Using walkability measures to identify train stations with the potential to become transit oriented developments located in walkable neighbourhoods

Dana Jeffrey\textsuperscript{a}, Claire Boulange\textsuperscript{b}, Billie Giles-Corti\textsuperscript{b}, Simon Washington\textsuperscript{c}, Lucy Gunn\textsuperscript{b, *}

\textsuperscript{a} Melbourne School of Design, The University of Melbourne, Melbourne, 2010, VIC, Australia
\textsuperscript{b} Centre for Urban Research, RMIT University, Melbourne, 3000, VIC, Australia
\textsuperscript{c} School of Civil Engineering, The University of Queensland, Brisbane, 4072, QLD, Australia

A B S T R A C T

Classifying train stations into typologies is a useful way to simplify their complex characteristics to assess their potential to become Transit Oriented Developments (TODs). However, researchers are yet to fully explore the walkability of train station neighbourhoods. Walkable areas have many features but typically include high residential density, greater street connectivity, and mixed land uses. Such characteristics facilitate access to the train station and can indicate both the effectiveness and potential of a train station to function as a TOD. This research explores the walkability of 230 train stations in metropolitan Melbourne, Australia using 14 different walkability measures. A two-stage cluster analysis was employed to group the train stations to determine their degree of walkability. The train station typology was validated using train station patronage data by different transport modes. Three clusters were found: cluster 1 train stations were more walkable and generally located in inner city areas, whilst those in cluster 2 were the least walkable and were generally located in middle and outer suburban areas. Cluster 3 train stations had the most potential for development as a TOD, having similar walkability features to those in cluster 1 but more car parking facilities and local living destinations when compared with the least walkable cluster 2 train stations. These findings suggest potential for cluster 3 train stations to become TODs, particularly if residential densities were increased in the surrounding neighbourhood. The patronage data validated the cluster findings in that the most walkable cluster 1 train stations had the highest percentage of pedestrian entries. TODs offer a way for planners to increase public transit use by co-locating a variety of services, destinations, residences and places of employment. Our findings provide a typology useful for exploring development strategies for developing train stations into TOD as a means of managing population growth and creating healthy, liveable and sustainable cities.

1. Introduction

Transit Oriented Developments (TODs) co-locate high density residential housing with public transit and a variety of destinations such as retail shops, supermarkets, employment, health and community services, and are intended to encourage walking and produce community benefits (Higgins & Kanaroglou, 2016; Kamruzzaman et al., 2014; Vale, 2015). TODs facilitate local living by allowing people to live, work and play close to home with access to other destinations via transit (City of Melbourne, 2015; Commission on the Social Determinants of Health, 2008; Newman et al., 2015). TODs contribute to and promote a range of co-benefits including reduced single occupant vehicle use, road congestion, and transportation costs (Renne, 2009; Stutzer & Frey, 2008), as well as improved air quality (Woodcock et al., 2009), enhanced social cohesion, (Giles-Corti et al., 2015) and active lifestyles conducive of improved health and well-being (Higgins & Kanaroglou, 2016).

Typologies for transit stops and the surrounding area are commonly used to assist policymakers, planners and urban designers understand which TODs have similar characteristics, enabling further analysis to identify those train stations ripe for development and revitalization (Center for Transit-Oriented Development, 2010; Lyu et al., 2016; Reusser et al., 2008; Zemp et al., 2011). To develop typologies, TODs are typically assessed using node and place measures using Bertolini’s node-place model (Bertolini, 1999). This typology provides a framework for assessing the transit stop (‘the node’), the surrounding area (‘the place’), and the relationship between the two (Bertolini, 1999). Measures such as transit frequencies and diversity of transit services measure the accessibility of the node, whilst built environment measures such as street connectivity and employment density are used to assess the quality of the place (i.e., the neighbourhood). Under the node-place model, transit stops that balance node and place values are ‘balanced’ and have development potential because they match transit accessibility with demand, whilst unbalanced nodes require further intervention to optimise TOD outcomes (Bertolini, 1999).

Whilst the node-place model provides an understanding of the transit stop itself, it overlooks the importance of walking accessibility between the node and the place (Lyu et al., 2016; Schlossberg & Brown, 2004). Without this additional information, a future development may fail to achieve its full potential as a TOD. This shortcoming is
The walkability of the neighbourhood surrounding a transit stop is a key aspect of its success, as it facilitates convenient, safe and direct access to and from the transit node (Jacobson & Forsyth, 2008; Vale, 2015). The walkability of the neighbourhood surrounding a transit stop is a key aspect of its success, as it facilitates convenient, safe and direct access to and from the transit node (Jacobson & Forsyth, 2008; Vale, 2015), encouraging greater use of the transit stop and ensuring that TOD residents have access to more local amenities to meet their local needs. Successful TODs maximise access by active modes. A walkable neighbourhood can help achieve this, because it is characterised by well-connected streets, high residential densities, and proximate access to jobs, services and public transport (Giles-Corti et al., 2015). However, one key difference between a TOD and a walkable neighbourhood is that, although a walkable neighbourhood has access to transit, it is not necessarily centred on a transit stop. These urban environment features are often presented in the literature through Ewing and Cervero’s ‘5Ds’ framework which includes: Density, Diversity, Design, Destination accessibility and Distance to transit (Ewing & Cervero, 2010; Ewing et al., 2009; Giles-Corti et al., 2016). In combination, these features promote walking and cycling and lead to health and community benefits (Frank, 2005; Giles-Corti et al., 2015; Saelens et al., 2003).

