036$) but not boys ($r = 0.160$, $p = .078$). There were no significant effects for age, sibling number (categorized for analysis as 0, 1, or 2+), or parents' occupations.

Table 4. QAP Correlations

	Dyads	Age	Race	Occupation	Gender	Siblings
Dyads	1.000	-.025	0.139*	-0.008	0.236**	0.043
Age	-	1.000	0.048	-0.009	-0.033	-0.018
Race	-	-	1.000	-0.039	-0.037	0.046
Occupation	-	-	-	1.000	-0.041	0.046
Gender	-	-	-	-	1.000	0.008

* Significant at the .01 level. ** Significant at the .001 level.

Because the QAP (Borgatti et al, 2002; Whitbred, 2011) is aimed at analysing *pairs* rather than individuals, it is relevant to calculate that a group of 31 children allows for 465 pairings, as shown from equation (1).

$$n(n-1)/2 \quad (1)$$

Furthermore, the QAP involves an additional procedure whereby a large number of additional permutations of the rows and columns were generated (Borgatti et al, 2002; Whitbred, 2011): the results in Table 4 were derived when the Pearson correlation was compared against the distribution of correlations that result from the permutation process in UCINET.

DISCUSSION

We found that gender was the strongest predictor of social interaction. Contrary to expectations, race was the only other significant predictor. This supports Laursen's (2007) claim that physical traits are more strongly associated with homophily – although of course race and gender are not solely physical traits, but do usually include some physical markers. Girls appeared to exhibit higher rates of racial homophily than boys (in accordance with McPherson et al., 2001; contra Lee et al., 2007), inferred from our result that racial homophily was significant in girls but not in boys. However, the two groups were not directly compared.

Our findings appear to support the "Selection Hypothesis", which suggests that homophily occurs due to individuals selecting friends who resemble them physically, *assuming* them to have similar attitudes and characteristics – as opposed to selecting friends who they *know* to have similar attitudes and characteristics in the first place (McCormick et al.,

2015). In contrast, the Socialisation Hypothesis posits that individuals' behaviours and attitudes influence each other over time, and therefore the individuals become more similar as a result of spending time together (Prinstein & Dodge, 2008). If true, these two hypotheses taken together may imply that friends select each other based on perceived similarity, but then become even more similar over time due to increased exposure to one another's influence (Laursen, 2017). An example of a possible implication of this is as follows: if children play with specific toys on the basis of their gender, and go on to seek out others who enjoy those same toys, they are likely to make friends of the same gender. Consequently, the more time they spend together, the more their gender-typed play will increase. This results in not only the emergence of sex-segregation, but its perpetuation too, thus children who show higher rates of gender homophily show higher preferences for gender-based toys and activities (Martin & Fabes, 2001; Martin et al., 2005). Conversely, those who show less sex segregation demonstrate less gender-typed preferences (Serbin et al., 1994; as cited in Mehta & Strough, 2009). There are many possible positive and negative effects of such gender, race, and other group segregation – such as the encouragement and perpetuation of positive versus negative behaviours, and the increasing of outgroup stereotyping and dislike. For more in-depth discussion of these effects see: Mehta and Strough (2009) and Nangle et al. (2004).

Our study had some limitations. Unforeseen challenges in the ability to decipher who could be said to be interacting with whom occasionally required the observer's discretion. For example, when one child was interacting with another who was interacting with another (e.g. A→B→C) it became unclear how to assign the clusters, considering that child A and C were not necessarily interacting with each other and thus shouldn't technically appear in the same "inner box", but child B could not appear twice on one data sheet. Similarly, when a child A is standing behind a dyad B and C, watching them play (i.e. A→B↔C), then the interaction is actually one directional (between A and B↔C) and it was again unclear how to assign them to a cluster. Another issue is that occasionally a child would move between more than one friend/group in the space of one minute, making it difficult to capture that minute's data accurately, but attempts were made to capture each child's position as closely to the beginning of each minute as possible (see Altmann, 1974, pp. 258-260 for a discussion of timing issues in this type of sampling). A further problem was that instances of arguing or fighting were technically included as interactions (because of the physical proximity), which are likely not good measures of the interaction "choices" that imply friendship. Finally, our study was limited by its small sample size, and,

as there was only one observer (SRW) and video recordings were not taken, we were unable to calculate an inter-observer reliability score.

