Abstract
Informed by critical realist philosophy of science, this paper aims to contribute to a better understanding of the issue of causality within urban and planning research. The concept of causality dominating within certain influential disciplinary and philosophical traditions is difficult to reconcile with research into influences of the built environment on human actions. This paper promotes a conceptualising of causality in terms of generative mechanisms operating in different combinations in normally non-closed systems, and discusses in what sense the built environment can be said to exert causal influences on human actions. In order to integrate knowledge about causal influences at the level of the individual and at the city level, a combination of qualitative and quantitative research methods is recommended.

Keywords: Built environment, causality, tendencies, critical realism

1. Introduction
Planning built environments in ways which contribute to more sustainable urban development has long been high on urban planners’ agendas. Such planning presupposes knowledge of how different built environment characteristics have an influence, often via actions of the inhabitants, on relevant impact categories. Better knowledge of how and why urban built environments influence human actions is important to this end. Interdisciplinary integration is necessary to improve such knowledge, since many sustainability-relevant actions, for example, activities resulting in greenhouse gas emissions and activities of importance to public health or urban social justice, involve spatial, infrastructural, socioeconomic, cultural, as well as demographic factors of influence.

The notion of influences implies that there is causality involved. However, different disciplines have different, and sometimes fundamentally contradictory understandings of the relationship between the spatial/physical and the social. This hampers interdisciplinary research and can create distrust and scepticism between planning practitioners from different disciplines.
The urban planning tradition clearly assumes a causal status of the built environment. This is reflected in academic literature within a number of substantive fields such as New Urbanism and smart growth debates in North America, various compact city discourses in Europe, literature on eco-villages in the UK, China and the Middle East, and post-apartheid planning to counteract social and ethnic segregation in South Africa. However, although an ‘ontological turn’ is arguably underway within planning research (Næss et al., 2013), this literature rarely elaborates on the notion of causality. The idea that the built environment exerts causal influences is still controversial within important disciplines involved in planning research and theoretical debates. The very notion of causality does not sit well with certain philosophical positions that have become influential within planning theory.

Informed by critical realist philosophy of science, this paper aims to contribute to a better understanding of the importance of the urban built environment and land use (hereafter ‘the built environment’) to human actions. In this context, the notion of built environment applies to:

- The geographical distribution, fabric and design of the building stock
- The mutual location of different functions (such as residences, workplaces, public institutions and service) within the building stock
- The transport infrastructure system (road/street network, public transport infrastructure, and parking conditions)

The main line of argument goes like this:

Knowledge of how and why urban built environments influence human actions and social life is important in order to form sustainable strategies in urban land use planning, and this presupposes an adequate understanding of the causal nature of such influences (this section). However, some influential branches of contemporary planning theory seem to reject, or at least downplay, the notion of causality itself and/or the causal status of the built environment. The roots of this disinterest in or denial of causality can arguably be traced back to disciplinary traditions and philosophical positions completely rejecting the notion of causality in studies of humans and the social world, or denying the specific causal status of the physical world on social life (Section 2). Such assumptions are at odds with the very idea of urban planning. Instead, this paper supports an understanding of causality in terms of causal powers and generative mechanisms operating in different combinations in normally non-closed systems (Section 3). Causal powers operate within the natural as well as in the social world, and in the social world they include the abilities of human agents as well as social structures and discourses to influence human actions, their well-being and social life in general. The built environment is perhaps the most obvious manifestation of socially constructed entities and should arguably be considered as a particular sub-set of social structures (Section 4). Although the built environment and human agents have their own causal powers and properties, and influence each other mutually, this happens with a time lag, which implies that our actions on a given day take place under pre-existing structural conditions. Urban structures influence the actions of individuals but also give rise to emergent city-level causal relationships. Some brief concluding remarks end the paper (Section 5).
Reflecting my own research background, many of the illustrative examples refer to influences of the built environment on travel. The causal-theoretical points are, however, in my opinion equally relevant to a number of other fields of research.

2. Sources of denial and disinterest in causality

In much contemporary planning literature, causality between influential factors and effects seems to be considered a more or less discredited concept insofar as social phenomena are dealt with. Instead, the emphasis is placed on interpretive understanding (Bransen, 2001). In particular, this appears to be the case among debaters making efforts to replace instrumental rationality with communicative rationality and among poststructuralist theorists rejecting the possibility of demarcating objects from each other. Alongside this, and possibly partially as a result, there has been a decades-long emphasis in planning theory literature on planning processes rather than on the contents and consequences of the spatial solutions promoted in urban planning (Healey, 1996; Fainstein, 2010). The very term ‘planning theory’ has almost become synonymous with what Faludi (1973) called theory ‘of’ planning (i.e. on the functions/tasks, procedural methods and results of planning), as distinct from theories ‘in’ or ‘for’ planning (dealing with - respectively - connections and relationships within the part of reality that is the object of planning, and methods included in the ‘craft of planning’). If discussed at all in planning theory, causality tends to be conceived in terms of relational fields where the directions of influence cannot be determined (Giddens, 1984).

The comments and critiques raised by scholars from the tradition of interpretive social research should not be ignored. Although compelling geometrical arguments can be put forth about the ways in which built environments may facilitate or hamper human actions, the ones who act are people, not buildings or geographical distributions of urban facilities. Urban life is a social phenomenon, and insights from the social sciences about people’s various motivations and considerations influencing actions such as choice of residence, use of urban facilities or travel behaviour, should obviously be welcomed. On the other hand, urban planners’ often qualitative arguments about likely impacts have been attacked by traditional transport planners demanding more precise and quantified effect estimates (e.g., Van Wee and Boarnet, 2014). The latter criticism is based on the belief that causal relationships can be identified and relatively context-independent estimates of their strengths can be measured through statistical analyses. I have elsewhere (Næss, 2015a) criticized the prevalent correlationist approach within transportation research. This critique will not be repeated here. Instead, the target audience of the present paper is made up of scholars and practitioners within the wide field of urban and regional planning.