Despite the importance of neighbourhood walkability to TOD development, only a few researchers have explored how the walkability of neighbourhoods surrounding transit stops could inform future TOD development (Schlossberg & Brown, 2004; Vale, 2015). This lack of attention may arise because a detailed examination and calculation of complex built environment features surrounding transit nodes goes beyond the broad definition of a TOD.

Existing TOD typologies to date have not fully accounted for this complexity. Hence, in this paper we propose and describe a new typology that better accounts for the complexity of walkability of the TOD’s neighbourhood and its role in the TOD’s success. This research seeks to develop and explore a train station typology based on walkability measures with the aim of providing refined and improved metrics that can contribute to identifying potential successful TOD developments. We develop and present a train station typology analysis using 14 objective walkability features measured within close proximity to 230 train stations in metropolitan Melbourne, Australia.

2. Literature review

2.1. Typology literature relating to walkability

Walkability within active living research primarily focuses on the area surrounding people’s homes (Leal & Chaix, 2011) with only a few researchers considering other important locations such as workplaces, neighbourhood/retail activity-centres (Gunn et al., 2017b) or transit stops (Schlossberg & Brown, 2004; Vale, 2015). A walkable environment increases access to these high use locations without the need for a motor vehicle (Giles-Corti et al., 2016; Higgins & Kanaroglou, 2016; Sallis et al., 2016) leading to numerous health, environmental and economic co-benefits.

In seeking to understand the features in such environments, active living researchers have used cluster and latent profile analysis to derive typologies. Both methods create homogenous groups (or clusters) based on a set of built environment measures that help characterize the areas of interest. Generally these studies use objective measures such as ped-sheds ratios, population density, land use mix, intersection density, cul-de-sac density, proximity to services, ratio of commercial area to total area, and various other density measures including density of businesses, bus stops, park types, recreational destinations, sidewalks, and cycle-ways (Charreire et al., 2012; Higgins & Kanaroglou, 2016; McCormack et al., 2012; McNally & Kulkarni, 1997). The work of Cerin et al. (2007), also incorporated the proportion of land use mix relating to residential, recreational, commercial or industrial land in their study using built environment features. Whilst Adams et al. (2011) used perceived built environment measures relating to residential density, land use mix, street connectivity, walking and cycling facilities, aesthetics, safety, transit stops and proximity to parks and recreational facilities from the Neighbourhood Environment Walkability Scale (NEWS) in a latent class profile analysis of individuals and their neighbourhood built environments.

Research typically finds that more walkable neighbourhoods and areas are associated with more walking (Cerin et al., 2007; Gunn et al., 2017b; McCormack et al., 2012). Furthermore, a common finding to date, is that more walkable areas are located in older more established areas, often located closer to the inner-city or downtown areas which typically have more commercial zoning, retail destinations and higher residential densities (Gunn et al., 2017b; McCormack et al., 2012).

2.2. Typology literature relating to TODs and commonly used measures for node and place

Understanding the node and place features of transit nodes can reveal their context-specific structure leading to improved selection, design and development of TODs (Vale, 2015). TOD typologies derived from cluster or latent class analysis assist in this process, as they simplify large datasets (based on individual transit nodes and their area-based features) into a small set of clusters based on similar features. This distillation helps policymakers and planners understand the current state of TODs relative to planning expectations, and can assist in identifying transit stops and their surrounding areas that require context-specific development (Higgins & Kanaroglou, 2016).

A variety of node-place measures that capture the features of transit stops are typically used in the TOD literature. Kamruzzaman et al. (2014) calculated six built environment measures for administrative areas based on 1734 Census Collector Districts (CCDs) in Brisbane, Australia. These included net employment density, net residential density, land use diversity, intersection density, cul-de-sac density and transport accessibility. Cluster analysis was then used to create a typology resulting in four clusters. However, by using CCDs, this research did not explicitly focus on train station precincts.

Other researchers have used similar measures for characterizing the area surrounding train stations. Using data from Switzerland, Reusser et al. (2008) used population, the number of primary and tertiary sector jobs, and land-use diversity within 700 m of 1684 train stations, and from questionnaires they identified number of full time jobs in education, distance from town centre and the presence of grocery, restaurants, pharmacies and florists as being important measures of place. Similarly, Zemp et al. (2011) used population and number of jobs within 700 m of 1700 train stations. Atkinson-Palombo & Kuby (2011) used measures based on parcels of residential, vacant, TOD-compatible and TOD-incompatible land and measures relating to people which included jobs, population density, household income, percentage of owner-occupied housing units, and percentage of people with a bachelor degree for their study of TOD potential conducted in Phoenix, Arizona. These authors also included various node measures for transit frequency and accessibility, which were combined using factor analyses prior to undertaking cluster analysis. Higgins & Kanaroglou (2016) used density, street connectivity, land use mix, development mix and a gravity measure of station interaction-potential based on population, employment and station-to-station transit time for exploring TOD typologies on 372 train stations in Toronto, Canada. They measured connectivity using ped-shed ratios based on 10 min of walking. Instead of using factor analysis and prior to undertaking the latent profile analysis, they used a latent class structure based on the two most correlated input variables of density and interaction potential. Despite arguing for the incorporation of measures relating to the walkability of an area, in a study in Lisbon, Portugal, Vale (2015) only included a ped-
shed ratio, measured as the area in a Euclidean buffer to that from a street network buffer, as an add-on to another 13 measures used in his node-place assessment and cluster analysis of train and ferry stations.

