Assessing the degree of urbanisation using a single-item self-report measure: a validation study

Lotte P. Brinkhof, a,b,c K. Richard Ridderinkhof, a,b,c Harm J. Krugers, b,c,d Jaap M. J. Murre a,b,c and Sanne de Wit a,b,c

Department of Psychology, Faculty of Behavioural and Social Sciences, University of Amsterdam, Amsterdam, Netherlands; Centre for Urban Mental Health, University of Amsterdam, Amsterdam, Netherlands; Amsterdam Brain & Cognition (ABC), University of Amsterdam, Amsterdam, Netherlands; Swammerdam Institute for Life Sciences, Cognitive Neuroscience, Faculty of Science, University of Amsterdam, Amsterdam, Netherlands

ABSTRACT

The differential impact of rural versus urban residence on mental health remains a controversial topic that requires more in-depth investigations. This calls for a valid and easy measure to assess the degree of urbanisation. The purpose of the present study was to determine the utility of a single-item self-report measure (SIDU) as a tool to classify areas along the rural-urban continuum. The validity of the SIDU was assessed by comparing its scores (1–7) to a commonly used objective surrogate measure of the degree of urbanisation (i.e. surrounding address density, SAD) in two independent older adult samples (A: N = 36, 65+; B: N = 121, 55+). SIDU scores approximated SAD scores, with r = .77 to 0.82, (A), and r = .79 to 0.83 (B). A SIDU threshold score of 6 most accurately distinguished extremely urbanised areas from other areas. Altogether, our findings suggest that SIDU scores could be used as proxy of SAD. Since self-report leaves room for the consideration of additional aspects that confer an urban settlement, this single-item scale may be even more comprehensive, and circumvents the collection and handling of highly sensitive location data when the primary goal is solely to distinguish urbanisation subgroups.

Introduction

More than half of the world’s population currently lives in urban settings and this number is expected to rise to nearly 70% by 2050 (United Nations 2019). On the one hand, there are clear benefits of urban living, such as opportunities for education, work and social interactions, and (health) amenities and entertainments in close proximity. On the other hand, it also poses significant challenges that have been shown to put urban residents at greater risk of developing mental health problems (e.g. common mental disorders such as anxiety, depression, addiction, as well as psychosis and loneliness; see van der Wal et al. 2021 for review). This may particularly apply to those individuals with increased (genetic) vulnerability (e.g. psychosis sensitivity, low social capital or socio-economic status, urban upbringing, early life stress; Galea et al. 2011; Lederbogen et al. 2011; Heim and Binder 2012; Gruebner et al. 2017), for whom the urban life could act as a trigger (Lecic-Tosevski 2019). At the same time, individual differences in predispositions or characteristics that determine one’s capacity for resilience (i.e. successful adaptation in the face of stressors) may also determine the extent to which someone will either flourish in the city or...
struggle with the challenges it poses (Smith et al. 2013; Christodoulou et al. 2017). Yet, the exact nature of such associations remains poorly understood, and further research is needed to improve our understanding.