As mentioned in the introduction, some influential branches of contemporary planning theory seem to reject, or at least downplay, the notion of causality itself and/or the causal status of the built environment1. Such disinterest in or denial of causality appears to have evolved from philosophical positions and disciplinary traditions that completely reject the

---

1 The authors referred to below as examples of such rejection or disinterest may not have meant their explicit or implicit denial of built environment causality seriously. But some of their statements nevertheless say so, taken literally, and these argument can be (and have sometimes been) used to discredit research and planning focusing on the influences of built environment on human actions.
notion of causality in studies of humans and the social world, or deny that the physical world can have any causal influence on social life. The main types of such denial/rejection are:

- the erklären/verstehen dichotomy within hermeneutics, leading to rejection of the relevance of causality in studies of humans and social life
- the sociological taboo against attributing any influence on human actions/social life from the physical environment
- a tradition within human geography deeming the attribution of causal influences to space as environmental determinism or spatial fetishism
- a strong individualist focus in the microeconomic approach within neoclassical economics
- the conflation of structure and agency found in Giddens’ structuration theory
- poststructuralist radical contextuality where objects cannot be distinguished from each other

The following sub-sections will briefly discuss each of these positions.

**Non-applicability**

Throughout the twentieth century there has been a tradition within parts of social science (especially those associated with hermeneuticism) of rejecting the very notion of causality in the study of human actions and social life. A dichotomy between verstehen (interpretive understanding)\(^2\) and erklären (causal explanation) has been advocated, where the latter approach has been said to belong only to the natural sciences (Bransen, 2001). According to this view, there are no causes of human action (but there are motives, intentions, meanings, etc.). The power the material world has is typically a power exerted by our interpretations of the material world. ‘Action determinants’ (including spatial frame conditions) only become relevant mediated by subjective patterns of interpretation (Scheiner, 2005). The focus of the researchers within this tradition tends, however, to be limited to the patterns of interpretation themselves, whereas the ways in which different physical surroundings tend to cause different interpretations are left outside of the scope. The approach thus privileges constructivist and subjectivist analyses of relationships between built environment and human actions.

In a similar vein, architectural urban theories are typically either persuasive ‘action theories’ aiming to guide future design of buildings and urban environments, or ‘interpretation theories’ seeking to interpret what has already been built (Nygaard, 2002). In either case, there is little emphasis on causal explanations.

The dichotomy between explanation and understanding posited by classical hermeneuticists like Weber and Dilthey appears to be rooted in an understanding of causality where humans and the social world are elevated, so to speak, above the mundane world in which causation applies, be it as causal agents or as subjects of causal influences. The position thus represents a kind of anthropocentric exceptionalism. Its underlying understanding of

\(^2\) I.e. understanding the meaning of action from the actor’s point of view (Wikipedia, 2015).
causation is David Hume’s conceptualisation of causality as constant conjunctions between subsequent events, with colliding billiard balls as the paradigmatic example.

This theory of causality is however, as shown in Section 3, not uncontested – and I would argue that it is not a very fruitful conceptualisation either. For land use planning and research, the position of non-applicability of causation to phenomena in the social world is a deleterious position. If no causation exists in the social world, different land use patterns cannot conceivably affect social phenomena and human actions. Research into the nature and magnitude of such impacts would then be meaningless, as would any attempt to influence human actions, welfare and social life through land use planning.

The sociological taboo
Within sociology, a dominating view was for a long time that the physical surroundings have no significant influence on social phenomena. This position can be traced back to the tradition of the sociological classic scholar Emile Durkheim. According to the Durkheimian tradition, social phenomena (termed ‘social facts’ by Durkheim) can only be explained by linking them to other social phenomena (as opposed to referring to psychological, biological or physical conditions) (Tonboe, 1993, p. 4; Lidskog, 1998, p. 22). The position is partly also rooted in the legacy from Max Weber. Although the scholars of the Weberian tradition emphasize that human actions are influenced by the way the actors perceive their surroundings, these perceptions are assumed to be formed mainly through influence from the surrounding actors, and not from the physical characteristics of the situation (Lidskog, ibid.). According to Dunlap and Catton (1983, p. 118), “The Durkheimian legacy suggested that the physical environment should be ignored, while the Weberian legacy suggested that it could be ignored, for it was deemed unimportant in social life.” Tonboe (1993, pp. 168-169) considers this legacy as rooted in the anthropocentric dogma that humans and social life are in control over nature, the material and the development, i.e. the perspective of modernity – an interpretation also shared by Catton and Dunlap (1978).

A strong focus on attitudes and a corresponding downplaying of structural enablements and constraints can be seen in certain strands within contemporary mobility sociology. Some researchers have claimed that travel in late-modern society is increasingly becoming a goal in itself rather than a means for transporting oneself to destinations where desired or necessary activities are carried out (Urry, 2000; Steg, 2005). Although not denying that travel also occurs for instrumental reasons, this approach has contributed to drawing research attention toward lifestyle factors and away from the influences of socio-spatial structures.

Accusations of environmental determinism and spatial fetishism
Debaters within sociology and human geography have often used the label of environmental determinism about researchers and designers who attach importance to influences from the physical environment on social phenomena. Admittedly, a strong belief in the possibility of shaping or controlling human behaviour by manipulating the physical environment has sometimes been expressed, especially within architecture and engineering. Among

3 Ironically, the hermeneuticist tradition seems to have adopted the classical positivist notion of causality, which, according to Comte, rejects concerns with the ‘intimate nature’ of causal relationships as well as the ‘essential mode’ by which they have been produced. Instead, research should, according to classical positivism, focus on ‘constant relations of succession or similitude’ (Comte, 1835). This positivist understanding may have paved the way for the denial of the concept of causality in the hermeneutic tradition.
architects, such ‘physical determinism’ has been quite pervasive in certain periods, for example in the design and urban planning program of the functionalist movement (Le Corbusier, 1973; see also Chermayeff 1982). In contemporary planning literature, such positions are rarely found, if at all.

Referring to the Durkheimian and Weberian traditions within sociology, Dunlap and Catton (1983) claim that:

“Should one violate these traditions and suggest that the physical environment might be relevant for understanding human behaviour or social organization, one risked being labelled an ‘environmental determinist’.”

