

Group selfies and Snapchat

Page, Ruth

DOI:

[10.1016/j.dcm.2018.10.003](https://doi.org/10.1016/j.dcm.2018.10.003)

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version

Peer reviewed version

Citation for published version (Harvard):

Page, R 2019, 'Group selfies and Snapchat: from sociality to synthetic collectivisation', *Discourse, Context and Media*, vol. 28, pp. 79-92. <https://doi.org/10.1016/j.dcm.2018.10.003>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

Checked for eligibility: 21/03/2019

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Group selfies and Snapchat: From sociality to synthetic collectivisation

Abstract

This paper sets out a multimodal framework that can be used to explore the ways in which people are positioned as individuals and groups within the selfies and other kinds of representations found in video-sharing practices. Unlike earlier, monomodal accounts of selfies, the framework accounts for the visual, aural and verbal resources that are used in video-sharing. The analysis focuses on the videos and photos that are produced and consumed in the Featured Stories of Snapchat as collective accounts of public events of different kinds. The results show that, in Snapchat, constructing group identities is prominent, both in selfies and quasi-selfies. This reflects a discourse of ‘us-ness’ current in many forms of social media, and which prizes particular forms of sociality. The uses of this discourse are ideologically charged, and include the strategic use of synthetic collectivisation as an emerging form of audio-visual communication which contrives to position the individual member of the audience as if they were part of a larger group, sharing the same experience and perspective as the person creating the video. The framework is used in this paper to examine the ways in which the collective identities of fans, mourners and protestors are constructed in Featured Stories from Snapchat, but is of relevance to many other forms of multimodal communication that are shared through social media sites and services.

Keywords:

Snapchat, Sociality, Selfies, Synthetic collectivisation.

Introduction

Sharing images and video online via social media services has become pervasive practice, particularly associated with social network sites and messaging apps. Chang (2016) reported that over 9000 photographs are shared on Snapchat every second and daily video views are estimated at 10 billion. Selfies are one genre often shared through these videos and photographs. Selfies are characterised as a digital type of self-portrait, taken using a smart phone camera and shared through social media sites. They incorporate a complex range of phenomenon which varies according to whether a selfie represents an individual or a group of people, the style of the selfie and the multimodal resources that are used (for example,

produced as video or as still photography). The group selfie is a digital self-portrait of an individual within a larger group of people and has attracted attention within the mainstream media, for example, in relation to Ellen De Generes's notorious photograph of herself and other celebrities taken at the 86th Academy Awards ceremony in 2014. Group selfies are by no means the preserve of celebrity practice and as they are shared by ordinary people through many kinds of social media services, they open up the possibility for selfie-takers to represent their identity as part of a group and to share that content with groups of different sizes.

The flexibility with which a selfie-taker can position his or herself within a group can be thought of in relation to what Miller et al. (2016) have recently described as 'scalable sociality', that is, the ways in which social media sites and services enable a person to interact with others in small or large groups and across public and private communicative contexts. Better understanding the group selfie within this context of scalable sociality is important for a number of reasons. First, analysing both personal and group selfies is an antidote to the widespread moral panic that promotes a simplistic view of selfies as individualistic resources which promote narcissism (Arpaci et al., 2018) and vanity (Abidin, 2016) - an individualistic emphasis that has been perpetrated by earlier research on selfies within discourse studies. For example, Zappavigna's (2016) and Zappavigna and Zhao's (2017) work on 'mommy blogging' and Eagar and Dann's analysis of self-branding (2016) concentrate on images of individuals. On the rare occasions that group selfies have been included in discourse analytic research, they have only been treated as a subordinate element within a larger project, as in Georgakopoulou's (2016) reframing of selfies as a type of 'small story'. Second, even in media studies, where selfies have been discussed in relation to collective action such as such as activist campaigns (Deller and Tilton, 2015) or political protests (Kuntsman, 2017), these studies have not examined the construction of the group selfie per se (for example, in relation to the image composition), but concentrated on the macro-social outcomes of the selfie-sharing. As Veum and Undrum (2018) rightly point out, there is a gap between media studies' emphasis on these macro-social issues (such as the feminist use of selfies) and the micro-analytic foci of discourse-analytic scrutiny of (individual) selfies. Taking a Critical Discourse approach bridges this divide, and means that in shifting analytical attention away from the individual and towards the sociality of the group selfie we can begin to interrogate the discourses of 'us-ness' that are at stake in social media, both in terms of the resources that are used to construct them and to question critically the ends to which this sociality might be put.