The relevance of urbanisation levels for mental health calls for valid and reliable measurement tools to define, measure, and compare degrees of urbanisation across settings. Urbanisation is typically defined as the percentage of a given population living in areas classified as urban (ChaoLIN 2020). Yet, there are multiple definitions of what constitutes an urban settlement, and there is no universal consensus on the criteria for whether an area should be defined as urban or not (Vlahov and Galea 2002). Nonetheless, minimum population thresholds, and/or population density, or some derivative thereof, have often been used as surrogates to classify areas along the rural-urban continuum (Champion and Hugo 2017; United Nations 2019; ChaoLIN 2020), either in isolation or in combination with other criteria (e.g. Dahly and Adair 2007; Jones-Smith and Popkin 2010; Allender et al. 2011; Novak et al. 2012). In the Netherlands, the so-called degree of urbanisation (or urbanisation grade) is often assessed by looking at the surrounding address density (SAD) of a neighbourhood, district or municipality (Den Dulk et al. 1992). Areas are considered to be ‘extremely urbanised’ when the average number of addresses per square km within a circle with a radius of one km is 2500 or more. Remaining categories are: ‘strongly urbanised’ (1500 to 2500 addresses/km²), ‘moderately urbanised’ (1000 to 1500 addresses/km²), ‘hardly urbanised’ (500 to 1000 addresses/km²) and ‘not urbanised’ (fewer than 500 addresses/km²). While such measures are obviously objective and widely accepted, their applicability is limited by the fact that linking objective environmental estimates to other research data requires personal, and potentially highly sensitive, information on participants’ location of residence (e.g. postal codes). Ever since the General Data Protection Regulation (GDPR) has been introduced to protect data and privacy of individuals in the European Union, researchers have been required to adhere to the principle of data minimization. This comprises that processing of personal data is not allowed if one can achieve the same goal without the concerning data. Hence, research into the interplay between mental health and urbanisation would benefit from an alternative measure of the degree of urbanisation that does not require precise information about the location of residence. This is a particularly important issue in the older population, as the chances of successfully tracing personal data back to an individual are considerably higher for individuals of higher age (e.g. there may be only one 94-year old man living in a certain village in the Netherlands).

One way to accomplish this is to rely on subjective reports. While there has been an ongoing debate on the reliability of subjective measures, some preliminary evidence suggests that individuals’ perceptions of the degree of urbanisation can approximate objective measures (Dijkstra and Papadimitriou 2020; Castle and Tak 2021). Yet, these and other studies (Lin et al. 1997; Onega et al. 2020) also illustrated that forcing individuals to match their living environment to one specific label (e.g. urban, suburban, large rural, small rural, isolated) can lead to ambiguity about the category to which they belong, especially for those who reside in sub-urban or sub-rural areas (also see: Kolko 2015; Bucholtz and Kolko 2018). Nowadays, many rural areas experience factors that are traditionally associated with urban environments, and vice versa, resulting in an increased blurring of rural-urban distinctions (Champion and Hugo 2017). Hence, a bipolar rural-urban scale, allowing for a fine-grained classification of individuals along the rural-urban continuum based on individual perceptions, is particularly worthy of further investigation.

As previously stated by Pateman (2011), “Most people have a clear impression of what the cities, towns and countryside look like [in the UK], both physically and in terms of the lives of the people who live there”. Not only do people’s preferences underpinning their residential decision show that they are conscious about the favorable and less favorable characteristics of different types of

---

environments (Jansen, 2020; Thomas et al., 2015), but they also align with the characteristics that are generally used to objectively distinguish urban from rural areas (e.g., (population) density, number of/proximity to amenities; Jones-Smith & Popkin, 2010; Vlahov & Galea, 2002). Moreover, similar factors are taken into account by individuals when defining their living environment by the degree of urbanisation (Bucholtz et al., 2020), and objective neighborhood conditions often correspond to individual perceptions (Ellaway et al., 2001; Sampson & Raudenbush, 2004). Altogether, these findings suggest individuals should indeed be capable of classifying their own living environment along a rural-urban continuum.

The present study set out to examine the validity of a single-item self-report measure of the degree of urbanisation (SIDU), and to assess whether it can serve as an easy, minimally burdensome alternative for the SAD. Specifically, our goal was to identify whether the SIDU could be used to distinguish individuals along the rural-urban continuum based on a 7-point scale, ranging from not urban at all/rural to very urban. Such a brief instrument may substantially alleviate participant burden in comparison with multi-item measures (e.g. Jones-Smith and Popkin 2010). The validity of this scale was assessed in two independent samples of older adults (N = 36, 65+; N = 121, 55+), living in areas ranging from predominantly rural to extremely urbanised. We expected to find a moderate to strong relationship between the objective and self-report measure, underscoring its potential as a mean to investigate the effects of different environmental surroundings on, for instance, mental health and functioning.