However, few of the scholars accused by Weber, Durkheim and their followers of being environmental determinists actually believed that the physical surroundings exerted any deterministic influence on human actions or social life. For example, the 19th century geographers Vidal (Weber’s antagonist) and Ratzel, as well as most of the later so-called environmental determinists within sociology and geography, rather held the physical environment to be influential along with social and cultural factors (Tonboe, 1993: 145-166). The same applies to the tradition in urban planning inspired by Patrick Geddes and Lewis Mumford, with its emphasis on understanding the specific geographic context of planning and the need to work with the ‘relationships among place, work and folk’. Even the earlier mentioned proponents of the functionalist movement hardly considered the physical environment capable of shaping human behaviour in a fully deterministic way, although they may have held exaggerated beliefs about its importance compared to cultural and social factors. A related position exists in the long-standing tradition of attributing the decisive influence of new technologies on societal development, often referred to as technological determinism although its proponents have hardly believed technology to be the only cause of social change.

A denial of the importance of the physical/spatial surroundings to human behaviour can also be found within important traditions of human geography, among others, in the (early) writings of David Harvey (1982) and Manuel Castells (1977). In Castells’s view, space has no meaning independent of its social background. Everything that space might possibly mean, lies in the social, economic and political powers by which it has been produced. Attributing any importance to space beyond this was, in Castells’s view, illegitimate space fetishism. In line with this, he refused to separate the spatial from the social, even for analytical purposes (Castells, 1977, quoted from Tonboe, 1993, pp. 4 and 531). This kind of denial could be characterized as social reductionism, where all causal powers and properties of the built environment are reduced to the societal conditions from which they emerged.

The label of environmental determinism appears to be mostly a pejorative characteristic used by antagonists of researchers and designers attributing importance to the physical environment for human life. Although the causal influence of the built environment may have been overstated in some of this literature (e.g. the belief among ‘Shared Space’ proponents in the possibility of increasing urban social interaction and people’s general considerateness through changed street design principles), and although the causal influences of the built environment may have been taken for granted without further reflection, few, if any contemporary writers in the urban design and planning literature assume such influences to be of a deterministic nature. At least insofar as determinism is
understood in the strong sense that one factor will entirely determine behaviour, the arguments of social scientists against environmental determinism in urban research must be considered largely as attacks against a straw man. If determinism is instead understood in a weaker sense - highlighting the importance of the physical environment without claiming this to be the only factor of influence, we are no longer referring to an ontological, a priori condition but an empirical question that needs to be investigated in specific contexts.

‘There is no such thing as society’
The infamous statement by the former British Prime Minister Margaret Thatcher that ‘there is no such thing as society’ illustrates a downplaying or denial of the causal importance of social structures typical within neoclassical economics. According to Archer (2000:6), several theorists within this school of thought consider social structures as mere epiphenomena of the aggregate preferences of instrumentally rational agents. Logically, the structures thus created include not only ‘immaterial’ social structures but also material structures resulting from the preferences and actions of individuals (see section 4). This kind of denial could aptly be characterized as individualist reductionism.

For a long time, transport economists tended to omit urban form variables in their analyses. For example, in a study of long-term automobile fuel demand, Schipper and Johansson (1994, p. 5) reluctantly chose to include a variable to represent population density (for the country as a whole), but, as they write, "there has been much discussion as to whether to include a non-economic variable, i.e., since such variables cannot be derived from economic theory, they should not be imposed in the regression model." Today, researchers studying relationships between the built environment and travel from a microeconomic point of view usually include urban form variables in their analyses, but their explanations tend to privilege economic variables like income, car ownership and consumer preference. The way this is done is often by introducing ‘control variables’ that could themselves, from a non-reductionist theoretical perspective, be expected to be influenced by built environment characteristics. Examples of this are the frequent use of car ownership as a control variable in studies of relationships between residential location and travel (see, e.g., Hjorthol, 1998; Krizek, 2003; Rajamani et al., 2003). Another example is the argument about pre-existing travel preferences as the (main) causes of geographical differences in travel behaviour rather than the urban structural situations of the dwellings (Kitamura et al., 1997; Van Wee, 2009). I have discussed these issues elsewhere (Næss, 2009 and 2014a).

Conflation of structure and agency
Another position, represented by, among others, Giddens’ (1984) structuration theory, claims that structure and agency are mutually constitutive and cannot be untied. According to Archer (2000), this seriously hampers any analysis of how structures and agents influence each other, as the specific properties and powers of neither the structures nor the agents can be identified. The context of Giddens’s theorizing about structure-agency relationships is the social structures of everyday social interaction and the concepts people use to make sense of it. Giddens’s focus on social life and subjects’ reflexivity within it (Delanty & Strydom, 2003) may be useful for such a context but has important shortcomings when taken as a general theory of structure and agency. Notably, Giddens’s claim that structures are not material but only exist as ‘memory traces orienting the conduct of knowledgeable

---

4 A secondary phenomenon that occurs alongside or in parallel to a primary phenomenon.
human agents and as the instantiation of rules in the situated activities of agents’ (Giddens, 1987, p.21) is hardly useful for analyses of the built environment.

Giddens’s structuration theory largely disregards the important time lag between the influence of our surroundings on our actions and the influence of our actions on our surroundings. In a similar vein, Scheiner⁵ writes: “…. residential location choice and daily travel are types of human action. It follows that they are guided by rationales (or motives, reasons, purposes...), and they produce meaning. This understanding implies that residential location cannot be a cause for travel behaviour, as residential location itself is based on human action.” However, this statement disregards that a residential location, once chosen, influences the lengths, and often also the travel modes, of a large number of subsequent trips originating from the dwelling. Scheiner’s argument, like Giddens’s structuration theory itself, downplays the fact that we are born into and living in already-existing socio-material contexts⁶ (see Section 3). We can potentially change such structures, but we are nevertheless in any concrete action situation subject to the enablements and constraints given by the specific structural context (for example the urban built environment).