As we mentioned in our introduction, homophily can have some negative effects and therefore it could be useful to seek practical measures to reduce gender and racial homophily in children and instead encourage cross-group friendship formation. These include grouping students of different backgrounds together to work towards common goals, designing activities that emphasize similarities among students of different backgrounds, structuring activities that allow opportunities for boys and girls to play/learn together, and reducing the emphasis on gender as a social grouping category (McCormick et al., 2015). In fact, Maccoby and Jacklin (1987) claimed that gender labelling itself is what pushes girls and boys into gender-stereotyped behaviours. An equivalent claim can likely be made for racial labelling. Our results, which show evidence of significant race- and gender-based homophily in young children, suggest that such practical steps should be taken as early as possible.

REFERENCES

Altmann, J. (1974). Observational study of behavior: sampling methods. *Behaviour*, *49*, 227-266. https://doi.org/10.1163/156853974X00534

Bobbitt-Zeher, D., & Downey, D. B. (2013). Number of siblings and friendship nominations among adolescents. *Journal of Family Issues, 34*, 1175–1193. https://doi.org/10.1177/0192513X12470370

Borgatti, S.P., Everett, M.G., & Freeman, L.C. 2002. *Ucinet 6 for Windows: Software for Social Network Analysis.* Harvard, MA: Analytic Technologies.

Connolly, P. (2004). *Boys and schooling in the early years.* London, UK: Routledge.

Crystal, D. S., Killen, M., & Ruck, M. (2008). It is who you know that counts: Intergroup contact and judgments about race-based exclusion. *British Journal of Developmental Psychology*, *26*, 51-70. https://doi.org/10.1348/026151007X198910

Eivers, A. R., Brendgen, M., Vitaro, F., & Borge, A. I. H. (2012). Concurrent and longitudinal links between children's and their friends' antisocial and prosocial behavior in preschool. *Early Childhood Research Quarterly*, *27*, 137–146. https://doi.org/10.1016/j.ecresq.2011.05.001

Holder, M. D., & Coleman, B. (2015). Children's friendships and positive well-being. In Demir, M. (Ed.), *Friendship and happiness: Across the lifespan and cultures* (pp. 81-97). Dordrecht: Springer. https://doi.org/10.1007/978-94-017-9603-3_5

Karpiak, C., Buchanan, J., Hosey, M., & Smith, A. (2007). University students from single-sex and coeducational high schools: differences in majors and attitudes at a catholic university. *Psychology of Women Quarterly, 31*, 282–289. https://doi.org/10.1111/j.1471-6402.2007.00371.x

Laghi, F., Baiocco, R., Cannoni, E., Di Norcia, A., Baumgartner, E., & Bombi, A. S. (2013). Friendship in children with internalizing and externalizing

problems: a preliminary investigation with the Pictorial Assessment of Interpersonal Relationships. *Children and Youth Services Review*, *35*, 1095-1100. https://doi.org/10.1016/j.childyouth.2013.05.007

Laursen, B. (2017). Making and keeping friends: the importance of being similar. *Child Development Perspectives*, *11*, 282–289. https://doi.org/10.1016/j.childyouth.2013.05.007

Lee, L., Howes, C., & Chamberlain, B. (2007). Ethnic heterogeneity of social networks and cross-ethnic friendships of elementary school boys and girls. *Merrill-Palmer Quarterly, 53*, 325–346. https://www.jstor .org/stable/23096124

Maccoby, E. E., & Jacklin, C. N. (1987). Gender segregation in childhood. *Advances in Child Development and Behavior*, *20*, 239–287. https://doi.org/10.1016/S0065-2407(08)60404-8