Methods

Sample characteristics and design

Data were derived from two independent studies. Data of the first sample (A) originate from an as-of-yet unpublished study (approved by the local ethics committee of the University of Amsterdam (UvA), 2020-CP-12248; N = 36, 7 males, 65–83 years, M = 71.1, SD = 4.43). Based on the outcomes for sample A, a posteriori power analysis confirmed that 36 participants sufficed to yield high power (.940). Yet, we wanted to confirm this outcome in a substantially larger sample. Thus, for the second sample (B), data from a larger study on successful aging and resilience were used (N = 121, 50 males, 55–90 years, M = 69.1, SD = 7.06; see Brinkhof et al. 2021 for study protocol; approved by the local ethics committee of the UvA, 2020-DP-12556; 2021-DP-13798). In both studies, participants were only eligible to participate if they were free from any (severe) memory impairments.

Design

In both samples, participants were asked to enter their postal code, and subsequently to subjectively rate the degree of urbanisation of their living environment using the SIDU (see Materials). These were followed by other questionnaires that are beyond the scope of the current study. All participants provided informed consent for the collection and (anonymous) use of their provided data.

Materials

Objective measure of the degree of urbanisation

The objective measure of the degree of urbanisation was operationalized as the surrounding address density (SAD), as previously described: the mean number of addresses per km² within a circle with a radius of one km. These values were used to assign individuals to five urbanisation grade categories: (1) not urbanised, (2) hardly urbanised, (3) moderately urbanised, (4) strongly urbanised, and (5) extremely urbanised. Data were provided by Statistics Netherlands, an autonomous governmental institution that gathers and offers a wealth of
(openly available) data about the Netherlands, and retrieved at the neighbourhood level (sample A: 2020; sample B: 2021; Statistics Netherlands 2020, 2021). Subsequently, scores were aggregated to mean values for each postal code area, and subsequently linked to participants’ four-digit postal codes. Because of the aggregation, non-discrete category values were also possible (e.g. 4.1).

Of note, in the Netherlands, four-digit postal code areas (average area size: 8.9 km$^2$) are geographically delineated areas within municipalities (average size: 117 km$^2$) and include on average 2027 households (data from 2020; Statistics Netherlands 2020).

**Single-Item self-report measure of the degree of urbanisation (SIDU)**
The SIDU comprised the following question: ‘Please indicate how urban your living environment is on a 7-point scale from (1) not urban at all to (7) very urban’. For clarity, the ends of this bipolar scale were supported by additional labels. Option 1 was labelled with ‘hamlet’ (tiny village in rural area) and option 7 was supported with examples of the bigger cities within the Netherlands (‘city centre of Amsterdam, Rotterdam, Utrecht or The Hague’).

**Statistics**
Statistical operations were carried out in R (R Core Team 2020), with alpha set at 0.05. In both samples, the relationships among the SIDU on the one hand and both the SAD and urbanisation grade categories on the other were evaluated using Pearson’s correlation coefficients ($r$), or Spearman’s correlation coefficient ($r_s$) in case of non-normal residuals, as established with the Shapiro-Wilk normality test. A Bonferroni-correction was applied to correct for multiple testing ($\alpha/4 = 0.0125$). Sample B was also subjected to some exploratory analyses, where two different age subgroups (<70; ≥70), and men and woman were evaluated separately (with $\alpha/8 = 0.0063$).

To further assess the coherence between the SIDU and objective surrogate measure(s), we also examined the sensitivity (true positive rate) and specificity (true negative rate) of different SIDU threshold points for extremely urbanised areas, defined with the objective SAD (2500 or more addresses/km$^2$). This allowed us to determine how the maxima of both scales were associated with one another. The threshold score showing the most balanced combination of sensitivity and specificity (i.e. highest sum) was considered to most accurately distinguish individuals living in extremely urbanised areas from the rest. For each threshold point, the positive predictive value (PPV; proportion of individuals who actually were living in extremely urbanised areas according to the objective guidelines out of all individuals with score above the threshold) and negative predictive value (NPV; proportion of individuals who actually were not living in extremely urbanised areas according to the objective guidelines out of all individuals with a score below the threshold) were calculated as well. A comparable analysis for the minima was not performed because of the relatively small number of participants who had a SAD of 500 or less addresses/km$^2$, which would most likely result in unreliable findings.