**Poststructuralist dissolution of categories**

Within postmodern and post-structural planning theory too (e.g. the large bulk of scholarly work since the 1990s building on thinkers such as Derrida, Foucault and Deleuze), theorizing on the causal influence of the built environment on human activity is hard to find. According to Williams (2005), a key feature of poststructuralist thought is the deconstruction of ontological concepts and the demarcation of objects. This implies a radical contextualism where ‘the outside is constitutive of the inside, always leaving its trace within the boundary of the category’ (Derrida, 1988, quoted from Dixon & Jones, 1998). Any categorization into causes and effects is considered as contingently constructed through differentiated systems of power. Consequently, the distinction is blurred between objects that cause things to happen and objects that are subject to the influence of causally efficacious objects. The awareness of context and interrelatedness between numerous objects in open systems is in line with critical realism (see section 3), but the poststructuralist take on contextuality is so radical that any social categorization becomes an arbitrary result of the socio-historical processes of power play. Causation as such may not be denied, but this ontology hampers, if not makes impossible, analysis of causal processes: “You will never know .... whether it was X that caused Z or X’s haunting trace Y?” (Dixon & Jones, 1998:255). Although this ontology may not preclude causal analyses within a given system of thought, there will be no basis for judging which one(s), among competing analyses carried out within different systems of thought, should be considered as more or less credible. Anyhow, whether for ontological reasons or because other questions have been at the core of interest, there has generally

---

⁵ Personal e-mail communication May 8, 2012.

⁶ When developing his structuration theory, Giddens admittedly drew explicitly on Hagerstrand’s time-geography, which strongly emphasizes the constraints on human actions set by spatial contexts. However, Giddens criticized Hagerstrand’s theory for recapitulating the dualism of action and structure (Van Schaik, 2009:53). Giddens was influenced by Janelle’s theory of space-time convergence, and this may have led him to downplay the time lag between the influences of structures on agency and agents’ subsequent reproduction, modification or change of the structures in question.
been little emphasis on causal explanations in poststructuralist literature, including such literature dealing with planning and the built environment.7

The built environment is sometimes theorized in terms of *semiosis*, where certain physical forms express symbolic messages, often reflecting power relations and a hierarchical social order (Gotttdiener and Lagopoulos, 1986). The ways in which the built environment may influence such mundane and physical actions as travel between different destinations within a city are, however, hardly ever theorized within this literature.

### 3. Causes as Tendencies

Instead of conceptualisations of causality as being non-applicable in studies of human actions or as correlations between subsequent events, understanding causality in terms of tendencies and generative mechanisms (Bhaskar, 2008) appears much more apposite in planning studies as well as generally. Such a concept of causality is often associated with realist philosophical positions. Theories based on such understanding are sometimes termed dispositional theories of causality (Mumford & Anjum, 2012). The conceptualisation of causality presented below draws chiefly on the philosophy of science position known as critical realism.

Critical realism, as presented by, among others, Bhaskar (1993, 1998 and 2008), Sayer (1992), Archer (2000), and Danermark et al. (2001), is a position within philosophy of science which, more than many other such positions, appears relevant for research into the ways in which structural conditions (including built environment characteristics) influence human actions. According to critical realism, the world exists independently of our knowledge of it, and this knowledge is both fallible and theory-laden. Critical realism distinguishes between three different domains of reality: the empirical (consisting of what we experience directly or indirectly), the actual (where events occur whether or not we experience them) and the real (including both experiences, events and the causal powers producing the events) (Danermark et al., ibid.).

Moreover, according to critical realist ontology, reality consists of different levels (strata), where new properties and causal powers emerge at each level, compared to the level below. The causal powers of the lower levels still exist at the higher levels. Different levels often correspond to the research fields of different disciplines. Moving from the most basic (lowest) to more complex (higher) levels, we can, for example, refer to phenomena dealt with within physics, chemistry, biology, psychology, and sociology as different levels of reality.

Patterns of car travel among inhabitants of urban neighbourhoods depend, for example, on phenomena within physics and chemistry (e.g. the qualities of materials and fuels used for vehicle propulsion), biology (e.g. drivers’ sight and physical ability to use steering wheel, accelerator and brakes), psychology (e.g. calm driving vs. road rage), social structures and institutions (e.g. affordability of cars, employment opportunities, traffic rules), built

---

7 Foucault’s discussions of how the layout of built structures such as prisons and asylums can discipline the behavior of inmates obviously imply that built environments are assumed to exert influence on humans and social life. His broad concept of *dispositif* (Foucault, 1977) includes architectural forms and arguably also the built environment generally, but the causal status of the built environment as such is, to my knowledge, not discussed in Foucault’s writings.
environment (e.g. infrastructure provision, proximity vs. distance), as well as cultures and discourses (e.g. a ‘car culture’ vs. high popularity of public transport and non-motorized modes). How we choose to define the levels depends on the research topic, and the specific classification is a social construction. However, the very phenomenon of emergence – that new causal powers emerge at higher levels, depending on lower levels but also working back on them – is a feature of the real world (Bhaskar, 2008). The extent to which the higher levels influence the lower varies with the levels dealt with as well as between different epochs. Human societies are, for example, emergent from psychological, biological, nature-geographical and basic physical conditions, but today they work back on the natural basis with considerably stronger power than in the Stone Age.

In some philosophical traditions, the notion of cause is used only about active impulses that trigger a given event to happen, whereas other circumstances enabling, modifying or preventing the effect of the triggering factor are termed as ‘conditions’ rather than causes (Mackie, 1965). Critical realism does not make any such distinction. Causes are instead understood as ‘tendencies’ that may or may not be actualized. According to critical realist ontology, what happens in the world – in nature as well as in society – is a result of causal powers working via several mechanisms. Objects have properties enabling them to exercise certain forms of influences on other objects and/or make them liable to certain kinds of influences from other objects. Reality consists mostly of more or less open systems where empirical regularities rarely occur spontaneously. In open systems, many different causal powers operate at once. Some augment each other, others counter each other, and some are only actuated under the influence of other causal powers. All this varies with the specific context (Bhaskar, 2008).

The general condition of context-dependent multi-causality still does not rule out the possibility of identifying causal mechanisms within the social world. In a critical realist view, any human organization – from the family household to public administration, shopping centres and factories – are examples of partially closed systems. This limited and conditioned closure within some parts of society exists because humans are internally related to each other. Some of these relations and regulations are more significant than others are and can engender regularities displayed within many fields of social life (Danermark et al., 2001).