In contrast to these studies using cluster or latent class analysis, Schlossberg & Brown (2004) focused explicitly on the walkability of 11 train stations in Portland, Oregon using six place-based measures dichotomized into high and low categories measured at a quarter and half mile (400 m and 800 m) distance from each train station in their comparison of TOD sites. These included: quantity of accessible paths measured as the length of minor roads, quantity of impedance paths measured as the length of major roads, pedestrian catchment areas (i.e., ped-sheds), impediment pedestrian catchment areas (i.e., ped-shed based on area accessible once high volume, high speed roads are removed), and density of dead ends. Their analysis ranked the train stations on each of these measures and grouped the three best and worst train stations. The remainder of their analysis explores the walkability of two station areas as case studies using maps of the built environment variables. A major finding was that evaluating and understanding the node-place structure and how it relates to the transit-node, is key for facilitating the location of TODs.

Lyu et al. (2016) included a third aspect to the node-place model relating to orientation, which was used to explore the interrelation between the node and place. They explained that the area surrounding transit stations may not have adequate street connectivity to support walking or cycling or may have densities situated adjacent to the station instead of toward it, resulting in a TAD. Lyu et al. (2016) referred to these area-based aspects surrounding the transit node as “functional and morphological interrelations,” which have the potential to facilitate transit use but also increase the utility and value of the area surrounding the train station leading to improved transit services according to the land use feedback cycle (Wegener, 2004; Wegener & Fuerst, 1999). Measures included in the ‘orientation’ aspect include average block size, average distance from station to jobs, average distance from station to residents, length of paved footpath per acre, intersection density, and the walk score – features which are also related to the walkability of an area. Although using different terminology, Vale (2015) and Lyu et al. (2016) have essentially articulated the need to include the walkability of the TOD's neighbourhood into the node-place model (Bertolini, 1999; Vale, 2015).

2.3. TOD performance and validation

Node-place measures can be categorized as input and outcome measures. Input measures relate to the node and place explicitly whilst outcome measures relate to TOD use. TOD performance relates to whether a TOD is doing well on the outcome measures (Renne, 2009). In the literature, there is some confusion on whether outcome measures should be included in node-place indexes or within the creation of TOD typologies. Indeed, including outcome measures in typology analyses may strengthen the distinction between well performing and balanced TODs compared to those that are not.

For example, Reusser et al. (2008) used passenger frequency and vehicle miles travelled in their node-place index. However, Renne (2009) and Zemp et al. (2011) argued that such measures are actually outcome measures that arise from the station inputs. The distinction between these measures is important as it is the outcome measures that indicate whether a TOD is performing well as a transit stop.

Instead of directly incorporating outcome measures, several authors use them to validate their TOD typologies. Renne (2009), for example, argued that TOD performance can be validated in various ways according to a framework based on: (1) travel behaviour; (2) the economy; (3) the natural environment; (4) the built environment; (5) the social environment; (6) the policy context; and by comparing (a) TODs to other TODs; (b) TODs with non-TODs; or (c) TODs against regional averages. Kamruzzaman et al. (2014) used this approach when they validated their typology from a travel behaviour perspective using mode choice behaviour according to their four cluster TOD solution. Using a multinomial regression with mode choice as an outcome measure and the cluster solution as the exposure, they found that people living in non-TOD areas were less likely to use active modes of transport.

In this study, the focus was to explore the typology of metropolitan Melbourne's train stations based on measures of walkability. We built on previous research, and used a comprehensive and established set of walkability measures based on Cervero and Ewing's 5Ds framework (Cervero & Kockelman, 1997; Ewing & Cervero, 2001) that have been used widely in public health literature (Saelens & Handy, 2008). The specific measures chosen for this study have been found to be associated with walking and wellbeing outcomes (Gunn et al., 2017a; Gunn et al., 2017b; King et al., 2015; Davern et al. 2018). In validating our cluster solution and typology, we followed the approach of Higgins & Kanaroglou (2016) and measured the performance of the train stations in metropolitan Melbourne separately to the derivation of TOD typologies, using train station patronage data. We used the percentage of train station entries made by transport mode (Higgins & Kanaroglou, 2016).

3. Data and methods

3.1. Study area

The metropolitan train network in Melbourne, Australia was the case study for this research.

Melbourne is a large sprawling city of approximately 10,000 km², with a hub and spoke train network, and one of the largest tram networks in the world. Nevertheless, access to transit is limited because density overall is relatively low at 13 dwellings per hectare (Arundel et al., 2017). Melbourne's population is projected to reach 8 million by 2050. This rapid growth will require an additional 1.6 million residential dwellings to be built in the next 32 years (State Government of Victoria, 2017a; State Government of Victoria, 2017b). Until recently, Melbourne, like most Australian cities, has favoured urban sprawl as a development model (Newman & Kenworthy, 1999), resulting in high rates of car dependency, poor health outcomes and social inequities (Dodson & Sipe, 2008). However, recent state government planning has recognised the need to accommodate future population growth in established areas including the development of TODs (State Government of Victoria, 2017a; State Government of Victoria, 2017b).