Martin, C., & Fabes, R. (2001). The stability and consequences of young children's same-sex peer interactions. *Developmental Psychology, 37*, 431–446. https://doi.org/10.1037/0012-1649.37.3.431

Martin, C., Fabes, R., Hanish, L., & Hollenstein, T. (2005). Social dynamics in the preschool. *Developmental Review, 25*, 299–327. https://doi.org/10.1016/j.dr.2005.10.001

McCormick, M. P., Cappella, E., Hughes, D. L., & Gallagher, E. K. (2015). Feasible, rigorous, and relevant: validation of a measure of friendship homophily for diverse classrooms. *The Journal of Early Adolescence*, *35*, 817–851. https://doi.org/10.1177/0272431614547051

McDonald, K. L., Dashiell-Aje, E., Menzer, M. M., Rubin, K. H., Oh, W., & Bowker, J. C. (2013). Contributions of racial and sociobehavioral homophily to friendship stability and quality among same-race and cross-race friends. *The Journal of Early Adolescence*, *33*, 897–919. https://doi.org/10.1177/0272431612472259

McGill, R. K., Way, N., & Hughes, D. (2012). Intra- and interracial best friendships during middle school: links to social and emotional well-being. *Journal of Research on Adolescence*, *22*, 722–738. https://doi.org/10.1111/j.1532-7795.2012.00826.x

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: homophily in social networks. *Annual Review of Sociology*, *27*, 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Mehta, C. M., & Strough, J. (2009). Sex segregation in friendships and normative contexts across the life span. *Developmental Review, 29*, 201–220. https://doi.org/10.1016/j.dr.2009.06.001

Nangle, D. W., Erdley, C. A., Zeff, K. R., Stanchfield, L. L., & Gold, J. A. (2004). Opposites do not attract: social status and behavioral-style concordances and discordances among children and the peers who like or dislike them. *Journal of Abnormal Child Psychology; New York, 32*(4), 425–434. https://doi.org/10.1023/b:jacp.0000030295.43586.32

Pahl, R., & Spencer, L. (2004). Personal communities: not simply families of 'fate' or 'choice'. *Current Sociology*, *52*, 199-221. https://doi.org/10.1177/0011392104041808

Prinstein, M. J., & Dodge, K. A. (2008). *Peer influence processes among youth.* New York, NY: Guilford.

Rubin, K.H., Coplan, R. Chen, X., Bowker, J.C., McDonald, K. & Heverly-Fitt, S. (2015). Peer relationships in childhood. In M. H. Bornstein and M.E. Lamb (Eds.), *Developmental science: An advanced textbook* (7th ed., pp. 591-649). New York: Psychology Press.

Scott, J. (2000). *Social network analysis: A handbook*. London: Sage.

Wang, Y., Palonen, T., Hurme, T.-R. & Kinos, J. (2019). Do you want to play with me today? Friendship stability among preschool children, *European Early Childhood Education Research Journal, 27*(2), 170-184, https://doi.org/10.1080/1350293X.2019.1579545

Whitbred, R. (2011). Quadratic assignment procedure (QAP). In G. A. Barnett (Ed.). *Encyclopedia of social networks* (pp. 733-734). Thousand Oaks, CA: Sage. https://dx.doi.org/10.4135/9781412994170.n301

Yuger, J. M. & Shapiro, E. S. (2001). Elementary children's school friendship: a comparison of peer assessment methodologies. *School Psychology Review*, *30*, 568-585. https://dx.doi.org/10.1080/02796015.2001.12086135

Acknowledgements: the authors would like to thank the editor and reviewers, Miss Sarah Vipond and the children and other staff from the nursery, Laura Moody, Farah Hassan, Suzanne Magnus, Donna Weisz, and Dr. Jackie Meredith.

Dedicated to the memory of Dr. Jackie Meredith (1957-2022).

The authors declare that there is no conflict of interest.