Acknowledgment of the importance of interpretive understanding does not, in a critical realist view, prevent inclusion of causal explanations in explanations of ‘purposeful action’. To want something is, for a human being, the same as being prone to act in a specific manner (Johansson, 1984, p. 88). For example, the want for employment is a basis for the tendency among people to seek jobs and travel to their workplaces. This want is, in its turn, strongly influenced by the wage labour system of our society and the need to earn money for subsistence. Which mean is the best to realise a wish will depend on the conditions under which the wish is to be realised. Moreover, according to critical realism, reasons may themselves be plausibly construed as causes (Bhaskar, 1998; Fairclough et al., 2002). If a teacher asks a student to raise her hand and she subsequently does so, the physical movement of the hand is not caused by an involuntary tick but by the student’s intention to raise it. This is in turn caused by the preceding communication between teacher and student, and the reasons the student might have for adhering to the teacher’s request (Bhaskar, 1998). Causal mechanisms can involve attitudes and knowledge resources of individuals, as well as intersubjective production of meaning (e.g. in discourses).
The conceptualising of causes as tendencies sits well with the multiple-cause situation a researcher is facing when trying to explain how the inhabitants of a city use its various urban spaces and their movements to get there. The built environment exerts its influence, but so do also a number of other conditions. The patterns of spatial behaviour in a given city are a result of people’s resources, needs, wishes, and obligations, modified by the constraints and opportunities given by the structural conditions of society. Any cause of spatial behaviour is only a contributory cause. Other causes may always add to or counteract the cause in question.

The above conceptualisation of relations between structural conditions and human behaviour implies an understanding of structure and agency attributing causal power to structures as well as agents. Critical realist ontology insists that structure and agency both have their own properties and powers and can be separated analytically to investigate their distinct influences and interaction (Bhaskar, 1993; Archer 2000; Danermark et al. 2001). There is a time lag between the creation of a social structure and its subsequent influences on human actions: although social structures are created by actions of agents, they form part of the conditions for future actions, sometimes for a long time beyond the lifetime of those who once created the structures. Except our natural environment, the structures surrounding us are in various ways 'socially constructed'. The ‘constructs’ may be physical artefacts like buildings or roads, or more immaterial structures like property ownership, economic state of affairs or prevalent belief systems and cultural traditions. Once created, the various types of structures hold emergent powers and properties different from and beyond the aggregate sum of agential powers by which they were produced. At the same time, human actions reproduce, amend and transform the structures. Such changes most often come about gradually and slowly, but sometimes more radically and fast.

The ways in which human actions depend on built environment characteristics should not be conceived of as either voluntarism or determinism, but instead in terms of tendencies or dispositions, where the ‘structural imperatives’ levied by the built environment are adapted, augmented or counteracted by a number of other structural and individual conditions. Furthermore, the relationships between the built environment and the actions of individuals are not unidirectional but a matter of two-way influences. As discussed more thoroughly in Section 4, the built environment of a city influences a number of daily-life activities among its residents, including travel. But the residents obviously also influence the built environment. They do so as market agents by demanding, purchasing and using certain kinds of dwellings, vehicles and urban facilities, and they may together influence the way the urban built environment changes by electing politicians who support a particular form of urban development. Some inhabitants may also take on a more influential actor role in urban development, such as urban planner, property developer, etc.

4. The Built Environment as a Causally Effectual Structure

Spatial planning assumes that spatial conditions have a potential to influence human actions. If the built environment did not influence people’s well-being, the economy, the natural environment, etc., there would be no point in trying to influence how built structures

---

8 There might seem to be a danger of reification here. However, the continuing existence of social structures depends upon activity. Although the structure has been there from the beginning, it is only there in the end if we continue to use or maintain it (Bhaskar, 1993).
develop. However, actors obviously do try, and for good reasons (Næss, 2015b). Buildings and physical infrastructure normally do not actively trigger things to occur, but they can (usually in interaction with other causal powers) enable, amplify, facilitate, restrain, suppress or prevent the occurrence of events and situations. They do not in this regard diverge from persons, non-material social structures or non-human nature. Applying the critical realist notion of causes as tendencies, built environment characteristics must be considered to have causal effects on human actions and social phenomena, although not in a deterministic way.

There are traditions within sociology, geography, anthropology, architecture and engineering that have implicitly presumed that built environments are capable of influencing people’s actions, living conditions, well-being or aesthetic experience. However, the built environment has rarely been theorized as a causally effective social structure.

In social theory, the notion of social structures is often understood to denote overall economic systems, social hierarchies, legislation and social institutions as well as prevalent norms and discourses, each operating at diverse scales. Material structures are typically overlooked, though not explicitly excluded. The predominant non-inclusion of the built environment in theories around social structures might be a remnant of the traditional taboo in sociology against attributing causal influence on social life from the physical environment (Benton, 2001).

Dissimilar to the natural environment, the human-made urban fabric is unquestionably socially constructed. It might be sensible to consider the urban built environment as a specific sub-set of social structures, with its own particular emergent powers and liabilities. As touched upon in the introduction, the term ‘built environment’ as applied in this article does not only comprise the building stock and infrastructures as merely physical objects, but also the location of various functions and activities relative to each other within these physical structures. This, together with its key role in conditioning the activities, functions and patterned relations between various groups of a society, implies that built environments are much more than just ‘material structures’. The built environment is an expression of and makes up conditions for the sorts of societies we have.

Understanding the built environment as a sub-set of social structure of course does not mean that its dependency on geographical features such as topography, streams, availability of land suitable to be built on, construction materials, water, energy, etc., is disregarded. In accordance with critical realist ontology, I consider social structures (non-material as well as material) as dependent on (and emerging from) more basic levels of reality, including (and not the least) the natural environment. In line with Bhaskar’s concept of the ‘four-planar social being’ I consider material transactions with nature as indispensable for societies and human life, along with social interaction between agents, the social structure proper and the stratification of embodied personalities of agents. The stratification of embodied personality

---

9 Some other authors, such as Østerberg (1974), explicitly distinguish between social structures and material structures.
connotes psychological aspects that may come into play, whereas the notion of material transactions with nature connotes ecological aspects\(^\text{10}\) (Bhaskar, 1993).