Melbourne's train network currently consists of 17 metropolitan train lines (with 230 train stations), which converge on the central business district located at the centre of the city via an underground loop that supports three underground stations. The hub and spoke network form means each of the train lines converge in the city. To date, there are no train lines that traverse the city and distances between train lines increase in the middle and outer city areas making access to train lines difficult. A new metro tunnel connecting the western and south-eastern sections of the city is currently under construction and a suburban rail loop has recently been proposed by the current state government (State Government of Victoria, 2018). The tunnel will facilitate an uplift of services across the network by 2025. Until recently, Melbourne's train network has been predominantly above-ground, inter-connecting with the road network across the city. This network configuration decreases efficient travel for road and rail users. To mitigate and alleviate road congestion, in 2017–2018, 50 rail crossings were removed to streamline rail and road transit (State Government of Victoria, 2017a; State Government of Victoria, 2017b), with sky-rails being constructed on one of the most heavily used lines (i.e., Cranbourne-Pakenham).

The analysis presented here serves as a baseline measure for the walkability of the train stations on Melbourne's rail network. It offers a novel and unique opportunity to explore both the walkability of areas surrounding train stations in a generally low density sprawling city and...
the implications this might have for future train station planning, development and policymaking. Given the timing of data collection, it does not take into account if, and how, removal of level crossings or sky-rail sections might have affected the walkability of the relevant train stations neighbourhoods.

3.2. Data and methods

Two types of data were used in this study: (1) spatial data on which walkability measures were derived using Geographic Information Systems (GIS) analysis; and (2) train station patronage data by mode representing annual counts of train station entries.

The spatial data were acquired from a variety of government and non-government agencies including:

- Public Transport Victoria: public transport data (open-access datasets) (Public Transport Victoria, 2014)
- VicMap: street network, parcel boundaries and zoning data from (open-access datasets) (State Government of Victoria, 2012a; State Government of Victoria, 2012b; State Government of Victoria, 2012c)
- Pitney Bowes: business points data (commercial dataset) (Pitney Bowes Ltd, 2014)

GIS-techniques in ArcGIS and analytical tools in ArcMap were used to derive the walkability measures. A geodatabase was deployed in ArcMap to store and organise all spatial layers and tables of data collected for this study.

The train station patronage data was sourced from the Public Transport Victoria (PTV) open dataset “Station patronage data for Melbourne’s rail network”. Each station was surveyed by PTV in 2012 for one weekday (excluding Friday) to represent an average day across the network. The patronage data provides the count and proportion of train station entries made by mode (Public Transport Victoria, 2015).

3.3. Walkability measures

Fourteen walkability measures were derived and categorized according to the 5Ds framework for each buffer and train station (Table 1). As discussed earlier most design, diversity and density related measures have been used in previous typology research. However, some measures, including car parking and bike facilities, whilst considered important in the literature, are often not included owing to data availability issues (Chorus & Bertolini, 2011; Griffiths & Curtis, 2017). Inclusion of these variables is important because it maximises the potential of using transit for those living further away. In this study, we also included walkability measures relating to destination accessibility that have been found to be associated with walking (Gunn et al., 2017a; Gunn et al., 2017b; King et al., 2015) and wellbeing outcomes (Daven, Gunn et al., 2018). These measures are related to the concept of ‘local living’ which ensures access to essential destinations allowing people to live within their immediate neighbourhood (Badland et al., 2017). Living close to essential destinations and amenity is important in shifting travel behaviour toward more active modes and for alleviating road congestion. Locating TODs in walkable neighbourhoods would maximise use of the TOD and its amenities by local residents; and provides TOD residents with access to more local amenities required for daily living.

Prior to analysis all freeways were removed from the road network spatial layer, as they are inaccessible to pedestrians. Areas around train stations were defined using an 800 m road network buffer. This distance was chosen as it has been associated with transport walking measures such as walking trips and sufficient physical activity and has been used in other typology analyses (Gunn et al., 2017b; Vale, 2015).

3.4. Statistical analysis

Summary statistics for the 14 walkability measures were computed and a two-stage cluster analysis of Melbourne’s train stations using standardized built environment variables performed. Cluster analysis was chosen ahead of latent profile analysis because it is an intuitive and well-established method for processing large datasets quickly. Pearson correlations were calculated prior to conducting the cluster analysis to check for multicollinearity amongst the walkability measures.

First, hierarchical cluster analysis was employed to help determine the number of clusters. Hierarchical clustering typically starts with single objects which are combined to create clusters based on similarity. Clusters are then combined until one single cluster is achieved. The appropriate number of clusters is determined using the distance measure between cluster formations from the agglomeration schedule where a large jump in the agglomeration schedule is indicative of the appropriate number of clusters. In undertaking hierarchical cluster analysis several linking algorithms exist and were used here including Ward’s, Complete and Average Methods. Each method links objects in different ways and the three algorithms were used to mitigate linking problems that can occur because of each linking method. Nevertheless, Complete and Average linkage methods in particular have been cited as being appropriate for many styles of data (Hair Jr et al., 2014).

In the second stage, the K-means clustering algorithm was used as it is more robust to outliers and irrelevant clustering variables (Hair Jr et al., 2014). To validate the number of clusters, the Calinski-Harabasz pseudo F statistic was computed. Following the cluster analysis, Tukey post-hoc tests were performed to determine statistically significant differences between built environment measures by cluster.