In continuation of the previous analysis, using a receiver operating characteristic (ROC) curve analysis, we assessed the area under the curve (AUC), reflecting the probability that residents from extremely urbanised areas scored higher on the SIDU than other individuals (Metz 1978; Hanley and McNeil 1982). The ROC plots the true positive rate against the false positive rate, and represents the accuracy of classification of, in this case, residents from extremely urbanised areas as scoring high on SIDU. An AUC near to 1 or to 0, indicates high and low classification accuracy, respectively.
**Results**

Our results show that participants were able to accurately judge the degree of urbanisation of their neighbourhood (see Figure 1). In both samples, and most convincingly in the larger sample, it appeared that the relationship between the SIDU and SAD was best described as an exponential relationship. Participants that passed a certain verge of SAD (as depicted with the red dashed line in Figure 1, left panels) tended to score relatively similar on the SIDU. Indeed, this is in accordance with the categorization guidelines of Statistics Netherlands, where all areas with 2500 surrounding addresses or more are considered to be extremely urbanised. By evaluating the sensitivity and specificity for possible SIDU threshold scores, we confirmed that a score of 6 most accurately distinguished *extremely urbanised* from *not urbanised* - *strongly urbanised* areas, in both samples (as reflected in the highest sum of sensitivity and specificity; see Table 1). Moreover, we found that the probability that residents living in extremely urbanised areas had higher SIDU scores than other individuals was high, with AUCs of 0.993 (95% CI 0.970–1) and 0.896 (95% CI 0.940–0.947) for the sample A and B, respectively.

After transforming the exponentially distributed SAD scores to log values, a strong linear relationship with SIDU was established for sample A, $r(34) = 0.82$ (95% CI 0.67–0.90), $p < 0.001$, and B, $r(119) = 0.83$ (95% CI 0.76–0.88), $p < 0.001$ (Figure 1, left panels). In addition, the urbanisation grade categories strongly correlated with the SIDU in both samples, with $r_s(34) = 0.77$ (95% CI 0.58–0.87), $p < 0.001$, and $r(119) = 0.79$ (95% CI 0.72–0.85), $p < 0.001$ (Figure 1, right panels). Finally, no major age- or gender-related influences on participants' estimation accuracy were found (Table 2).
Table 1. Sensitivity, specificity, PPV, and NPV for different single-item self-report measure (SIDU) threshold scores for sample A and B.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: N = 36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>0.00</td>
<td>0.81</td>
<td>NA</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>0.00</td>
<td>0.81</td>
<td>NA</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>0.00</td>
<td>0.81</td>
<td>NA</td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>0.29</td>
<td>0.85</td>
<td>1.00</td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>0.86</td>
<td>0.97</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>6</strong></td>
<td><strong>0.90</strong></td>
<td><strong>1.00</strong></td>
<td><strong>1.00</strong></td>
<td><strong>0.70</strong></td>
</tr>
<tr>
<td>7</td>
<td>0.00</td>
<td>1.00</td>
<td>NA</td>
<td>0.19</td>
</tr>
<tr>
<td>B: N = 121</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>0.00</td>
<td>0.38</td>
<td>NA</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>0.13</td>
<td>0.41</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>0.23</td>
<td>0.44</td>
<td>1.00</td>
</tr>
<tr>
<td>4</td>
<td>0.98</td>
<td>0.36</td>
<td>0.48</td>
<td>0.96</td>
</tr>
<tr>
<td>5</td>
<td>0.93</td>
<td>0.67</td>
<td>0.63</td>
<td>0.94</td>
</tr>
<tr>
<td><strong>6</strong></td>
<td><strong>0.76</strong></td>
<td><strong>0.87</strong></td>
<td><strong>0.78</strong></td>
<td><strong>0.86</strong></td>
</tr>
<tr>
<td>7</td>
<td>0.00</td>
<td>1.00</td>
<td>NA</td>
<td>0.62</td>
</tr>
</tbody>
</table>

PPV: positive predictive value; NPV: negative predictive value. Bold faced values reflect the threshold score showing the most balanced combination of sensitivity and specificity.