As mentioned above, the causal properties of objects comprise both causal powers and liabilities. The built environment consists of material objects that are often solid and can block passage (e.g. walls), or they may facilitate passage only through certain paths (e.g. roads, doors). When material structures are not passable, they do not yield to any pressure from humans that might try to pass through them. They are not liable to such pressure. If a person pushes against a solid wall in order to get through it, this is, in physical terms, countered by a counter-pressure from the wall resulting in no further movement of either the wall or the pushing hand. This non-liability and counter-force is a property of the object itself, but it will only manifest itself if someone actually tries to pass through the object. Similarly, the ‘passability’ of another object (e.g. a road) is a liability that does not manifest itself until someone (a pedestrian, a vehicle or whatever) actually tries to use this passage. Moreover, passable parts of the built environment may only be passable when making considerable efforts to move along them. For example, it takes more effort to walk up stairs than along a corridor in a building. But even the movement along a corridor or a flat pavement requires some effort – there is always some friction, and this friction increases the longer distances that have to be overcome. This is of course also the case when moving by means of a vehicle. For example, the friction is greater when driving along a narrow and winding road with many potholes than when driving on a straight and wide paved road, and in both cases the friction is greater the further away the destination is from the point of departure. ‘Passability’ is not only affected by physical characteristics but also by the activities they facilitate. For example, heavy vehicle traffic along a road can render it difficult for crossing pedestrians to pass.

Similar to the way that individuals and societies belong to different strata, we can talk about individual buildings and the urban built environment as belonging to different strata. The urban built environment is a structure that is more than the sum of individual buildings. It makes up a pattern characterized by, among others, a specific overall density, an internal distribution between denser and less densely built areas, a configuration of different building types, a network of roads, metro lines, and with different amounts of interspersed non-built, publicly accessible areas. These patterns depend, of course, to a great extent on natural topography and climatic conditions, for example, it the area is hilly or flat, has access to the sea or not, is cool or tropical, etc. Built environment characteristics imply, for example, different average distances between the buildings, different distances to the central part of the city, differences in the accessibility of open space, different conditions for private and public transport, and differences in the extent to which buildings are casting shade on each other. Such differences have a bearing on, among other things, travel distances and modes, physical activity, energy requirement in buildings, and the number of pedestrians and vibrancy of social life in public spaces and streets, just to mention a few examples.

The ‘non-material’ social structures add to material constraints and enablers. Property rights define who is entitled to access different parts of the built environment. Although it

\(^{10}\) The critical realist concept of four-planar social being is thus fundamentally different from the ‘nature-blindness’ of traditional sociology (Dunlap & Catton, 1983; Benton, 2001) as well as the one-dimensional ‘homo oeconomicus’ of neoclassical economics (Archer, 2000).
may be possible to trespass on other people’s property illegally, the owners, or those persons given delegated power to do so, are normally the only ones who can lock and unlock doors and gates. Land use regulations often define the allowed types of activities or purposes in different parts of the built environment. Together with the dispositions of the owners, the physical suitability of the built structures, the situation within different segments of the property market, etc., this largely defines the actual use of buildings. An office building, then, will normally be accessible only for those who work or have an errand there. On the other hand, due to the social organization of labour, the workers normally have to go to the office building at least some days during the week (often all the weekdays), and in many cases at fixed hours of the day. The above material and non-material causal properties are important premises for theories such as Hägerstrand’s (1970) time-geography and various location theories.

Some of the structure-agency relationships typical for other kinds of social structures may be different when the structure in question is the urban morphology and land use with all their material and spatial features. Among other things, human-made material structures (such as roads and buildings) often have a high permanence, although being gradually adapted and altered by human actions. Buildings often have a lifetime of many decades or even centuries, and the street pattern in the inner parts of many cities still resembles the street network established several hundred years ago. The overall spatial/material structures of cities arguably make up more stable conditions for human daily-life activities than do certain other social structures.

Furthermore, while it is possible to ‘defy’ non-material structures to a greater or lesser extent (for example, breaking the law or violating the code of behaviour in an organization), material structures are in some ways more compelling and may make up constraints that are more absolute. You cannot drive through the walls of a building, and you run the risk of being killed if you walk between the metro rails. Many of the impacts of the built environment on human behaviour and social life are still less absolute. Apart from making up some absolute limitations and enabling actions that would otherwise not be at all possible, the built environment facilitates, encourages and discourages, making some actions and adaptations attractive and easy while making others burdensome and difficult (see below).

In Gibson’s (1977) terms, urban built environments could be understood as kind of affordances offering opportunities or limitations for action that a subject might perceive. According to Østerberg (1998), cities should be seen as a socio-material field of action, where the facticity of the material structures (cf. Sartre, 1948) creates restraints on the inhabitants’ possibilities to unfold their wishes. The conceptualising of the built environment as causally effectual on human actions and social life promoted in this paper has some similarities to both Gibson’s and Østerberg’s theories. However, rather than using their frameworks I have situated my understanding of the causal status of the built environment within the critical realist ontology of the stratified world (cf. Section 3), which I find much more elaborate than those of Gibson and Østerberg.

The causally effective characteristics of the built environment include both its observable and non-observable parts. The characteristics of a dwelling influencing its practical indoor functionality (the layout of rooms, windows, interior walls, etc.) are obviously observable, while, for example, its accessibility to job opportunities and service facilities, or the ownership type in different residential neighbourhoods, are less so.
Possibilism, probabilism and aggregate-level emergence

Within human geography, the positions of possibilism and probabilism were proposed in opposition to the (alleged) environmental determinism that had been a dominant view in the late 19th century. Possibilism admits that the environment sets certain restrictions on human activities, but holds that social conditions and the choices people make are more influential (Fellman, 2000; Collins English Dictionary 2014). However, possibilism is not concerned about the differing degrees to which a given set of environmental conditions may enable or constrain a specific kind of human activity. Some of the possible actions may be hard to carry out while others may be easy. A possible versus impossible dichotomy ignores such differences.

According to a probabilist view, humans will, based on the characteristics of the environment, be more likely to make certain decisions over other ones (Fellman, 2000). Besides making certain actions impossible (such as driving across a river where there is no bridge or underwater tunnel), the built environment makes up a set of inducements facilitating some sorts of human actions and discouraging other kinds of actions, still without rendering the latter impossible. Rather than a dichotomy between the facilitating and discouraging circumstances, there is frequently a continuum. The result is also not only of the kind ‘occurs or does not occur’ (e.g. using or not using car as a travel mode), but is often a degree along a continuum (e.g. the daily distance travelled). There may be a greater or lesser element of friction. This friction influences how likely it is that the possible actions will be carried out at all; how often, for how long or how far away they will take place; and by which means of conveyance the spatial separation of activities will be overcome.