Furthermore, in both media studies and discourse analysis, researchers have only studied selfies in relation to still photography. Selfies are not just produced as images, but are a form of multimodal discourse which can include visual, aural and verbal elements when shared through video clips that can be created on smart phones. This presents a major challenge for discourse analysts, whose earlier work on selfies has been monomodal, concentrating only on photography (as in the work of Zhao and Zappavigna, 2018) or only on the verbal content which accompanies the images (Georgakopoulou, 2016). There has, as yet, been no attempt to analyse systematically the ways in which visual, aural and verbal elements might together contribute to the positioning of people vis-à-vis selfies, and in particular, no attempt to differentiate how this might be constructed for groups as compared with individuals. Finally, in discourse studies, debates about the methods used for analysing multimodal data still persist. Despite longstanding recognition that communication entails far more than words alone (Kress and van Leeuwen, 2001), researchers are still seeking robust and replicable methods for making sense of multimodal data. The photos and video clips in Snapchat are particularly complex with respect to the methods they demand for data collection and analysis. They provide us with an opportunity to demonstrate the levels of methodological rigour and transparency that will help the move forward future development of multimodal discourse analysis.

Our paper is situated as a response to the research gaps and challenges in this earlier work. We set out a new discourse-analytic account of how people are positioned relative to one another through social video practices, combining micro-level categorization with the macro-social interpretation of these categories. We develop a framework that can account for the multimodal positioning of groups and individuals through the perceptual resources of image, sound and words in social video-sharing. In this way, we redress the previously monomodal and individualistic emphasis within selfie research and critically evaluate the ways in which sociality is constructed. We then use the framework to illuminate data from a important but relatively under-scrutinised social messaging app where the representation of groups *and* individuals proliferate, Snapchat, one of the most popular social media sites used by young people today (Smith and Anderson, 2018). However, the framework is also applicable to audio-visual material broadcast through mainstream and other forms of social media and so will be useful to scholars in a number of fields including discourse studies, media and communication, and the visual arts.

Snapchat

Snapchat is a particularly rich context in which to examine the representation of individuals and groups, including those found in selfies of different kinds. It was designed from the outset as a way to share visual messages ('Snaps') privately, with a very minimal user interface focussing on the camera. In the last few years, the app has added a number of more public features, which has increased the interpersonal dimensions of the service and promotes the construction of a Snapchat community. This includes 'stories', where individuals can make Snaps visible to all their friends for 24 hours and the SnapMap, where users can share geolocated Snaps to 'Our Story', of which a selection become visible to all users, and can be viewed by browsing a map of the world. Some of these Stories are curated by Snapchat's staff, and become 'Featured Stories' that are visible both from the SnapMap and from the list of other stories.

The possibilities for representing people within a Snap are to some extent shaped by the affordances of Snapchat as a media platform. Snapchat is a 'camera first' technology, so the creation of the visual content precedes the ways in which users can shape that content. Figure 1 shows the steps involved in creating a simple Snap. First, users take a picture or record a video. Creating Snaps is different from most photography. Snaps are always vertical. This tall, narrow image format gives preference to the human body, and is much less effective for landscapes. Snap video clips have a time limit of 10 seconds. When the clip is recorded, users can write or draw on it, or add visual effects. A finished Snap is sent to the friends selected from a list. 'My Story' and 'Our Story' are also listed as possible recipients. Users may also see an option to share to a topical or location-based Story. Movement between these screens to create, access and view Stories is achieved through 'swiping' from one screen to another (see Figure 2).

Creating a snap:



1. The app opens with the camera. Press for picture, press and hold for video.



2. When the picture is taken, more controls appear.



3. Add text and visual effects, if desired.



4. Select recipients, including "My Story".

Figure 1: Creating a Snap

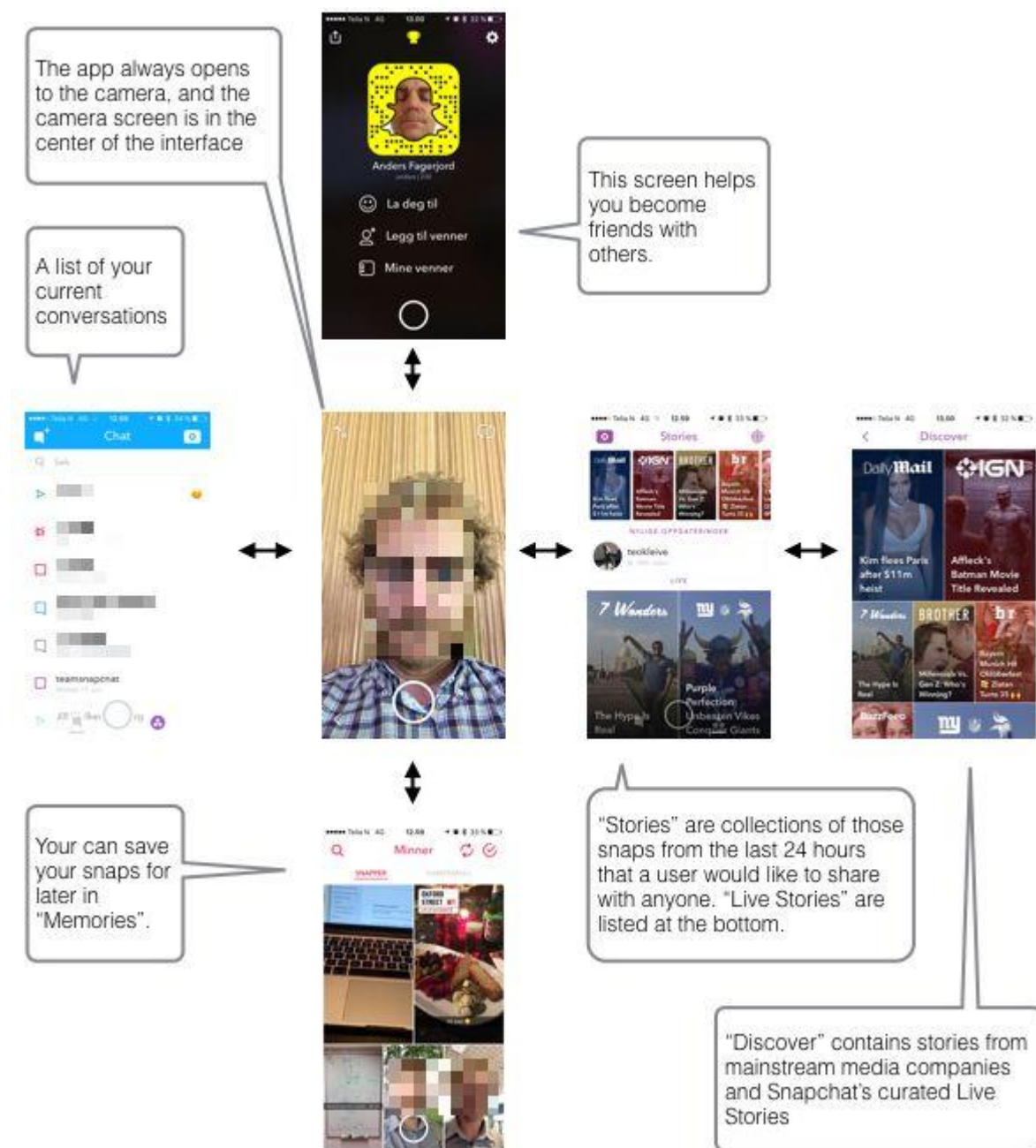


Figure 2: Snapchat's user interface. Arrows signify swipes between screens.

Featured Stories were previously known as Live Stories and built upon the previously tested 'Our Story', and similarly are only available for 24 hours. While Stories are individual records of individual users, Featured Stories are collective records. They are edited from clips contributed by many users who have participated in a certain event or been in a certain location, and so likely to contain representations of groups as well as individuals. Indeed, the announcement of the Live Stories in Snapchat's blog (News 29.8.2014) reflects a strong

connection between these particular types of Snap and the construction of a wider group identity, describing these as ‘*truly* a product of the Snapchat community’ (emphasis in the original). While there are other examples of videos assembled from user-generated content, Snapchat’s Featured Stories are the first such videos to have become an influential and constant media influence viewed by a mass audience. On average a Snapchat Featured Story is viewed by about 20 million users (Dodson, 2015), and many stories are released each week. These Featured stories are positioned within the interface of Snapchat within the ‘Discover’ screen, which was introduced in January 2015. The Featured Stories of collective, personal experience appear alongside a daily edition of stories produced specifically for Snapchat in collaboration with established media outlets such as CNN, Seventeen and Mashable (Snapchat news, 27.1.2015). Unlike other aspects of Snapchat which allow interaction between Snapchat members (e.g. via private chats), Featured Stories are somewhat limited in that viewers do not ‘chat’ with the person who has created the Snaps contained therein, although it is possible for the person whose Snap is contained in the Story to see how many times people have viewed it. As the platform has grown, it has also developed monetisation strategies, such as sponsored stories, ads positioned between stories (Snapchat News 17.10.2014) and sponsored filters and lenses that users can use on their own videos and photos. The potential for Featured Stories to construct a particular kind of sociality thus needs to be understood within the wider, critical history of social media, where user generated content (the Snaps) are interwoven with commercially-produced content (adverts and mainstream media content) as part of an economically motivated development of the service in question, typical also of other older examples of social media platforms (Gillespie, 2010: 348). The importance of the Snapchat community that underpins the creation of the Featured stories calls into question the rhetoric of ‘us-ness’ that appears, on the one hand, to celebrate the contribution of the Snapchat members but is, on the other hand, strategically driven by economic imperatives.