3.5. Validating the cluster solution

The final cluster solution was mapped using GIS and a desk-top audit using Google maps was undertaken to validate the cluster analysis solution to local knowledge of the characteristics of Melbourne’s train stations and urban form.

A final validation of the cluster solution was undertaken using patronage data by mode and includes both descriptive and visual analyses presented as the percentage of train station entries made by walking overlaid onto the GIS map of the final cluster solution.

Statistical analyses were performed in R and final GIS maps were produced using the R Leaflet package. GIS built environment data manipulation was undertaken in ESRI ArcGIS Desktop version 10.3.1.

4. Results

4.1. Cluster analysis

Table 2 presents summary statistics and the three cluster solution for the built environment measures surrounding Melbourne’s train stations chosen based on Tukey post-hoc tests, the Calinski-Harabasz pseudo F statistic, and the overall interpretability of the results. Pearson correlations revealed no issues with multicollinearity.

Overall, three different clusters of train stations were identified. Compared with other clusters, cluster 1 train stations had the highest values for gross dwelling density, community resources, educational facilities, convenience stores, entertainment venues, destination diversity, and number of tram stops. These high values suggest that the area surrounding these train stations is highly walkable. Cluster 1 train stations had similar values to cluster 3 train stations for the connected node and ped-shed ratios, and access to sporting facilities. Whilst there are a high number of bus stops in the area surrounding cluster 1 train stations, there were fewer bus stops than the areas surrounding cluster 3 train stations. The number of bike facilities were also low, although similar in value to cluster 2 train stations. Finally, cluster 1 train stations had the lowest amount of car parking in the area surrounding the
Table 1
Built environment measures calculated within 800 m of a train station.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Definition</th>
<th>Data source</th>
<th>GIS methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design</td>
<td>The connected node ratio was calculated using the following formula:</td>
<td>Intersection data were derived from the VicMap Transport road centrelines (State Government of Victoria, 2012a).</td>
<td>The join command was used to count the number of intersection points in each buffer.</td>
</tr>
<tr>
<td></td>
<td>[ \text{Connected node ratio} = \frac{x}{y} ] Where ( x ) is the number of three or more ways intersection and ( y ) is the total number of intersections.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ped-shed ratio</td>
<td>Ratio of area within 800 m street network buffer to the area within 800 m Euclidean buffer.</td>
<td>Street network data were sourced from the VicMap Transport road centrelines (State Government of Victoria, 2012a).</td>
<td>The ArcGIS Network Analyst extension was used to define a 1600 m road network buffer (i.e., areas that encompass all accessible streets within specified cut-off distances) created around each train station.</td>
</tr>
<tr>
<td>Diversity</td>
<td>Land use mix was calculated as an entropy measure based: ( LUM = - \sum_{i=1}^{n} \left( \frac{p_i}{n} \right) \ln \left( \frac{p_i}{n} \right) ) Where ( i ) is an index for land use category, ( n ) is the total number of land uses included, and ( p ) is the area of each land use ( i ). Five land use categories were included in the calculation: residential, retail, commercial, industrial and “other”.</td>
<td>A customised land use layer was created where parcel boundaries (State Government of Victoria, 2012b), zoning data (State Government of Victoria, 2012c) and geocoded business points data (Pitney Bowes Ltd, 2014) were combined. The parcel boundaries data and zoning data were readily available nationally while the business points were available on a fee-for-service basis.</td>
<td>The land use mix measure was calculated using the entropy formula, adapted from Frank et al. (Frank et al., 2005) (Frank et al., 2002) and is the same as the one used by Christian and colleagues (Christian et al., 2011). The intersect command was used to calculate the area of each land use type within each buffer.</td>
</tr>
<tr>
<td>Car parking</td>
<td>Land area dedicated to car parking (m²).</td>
<td>Car parking data were sourced from Public Transport Victoria data (Public Transport Victoria, 2012).</td>
<td>The join command was used to measure the land area dedicated to car parking in each buffer.</td>
</tr>
<tr>
<td>Density</td>
<td>Gross dwelling density is calculated as the sum of the commercial dwellings and residential area divided by the area within each buffer.</td>
<td>Dwelling count data were sourced from the meshblock-level census data (Australian Bureau of Statistics, 2011).</td>
<td>The join command was used to sum the number of dwellings in each buffer.</td>
</tr>
<tr>
<td>Destination</td>
<td>Community resources: Count of all the libraries and post offices.</td>
<td>Destinations data were sourced from geocoded business points data (Pitney Bowes Ltd, 2014).</td>
<td>The join command was used to retrieve the destination points within the boundaries of each buffer.</td>
</tr>
<tr>
<td></td>
<td>Education: Count of all the primary and secondary schools, childcare, kindergartens and creches.</td>
<td>Summary statistics were used to compute the number of destinations in each buffer.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Convenience stores: Count of all the supermarkets, specialty food stores, fishmongers, butchers, poultry stores, fruit and vegetable shops, milk bars, petrol stations and news-agenes.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Entertainment venues: Count of all the cinemas, theatres, art galleries, gaming venues.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sporting facilities: Count of all the tennis courts, swimming pools, gyms, ovals and fields.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Destination diversity: Sum of the presence/absence of each of the following 10 destination categories in a buffer:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. Specialty food stores: Fruit and vegetable grocers, fish stores, poultry stores and butchers</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Supermarket</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. Convenience store: Convenience stores, petrol stations and newsgagents</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5. Transport: Bus and Tram stops</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6. Pharmacy</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7. Medical: General practitioners, maternal health clinics, community health centres</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8. Dentist</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9. Library</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10. Post Office</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued on next page)
train station, which is commensurate with the high walkability of the surrounding area.