Section 3 emphasized the critical realist concept of causality in terms of tendencies. ‘Tendency’ is here understood as how a given causal mechanism tends to operate when activated (although it may not be activated or it may be counteracted by other causal mechanisms). As stated by Fleetwood (2011, p. 98), tendencies “are not of an either/or, discrete, dichotomous or discontinuous nature but are also continuous; there are stronger and weaker tendencies and powers”. Moreover, “when it comes to considering the interplay of several tendencies or powers ... it may be very important to try and get some idea of whether the tendencies are stronger than the countervailing tendencies” (ibid.). If some tendencies, for example to prioritize job quality (in terms of job content, working conditions, salary, etc.) above proximity to the dwelling, are stronger and/or occurring more frequently within a population than any countervailing tendencies, this can give rise to well-grounded probability distributions. In the example, this could be a higher likelihood of long commuting distances among suburbanites living far away from a city’s largest concentration of workplaces (see below).

The emergence of causal relationships at a higher ontological stratum presupposes – at least in some cases – the existence of this kind of ‘event regularity’ at a lower stratum. For example, from the regularity of a high amount of precipitation caused by the multitude of causal mechanisms influencing cloud formations, wind patterns and temperatures on the Western Norwegian coast, a local climate has emerged in the city of Bergen differing sharply to the climatic conditions of a Saharan town. Yet, both the Western Norwegian coast and Sahara experience periods of drought as well as rainfall. The difference is that rainfall is regularly a much more frequent event in Bergen than in Sahara (and conversely for drought periods). These different climatic conditions are, in their turn, causally efficacious in...
producing different vegetation patterns in the two areas, which in their turn react back on cloud formation and precipitation (Næss & Strand, 2012).

Similarly, the causal relationships between city-level built environment characteristics and aggregate-scale human behaviour (e.g. in terms of travel) are dependent on some event regularity (again, only in the form of a probability distribution, not a deterministic conjunction) in the inhabitants’ prioritisation of different rationales for their actions. The regularities underlying aggregate causal notions such as ‘car-dependent cities’ or ‘socially segregating built environments’ are of a probabilistic nature (in critical realist terminology often referred to as demi-regularities). Such probabilism does not imply that the likelihood or unlikelihood can necessarily be measured accurately, let alone predicted. The nature of the relationships between built environment features and various sorts of human actions and characteristics of urban life should still be understood as probabilistic rather than merely possibilistic.\(^{11}\)

**Examples of impacts**

Built environments can influence human actions, well-being and social life in numerous ways: the size, layout, technical appliances and design of dwellings is important to the residents’ material standard of living; the distances between urban functions and facilities can influence the need for travel, as well as the possibilities people have for participating in activities; the uniformity or variation of a local neighbourhood in terms of dwelling sizes and standards can affect the degree of social segregation or integration; the visual look of buildings can evoke positive aesthetic experience; and building styles and architectural forms can express symbolic meaning, convey a message about power and social order and have a disciplining effect. Built environment characteristics can also influence social interaction, play, aggressiveness, relaxation, stress, etc.

Human actions, well-being and social life can be influenced by the material qualities of specific places (local morphology) as well as the location of a neighbourhood or a facility relative to other neighbourhoods and facilities (‘relative space’). Furthermore, the influences of built environments vary across geographical scales: The distribution of cities and towns of different sizes at a national (or international) scale obviously has other kinds of social effects than the distribution of different types of neighbourhoods within a city or the built environment features of an individual neighbourhood.

Some of the influences of the built environment are easily observable and hardly contested. This is typically the case when one or a few causal mechanisms are so dominating that no other mechanisms are able to disturb this prevalent relationship. The effect of a locked door in preventing unauthorized people from trespassing may serve as an example. Other built environment effects may be less easy to perceive, partly because the relevant aspects of the built environment are themselves unobservable (see above), or because the effects are difficult to discern from the influences of other factors of influence operating at the same time. The impact of one sort of change in the built environment may also be different,

\(^{11}\) An analogy with medical research may help illustrate this point. Although smoking disposes for lung cancer, not all smokers actually develop this disease. Smokers are still more likely than non-smokers to get lung cancer, and more so the more they smoke.
depending on the specific urban context (including socio-cultural conditions as well as the existing configuration of the built environment).

The impacts of single-family house construction on social segregation could be an example of the latter. In a small town where most dwellings are single-family homes and there are only a few apartment buildings and row houses, choosing single-family houses as the only type of new dwellings to be constructed may increase a tendency of social stigmatization of those inhabitants living in apartments and row houses. In a big city dominated by apartments and other high- or medium-density housing types, giving priority to the construction of single-family homes over the next few years would hardly result in any stigmatization of those living in other housing types. However, it might pave the way for the emergence of new, spatially segregated enclaves of privileged population groups. If continuing over a long period, suburban single-family house development could engender a flight of affluent residents into the suburbs, leaving the inner city behind as an area populated by less privileged groups and with an eroding tax base, as experienced in many cities in the USA in the last half of the 20th century (Power, 2001).

The links between built environment characteristics and public health make up another example of non-straightforward effects of the built environment on human life. Inner-city residents can usually access a higher number of potential trip destinations within acceptable walking or biking distances than do their suburban counterparts (see below). On the other hand, inner-city dwellers are often more exposed to air pollution and noise and often have less access to green recreation areas. So although living close to the city centre facilitates the use of ‘active’ travel modes, this may be outweighed by the counteracting influences of noise, pollution and lack of outdoor recreation opportunities (Næss, 2014b).

**Residential location and travel as an illustration**

The influence of residential location on travel behaviour may serve as a closer illustration of the ways in which built environment characteristics, together with a multitude of other causal powers, influence human actions and urban life. By determining the distances between locations where different activities may take place, and by providing conditions suitable for different modes of travelling, the urban built environment constitutes a set of conditions facilitating some kinds of travel behaviour (e.g. in terms of trip distances and travel modes) and rendering other types of travel behaviour less attractive or likely. For example, a resident living in the outer part of a metropolitan area may choose to travel 33 kilometres by car in the morning because this action, according to the person’s opinions, is the best means to realise a wish to reach the workplace at the scheduled hour. Another person, living in the downtown area, may instead choose to ride 2 km by bike in order to realise a similar wish. Needless to say, causes of travel behaviour of course also include personal characteristics of the travellers, such as gender, age, professional status, income as well as their values, norms, lifestyles, responsibilities and acquaintances. Lifestyles and norms, including travel preferences, are themselves influenced by discourses within society on mobility, freedom of choice, self-realization, etc., as well as by the general political-economic conditions of society.