Table 2. Relationships between the single-item self-report measure (SIDU) and both the surrounding address density (SAD) and urbanisation grade categories, in Sample B.

<table>
<thead>
<tr>
<th>SAD a</th>
<th>Urbanisation grade categories</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;70 years old</td>
<td>n = 72, r(70) = 0.82 (95% CI 0.73–0.88), p &lt; 0.001</td>
</tr>
<tr>
<td>70+ years old</td>
<td>n = 49, r(47) = 0.84 (95% CI 0.73–0.91), p &lt; 0.001</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>n = 71, r(69) = 0.82 (95% CI 0.72–0.88), p &lt; 0.001</td>
</tr>
<tr>
<td>Men</td>
<td>n = 50, r(48) = 0.84 (95% CI 0.73–0.91), p &lt; 0.001</td>
</tr>
</tbody>
</table>

aThe relationship between the SAD and SIDU was best described as an exponential function in this sample. Hence, the relationship between log(SAD) and SIDU is reported here.

Discussion

The proportion of individuals living in urban settings is projected to increase tremendously, with a large share of older adults (World Bank Group 2021). In order to study the influence of the degree of urbanisation on mental health and functioning, as well as other potentially interesting variables, the field would benefit from a short and reliable, yet privacy-insensitive, measure. In two independent older adult samples, we have established that the degree of urbanisation of one’s living environment can be assessed accurately using a single-item self-report measure (SIDU). Subjective reports on the SIDU closely approximated two variants of a commonly used objective surrogate measure, providing support for the use of this self-report measure as a proxy. Individuals scoring 6 or higher on the SIDU can be considered as living in extremely urbanised areas.

While previous studies have reported discordance between subjective and objective environment classifications (e.g. Lin et al. 1997; Kolko 2015; Onega et al. 2020), our results demonstrate that individuals are in fact highly capable of assessing the degree of urbanisation of their environment when using the SIDU classification measure. We argue that this is because this bipolar measure takes into account the increased blurring of rural-urban distinctions (Champion and Hugo 2017), and is able to capture small, yet meaningful, differences among areas that may not fit well into a specific category.
It is noteworthy that the two samples in the present study differed in residency, with the first sample including relatively more (extremely) urban residents, whereas less densely populated areas were better represented in the second sample. Nevertheless, correlation coefficients of the relationship between both the surrounding address density (SAD) and urbanisation grade categories, on the one hand, and scores on SIDU on the other were highly similar for both groups. This suggests that living environment did not have a meaningful effect on the capability to assess one’s urbanicity. This also seems to apply to age and gender differences, as we found relationships of comparable strengths for those aged <70 and ≥70 years, and for males and females. An outstanding question for future research is whether younger age groups would be similarly capable of assessing the degree of urbanisation of their living environment.

Moreover, it is important to consider the impact of potential self-report biases that may occur in individuals with poorer (mental) health, psychiatric disorders or cognitive impairment. A deteriorated mental state could potentially alter the way in which people perceive their environment and thereby also impact one’s objective evaluation of the degree of urbanisation (Ellaway et al. 2001; Lee and Waite 2018; van der Wal et al. 2021). In line with this possibility, many cross-sectional studies have reported associations between mental health or cognitive functioning and perceptions of the environment. We would therefore encourage researchers to validate the SIDU in psychopathological and cognitively impaired samples if they would consider using this subjective measure in such groups of individuals.

Ultimately, this scale may allow for generalization to other countries as well, which would be a major advantage considering that there is no global consensus on the criteria for assessing the degree of urbanisation (Vlahov and Galea 2002). We therefore encourage other researchers to extend this line of research by establishing whether the classification of areas along the rural-urban continuum can be accurately done by using a similar self-report measure in other countries and populations. Obviously, slight adaptations to the labels of the scale have to be implemented to make the scale appropriate for non-Dutch populations.