Cluster 2 train stations had the lowest values of all clusters for the connected node and ped-shed ratios, gross dwelling density, sporting facilities, destination diversity and bus stops. The values for community resources, education, convenience stores, entertainment venues, tram stops and bike facilities were also lower than those found for cluster 1 train stations, but were not statistically different to those for cluster 3 train stations. The amount of car parking for cluster 2 train stations was between that of the other clusters. These findings suggest cluster 2 train stations are located in low walkable neighbourhoods.

The walkability attribute values of cluster 3 train stations were low and similar to cluster 2 stations for community resources, education, convenience stores, entertainment venues and tram stops. However, cluster 3 stations had the highest values for bike facilities and bus stops and had the highest values for car parking. The gross dwelling density and destination diversity values for cluster 3 train stations were between the other clusters.

Notably, there were no statistical differences between train station clusters for the land use mix entropy measure, with similar values found for each cluster of train stations. Land use mix is generally found to be associated with walking behaviour (Christian et al., 2011). However, in the context of train station neighbourhoods there did not appear to be sufficient variation in land use mix across any of the train station neighbourhoods. All three clusters differed statistically in their amount of car parking.

### 4.2. Location of train station clusters

Fig. 1 shows the location of the different train stations clusters. This map illustrates that cluster 1 train stations - typically located in high walkable neighbourhoods - are predominantly located in the inner and middle suburbs. Cluster 3 train stations - found in neighbourhoods with some but not all walkable features - are typically located in the middle suburbs. Whilst cluster 2 train stations are found in low walkable neighbourhoods and are typically located in the outer suburbs and regional areas of metropolitan Melbourne.

### Table 2

Descriptive statistics for the three cluster solution for built environment measures around train stations.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transit accessibility&lt;sup&gt;a&lt;/sup&gt;&lt;sup&gt;b&lt;/sup&gt;</td>
<td>The join command was used to retrieve the number of bus stop, tram stop and bike infrastructure points within the boundaries of each buffer.</td>
</tr>
</tbody>
</table>

<sup>a</sup> Transit accessibility is used in place of distance to transit.

<sup>b</sup> Bike parkiteers are large sheds secured through access cards capable of storing many bikes.

### Table 3 (continued)

<table>
<thead>
<tr>
<th>Measures</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transit accessibility&lt;sup&gt;a&lt;/sup&gt;&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Bus stops data were sourced from Public Transport Victoria data (Public Transport Victoria, 2014).</td>
</tr>
<tr>
<td>Bus stops</td>
<td>The join command was used to retrieve the number of bus stop, tram stop and bike infrastructure points within the boundaries of each buffer.</td>
</tr>
</tbody>
</table>

<sup>a</sup> Transit accessibility is used in place of distance to transit.
4.3. Train station entry patterns – validation of the cluster solution

Table 3 shows the percentage of train station entries by transport mode and cluster. Notably, 82% of train station entries were made by active transport modes: 56% by walking, 25% by public transport; and 0.68% by cycling.

Cluster 1 train stations located in highly walkable neighbourhoods were accessed by the highest percentage of pedestrians (64%), whilst the remaining clusters were accessed by fewer but also high percentages of pedestrians relative to other active transport modes (41% for cluster 2, and 43% for cluster 3).

Table 3 also shows that train station entries to cluster 1 train stations made by either train and tram were higher at 16% and 8% respectively than the other two clusters where the percentages were relatively small being less than 5%. Only 3% of train station entries to cluster 1 train stations were made by bus. The higher percentages of train and tram entries compared to bus entries for cluster 1 train stations is unsurprising as the tram network is predominantly based in the inner-city, and train lines become more disperse with limited service variety in middle and outer areas negating this mode as a connecting service and entry point. Supporting this finding is the jump in train station entries made by bus for cluster 2 and cluster 3 train stations at 11% and 16% respectively. Notably, the percentage of train station entries made by cycling was relatively small for all clusters, however the percentages were slightly higher for cluster 2 and 3 train stations at 0.8% and 1.0% respectively when compared with those made to cluster 1 train stations (0.5%).

Overall, 18% of all train station entries were made by private motor vehicle. However only 8% of drivers accessed cluster 1 train stations, with the largest percentage made to cluster 2 train stations (45%) followed by cluster 3 train stations (35%).

Fig. 2 Panels (a) – (d) overlay the cluster solution with the

---

Table 3

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Transport mode</th>
<th>Walk</th>
<th>Cycle</th>
<th>Bus</th>
<th>Train</th>
<th>Tram</th>
<th>Active Transport</th>
<th>Private Motor Vehicle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>63.94</td>
<td>0.54</td>
<td>3.19</td>
<td>16.09</td>
<td>8.43</td>
<td>28.25</td>
<td>7.58</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>40.53</td>
<td>0.77</td>
<td>10.58</td>
<td>1.90</td>
<td>0.39</td>
<td>13.64</td>
<td>45.47</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>42.90</td>
<td>0.96</td>
<td>15.80</td>
<td>3.16</td>
<td>0.96</td>
<td>20.95</td>
<td>35.45</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>56.29</td>
<td>0.68</td>
<td>7.44</td>
<td>11.40</td>
<td>5.73</td>
<td>81.56</td>
<td>18.06</td>
</tr>
</tbody>
</table>