People’s *transport rationales*, i.e. the backgrounds, motivations and justifications that people draw on when they make transport-relevant decisions about their participation in activities, location of these activities, travel modes and the routes followed, make up important links in the chains of causal mechanisms by which built environment features
influence travel (Næss and Jensen, 2005; Næss, 2013). For example, the tendency among a high proportion of residents in late-modern affluent cities to prioritize choice of the best facility above proximity to the dwelling implies that trip distances to jobs and other daily-life facilities will on average be considerably longer than what would be the case if people always used the closest opportunities. Especially, this is the case for journeys to work, since the location of the workplace is dependent on the double requirement that the employee must choose to apply for the job as well as be chosen by the employer among the applicants. Travel distances therefore depend more on the location of the dwelling relative to large concentrations of facilities than on the distance to the closest facilities. Since most cities have a higher concentration of workplaces in the inner areas, people who live close to the city centre usually have a larger number of potential job opportunities within a short distance from the dwelling than suburbanites have. This gives rise to a structural imperative which says that residents of the outer suburbs must be prepared to accept longer commuting distances than their inner-city counterparts in order to find a suitable job (Næss, 2013).

Aggregate-level causal relationships between built environment and travel, for example between suburban housing construction and growing car traffic, depend on some ‘demi-regularity’ (cf. above) in the residents’ prioritizing between different rationales for travel behaviour. Given the predominance of a given set of transport rationales, city-scale causal relationships between urban form and aggregate-level travel behavioural patterns are thus emergent from individual-level causal relationships between residential location and travel. These aggregate-level relationships depend on the continuing importance attached to rationales such as choice of the best facility, minimizing of friction of distance and limitation of physical efforts, giving rise to certain patterns of decisions resulting in certain patterns of travel behaviour under certain conditions (such as longer trips when distances to main concentrations of facilities are increased).

The transport rationales, according to which increased suburban residential development tends to increase car traffic, seem to make up an enduring condition of this socio-technical system. The various transport rationales manifest themselves in a pattern ensuring a stable aggregate correlation (i.e. in total for the many different tendencies represented by various transport rationales) between suburban residential location and long daily traveling distances, usually by car. The strength of these correlations vary across space and time, due to the different constellations of other causal mechanisms at work in each situation. In its turn, this emergent system (Bhaskar 2008, pp. 231–234) can react back on the rationales for the travel behaviour on which it is (partially) based. For example, the tendency of higher car dependency among suburbanites can contribute to the development of cultural norms in the suburbs according to which the car is more and more considered the ‘normal’ means of transport. Such prevalent attitudes cannot be understood detached from their spatial contexts, i.e. the physical and location-based facilitation for different travel modes (Næss, 2005).

5. Concluding Remarks
For urban planning to be meaningful and possible, urban spatial structures must exert some influence on people’s behaviour, well-being, the natural environment, the economy, etc. – otherwise there would be no point in trying to influence how these structures develop.
However, contemporary planning theory debates often shun the notion of causality. Rather than investigating and theorizing causal relationships between the built environment and human actions, the attention tends to be concentrated on agents’ patterns of interpretation of the physical/spatial surroundings. On the other hand, certain other traditions, such as transportation research or public health studies, are dominated by quantitative approaches where causality is, often without much reflection, conceived of in terms of constant conjunctions between subsequent events.

In order to enhance interdisciplinary cooperation and integration, it is necessary to get rid of dogma and a priori assumptions that in practice deny the knowledge represented by other relevant disciplines. Instead of continuing to be split in the two separate domains of quantitative/statistical and qualitative/interpretive research, inquiry into influences of the built environment on human actions should integrate knowledge across these domains. Such integration requires a more adequate conceptualising of causality than the ones now dominating within each of the two research traditions.

This paper advocates the view that causality in the social world should be understood in terms of generative mechanisms; that both social structures and human agency have their own properties and causal powers; and that built environment characteristics could be considered as a particular sub-set of social structures characterized by having a relatively high permanence (although usually being constantly subject to incremental change). Reflecting my own research background, many of the illustrative examples of this paper were about influences of built environment characteristics on travel behaviour. The main points are, however, equally relevant to a number of other fields of research (e.g. social interaction, relaxation/stress, etc.), some of which are domains where the denial of causality is more dominant than in the travel behaviour field.

Research into influences of built environment on human actions, human well-being and social life should aim at explaining the specific causal mechanisms through which built environment characteristics of a city influence observed patterns of human life. Qualitative studies are necessary in order to get such insight (cf. Røe, 2000; Næss, 2013). There are, however, also emergent causal relationships between city-level built environment characteristics and overall patterns of human actions at a city scale, for example in terms of travel behaviour. We can therefore talk about population characteristics, dominant transport rationales, configuration of dwellings, workplaces etc. at the city level engendering location-dependent variations in phenomena such as commuting distances, car dependency, levels of physical activity, or health risks. For urban planning, the latter group of influences of urban conditions are important. In order to investigate such patterns, the ‘demi-regularities’ forming the base for the overall urban-scale relationships must be uncovered. This would usually involve some sorts of surveys and statistical analyses. However, such aggregate-scale analyses of relationships between built environmental characteristics and, say, travel can tell little or nothing about causality unless the underlying mechanisms have been uncovered through qualitative research at the scale of individuals and households.

A combination of qualitative and quantitative research methods is therefore called for in order to integrate knowledge about causal influences at the level of the individual and at the city level. Such combination involves breaking out of the ‘cages’ of, on the one hand, the quantitative regularity-seeking tradition and its typical disinterest in digging deeply into why the observed regularities exist, and, on the other hand, the idiographic causality-denying
interpretivist tradition, which focuses only on individuals and their motivations, without considering what aggregate patterns and new emergent structures might result from the different individuals’ adaptations to their contexts.

Acknowledgments: The author would like to thank Sebastian Peters, Tim Richardson and the three anonymous referees of Planning Theory and Practice for valuable comments on previous versions of this article.

References