One important caveat to consider is related to the debate on whether single objective surrogate measures, such as the SAD, comprise an accurate reflection of the degree of urbanisation. Indeed, the use of such measures has been widely accepted, but it has been argued that a more comprehensive measure, based on the various dimensions of urbanisation, may be needed to avoid misclassification of environments (Judd et al. 2002; Dahly and Adair 2007; Jones-Smith and Popkin 2010; Cyril et al. 2013; Champion and Hugo 2017). Comparing SIDU scores to such multi-component measures may be a fundamental next step, providing an even broader validation of the self-report scale. In anticipation of this, such analysis could be expected to unveil an even stronger relationship among the subjective and objective measure. Subjective report leaves room for individual perceptions on what constitutes an urban area, which allows for the consideration of more aspects of urban density than SAD only. This may render the SIDU more comprehensive than the objective surrogate measure, thus approximating multi-component scores even more closely.

A final recommendation for future work arises from the exponential curve that was observed between the SAD measure and the SIDU. This relationship illustrated that individual’s subjective reports do not distinguish areas above a certain threshold (score of 6) of the degree of urbanisation. Indeed, this has also been acknowledged by the urbanisation grade category guidelines of Statistics Netherlands, where all areas with more than 2500 addresses/km² are considered to be extremely urbanised. Despite the fact that subjective and objective reports reveal highly similar patterns, it may be opportune if future work would more broadly evaluate the implications of neglecting potential differences among extremely urbanised areas that are not easily captured by such measures, but possibly still impact psychological functioning.

In sum, we found the single-item self-report measure of the degree of urbanisation (SIDU) adequate and useful in classifying older adults along the rural-urban continuum. This provides an easy alternative to objective, postal code reliant, measures and thereby safeguards
individuals’ privacy. This study aligns with previous successful applications of single-item instruments, such as with happiness (Abdel-Khalek 2006), depression (Skoogh et al. 2010; van Rijsbergen et al. 2014), satisfaction with life (Jovanović and Lazić 2018), and stress symptoms (Elo et al. 2003), and may inspire researchers from other fields to evaluate the potential of single-item measurements, rather than multi-item scales or other intensive measurements tools (e.g. structured interviews). It should be noted, however, that this measure is not optimally adequate when a main focus of a study is to gain detailed and precise information about specific environmental characteristics (e.g. availability of green space, noise, pollution). In that case, researchers would need information about the residency of participants in order to successfully link environmental data to individual answers on e.g. surveys. Otherwise, this single-item self-report measure may be a more favourable alternative, as it is easy to carry out and circumvents the collection and handling of highly sensitive data.

Acknowledgements
We thank both Prof. R. W. Wiers (Centre for Urban Mental Health & Dept. of Psychology, University of Amsterdam, Amsterdam, Netherlands) and Prof. C.L.H. Bockting (Centre for Urban Mental Health & Academic Medical Center, Dept. of Psychiatry, University of Amsterdam, Amsterdam, Netherlands) for providing comments on the manuscript, and Prof. R.W. Wiers for his valuable suggestions for the analysis as well.

Author contributions
L.B. designed the study protocol and arranged data collection of both samples; L.B. cleaned the data and performed the analyses and interpretation undersupervision of S.W., K.R.R., H.K., and J.M.. L.B. drafted the paper and S.W., K.R. R., H.K., and J.M. provided critical revisions. All authors approved the final version of the paper for submission.

Disclosure statement
No potential conflict of interest was reported by the author(s).

Ethics statements
This study was reviewed and approved by the local ethics committee of the University of Amsterdam (2020-CP-12248, 2020-DP-12556, 2021-DP-13798) and was conducted in accord with relevant laws and institutional guidelines. Informed consent from study participants was obtained online. Participants provided consent by clicking an acceptance box at the bottom of the agreement.

Open practices
All data and materials are made publicly available via FigShare and can be accessed at 10. 21942/uva.16836805.

Funding
This research was funded by the Centre for Urban Mental Health, a Research Priority Area at the University of Amsterdam.

References