Fig. 1. Map of metropolitan Melbourne train stations displayed by cluster type.
percentage of train station entries made by walking, cycling, bus and by private motor vehicle respectively. Similar to the location of cluster 1 and cluster 3 train stations, the proportion of train station entries made by walking shown in panel (a) is generally highest for train stations in the inner and middle suburbs, and declines by distance from the centre of Melbourne. The walkability measures for Cluster 1 train stations indicate that these train stations were located in highly walkable neighbourhoods. Indeed, neighbourhoods surrounding cluster 1 train stations had an average of 23 dwellings per hectare, which is consistent with the 25 dwellings per hectare recommended to support local walking (Boulangé et al., 2017). As a result, these stations were located in neighbourhoods that encouraged patrons to walk which was reflected in the patronage by mode results that showed most people accessing cluster 1 train stations by active transport, and in particular walking (64%).

Panel (b) shows the proportion of train station entries made by cycling, although there is no clear pattern, the map shows that some train stations located in the middle and outer areas of the city have a...
high proportion of train station entries by cycling. Panel (c) shows low proportions of train station entries made by bus, although there appear to be some stations in the middle and outer suburbs in the south east of the city that have slightly higher proportions of train station entries made by bus. Panel (d) shows train station entries made by private motor vehicle, which are generally low in the inner-city areas, and high for middle and outer suburban train stations. This reflects the walkability findings for cluster 2 train stations, which were generally located in low walkable neighbourhoods with poor pedestrian access. This was evidenced by the poor connected node and ped-shed ratios, the lowest gross dwelling density of all clusters (10 dwellings per hectare), low destination counts and diversity, and the lowest quantity of bus stops. Most entries in cluster 2 train stations were made by private motor vehicle (45%), even though there was less car parking available at these stations than around cluster 3 train stations.

Cluster 3 train stations were located in moderately walkable neighbourhoods. These train stations had some neighbourhood walkability attributes similar to those found in cluster 1, in particular the connected node and walkable ped-shed ratios. However, cluster 3 train station neighbourhoods differed significantly in the quantity and variety of local destinations available. This finding is not surprising, because the average gross dwelling density of neighbourhoods surrounding cluster 3 stations was considerably lower at 16.2 dwellings per hectare. Of the three clusters, cluster 3 train stations had the most car parking and bike facilities. The patronage results indicated that to get to these train stations 43% of people walked, 35% used a private motor vehicle, 16% used buses, and only 1% cycled.

5. Discussion

This research derived a typology for evaluating access to metropolitan Melbourne’s train stations based upon a cluster analysis methodology using 14 walkability measures. In deriving this typology, we sought to understand what constitutes a walkable train station neighbourhood and to identify possible interventions and redevelopment that could create more walkable neighbourhoods surrounding established and new train stations with the aim of identifying effective TOD location. We identified and defined three types of train stations, which we validated using Google maps and with train station patronage data by transport mode. With few exceptions, the three clusters were generally located in inner, middle and outer areas of metropolitan Melbourne. The degree of walkability declined with increasing distance from the Central Business District.

Generally, cluster 1 train stations were located in the most walkable neighbourhoods and as a consequence, were accessed by approximately 30% more pedestrians. Access by cyclists was the lowest of all the clusters, and notably these stations had few bike facilities. For Cluster 1 train stations located in inner cities areas, the on-road cycling infrastructure may have already existed, and for people who cycle, the distances to key destinations are close enough to avoid transit use altogether. However, for other cluster 1 train stations located in suburban areas, greater investment in safe on-road cycling infrastructure may increase both the number of train patrons overall, and those accessing the train station by cycling (Pucher & Buehler, 2009).

Overall, train stations with the most potential for redevelopment to become TODs were the cluster 3 stations. These were located in the next most walkable neighbourhoods, although the levels of gross density and access to destinations were considerably lower than cluster 1 train stations. Cluster 3 train stations had twice the availability of bike facilities compared to the remaining cluster 1 and 2 train stations, however only 1% of patrons cycled. More cycling and use of existing bike facilities could be encouraged if they were complemented by on-road cycling infrastructure provided within 5 km. Such infrastructure has been shown to connect large tracts of land across metropolitan Melbourne (Giles-Corti et al., 2014). This approach could also be applied to cluster 2 train stations which also had both a low number of bike facilities and a low proportion of train station entries made by cycling. Indeed, better integration of active transport networks and provision of storage and changing facilities within the train station could improve access by active transport and in particular entries by cyclists.

For cluster 3 train stations investment in on-road cycling infrastructure could also potentially reduce the need for nearby car parking, the space for which could be re-purposed into mixed land use developments that increase gross dwelling density, the number of destinations and employment density of these precincts, further increasing their walkability and potential to become fully fledged TODs. However, for cluster 2 train stations a different strategy may be required since they were located in the least walkable neighbourhoods and did not appear to be designed to maximise accessibility. They were poorly served by public transport (i.e., bus and tram stops); and had less car parking available than cluster 3 train stations. For these train stations, providing a greater variety and frequency of inter-connected public transport services may improve the accessibility and patronage of cluster 2 train stations, even if they are not accessed by active modes of walking and cycling. Similarly, for those living further away, cluster 2 train stations may benefit from having larger car parking facilities or park and ride facilities available to maximise the number of people who could make use of train travel (Olaru et al., 2014).

Research of this type is ideally positioned to inform the identification of context-specific strategies for enhancing the accessibility of individual train stations. Moreover, the results could be used to guide train station development in established and new areas. However, some built environment features are more difficult to upgrade than others, and are best planned in the site design phase, to ensure that train station areas are located in walkable neighbourhoods from the outset. This would make it easier for subsequent development to convert train stations into TOD at a later stage (Messenger & Ewing, 1996).

In this study, we found that train stations in the middle and outer areas may benefit from the provision of on-road and end-of-trip cycling facilities that would facilitate access to train stations from distances further away. However, we identified fewer options for supporting access by walking for train stations built in low walkable areas, and without massive redevelopment and investment, it is unlikely that these train stations could ever become well-functioning TODs (Vale, 2015). To increase the patronage of these stations, different strategies are required including enhancing the cycling facilities, providing more inter-connected transit services as well as providing more car parking facilities. However, preferably, future train stations will not be built in low walkable neighbourhoods.

Our findings, can be used to mitigate low walkability by using development strategies informed by the train station typology and based on the features of the most walkable train station neighbourhoods. This research supports a revision to the old adage, ‘if we build it, they will come’ to perhaps — “if we build it with the right features, they will come’. Train stations designed to encourage walking and cycling, require a supportive and broader land use strategy. This is important as cities grow because it maximises public transport use, provides more sustainable transport options and supports the health and well-being of nearby residents (Watts et al., 2018).

6. Strengths and limitations

The primary strength and contribution of this research is it developed a rigorous, evidence-driven typology derived using 14 measures of walkability of the train station neighbourhood that included measures for car parking and bike facilities. Many researchers advocate for the use of car parking and bike facility measures in analyses of train stations and TODs but lack the data to do so (Griffiths & Curtis, 2017; Renne, 2009). In this study the use of these two measures in addition to the remaining walkability measures revealed useful and important findings that could be used by policymakers and practitioners for
improving station access and train patronage by both of these transport modes. This typology can be used to identify train stations ripe for development, and for understanding which features could be changed and by how much. It can also be used by policymakers, planners and urban designers to derive a normative or standardized development, and for understanding which features could be changed and by which modes.

Another strength of this paper is the use of 800 m street network buffers for calculating train station features. The 800 m buffer is evidence-based: i.e., in Melbourne the policy is that residents live within 800 m of train stations. However, the network buffers also better represent people’s walking experience as they are based on connected streets rather than Euclidean or as the crow flies approaches used in other studies. Much past research draws upon Euclidean buffers (Kamruzzaman et al., 2014; Lyu et al., 2016), although research incorporating ped-shed ratios necessarily uses data calculated via both methods (Higgins & Kanaroglou, 2016; Vale, 2015). Although many papers use a distance of either 700 m or 800 m, we recommend that 800 m be used as a standard for future work as there is now supporting literature on 800 m being a walkable distance linked to food related and local living destinations (Gunn et al., 2017a; Gunn et al., 2017b; King et al., 2015). However, a limitation of our approach was that our street network layer used road centre-line data, which may not accurately represent the full extent of the pedestrian network potentially underestimating some of the walkability measures. There was also a difference in the timing of the GIS data, which is based on 2011 spatial layers and the train station patronage data which was collated in 2014. Whilst a temporal mismatch such as this is common practice given the availability and complexity of creating GIS data, it is not clear how this affects the analysis or the results.

Future research could include additional walkability measures such as perceptions of safety, the presence of street lighting and aesthetic aspects of the built environment such as street trees. We opted to use cluster analysis ahead of other methods for combining variables because it is an intuitive, simple and well-established method used in previous built environment studies. Other methods such as latent class analysis may have been preferable, and these could be explored in future studies.

Our validation of the clusters by transport mode is a final strength of this paper. However, we recommend that future work explicitly model walking and active transport behaviours in-conjunction with built environment features using other methods including the use of Poisson or Binomial regression models. Such approaches are now being used to measure the health benefits of good built environment design and are being used in economic evaluation frameworks (Boulangé et al., 2017; Veerman et al., 2016), which are now being used for advocating for healthy, liveable and sustainable city-wide development.

Acknowledgments

This manuscript is based on the research of Dana Jeffrey as part of her Masters of Urban Planning undertaken at The University of Melbourne.

Funding declaration

CB and LDG are support by the NHMRC Centre for Research Excellence in Healthy Liveable Communities Grant (#1061404). BGC is supported by an NHMRC Senior Principal Research Fellowship (#1107672) and Clean Air and Urban Landscapes Hub – National Environmental Sciences Programme. CB and BGC are also supported by the NHMRC Australian Prevention Partnership Centre (# 9100001).

Competing interests

The authors declare they have no actual or potential competing interests.

Authors’ contribution

DJ, LG, CB, BGC, and SW conceived the research. DJ, LG and CB conducted the analyses. LD, DJ, BGC, and CB drafted the manuscript. All authors contributed to data interpretation, added intellectual content, commented on drafts and approved the final version. All authors read and approved the final manuscript.

References


D. Jeffrey, et al.


Abingdon, Oxon New York, NY.


Public Transport Victoria, 2012. Metlink Database.


Res. 12 (3), 79–104.