Snaps, selfies and sociality

We conceptualise the videos and photographs shared as Snaps (including selfies) as forms of discourse that enable identity to enter the social world through interaction with others (Bucholtz and Hall, 2005: 586-7). This brings together two aspects of the Snaps: the identities that are constructed within their content and the relationship between the persons who produce and consume them. We explore this from a social semiotic perspective derived

from Halliday's (1994) work, where he proposed that language was meaningful simultaneously at three levels. For each Snap, meaning is created at the following levels:

- Ideational: What the Snaps represent;
- Interpersonal: How the Snaps construe relationships between people, including the Snap creator, the audience and other persons represented in the Snaps;
- Textual: How the Snaps are organised in terms of their semiotic resources (image, sound and words).

A social semiotic approach understands meaning-making as highly context-dependent, and thereby allows the analyst to move between micro-analytic scrutiny of particular Snaps and the interpretation of those videos in the light of wider social discourses. We are particularly interested in how the analysis of the Snaps contributes to the discourses of 'us-ness' (where sociality is prized) and 'me-ness' (where individualistic self-expression is valued). The combined ideational and interpersonal focus of our analysis allows us to explore the construction of sociality and the ends to which this is put within a particular social media context: the Featured stories of Snapchat.

Sociality, as the extent to which an individual aligns their position relative to others within a group has been a key theme in the research literature in many disciplines. Within computer mediated discourse analysis, researchers have examined the linguistic variation that can occur because of the number and type of people who are involved in the interaction, as well as the types of technology being used. For example, in Herring's (2007) influential scheme for analysing computer mediated communication, she includes the participation structure (which includes group size) as a factor which may affect the ensuing interaction. Others have been interested in the ways that relationality, as our 'degree of alignment with others' (Lambert Graham, 2015: 306), is negotiated through online interactions which create different kinds of individual and group identities. In media studies, scholars examined the ways in which sociality is valued. For example, Steinfeld et al. (2008) showed how the size of a Friend list on Facebook could be regarded as a form of social capital, and Marwick and boyd (2010) theorised the aggregated Follower lists in Twitter as 'fan bases' to be managed in the processes of micro-celebrity. As Van Dijck (2013) reminds us, once sociality became technological, it also became 'salable', pointing to the commercial exploitation hidden behind the rhetoric of sociality.

More recently, Miller et al. (2016) have conceptualised ‘scalable sociality’ as the increasing amount of choice that a person has over the size of group with whom they might wish to communicate with via social media, from the smallest unit of interaction between two persons through to messages that are broadcast to very large audiences (for example, public posts to a micro-blogging site like Twitter). The evolution of Snapchat neatly illustrates the flexibility of scalable sociality, where the publicly available Featured stories scale up the sociality of video and photo sharing (distributed to the large, imagined audience of the Snapchat community) as compared to the private, dyadic chats (where Snaps are sent between individual members) that were first available in the service. Scalable sociality might seem most readily analysed in relation to the interpersonal aspects of social media, that is, the ways in which a person might vary the number of persons with whom they are sharing content and which might in turn influence the ways in which that group are positioned relative to the Snap. In addition, we suggest that the concept of scalable sociality can also be applied to the ideational aspects of Snaps, where people can choose flexibly how to represent their identity within and perspective towards groups of different sizes and of different types.

Our interest in how the discourses of sociality are constructed in the practices of video-sharing brings into question the complex relationship between group and individual identity. We begin from the premise that identity is not a ‘thing’ but is a multifaceted, relational and dynamic process (Simon, 2004). We follow the distinctions between individual and collective identity set out within interactional pragmatics (following Spencer Oatey, 2005), where individual self-representation focuses on the attributes of the personal self, such as appearance, capability and so on (Culpeper, 2011: 27). Collective identity refers to the way a person positions themselves as members of a group or category (Klandermans, 2014: 2). We recognise the potential for collective and individual identities to intersect and overlap. For example, as Lewis (2018: 215) points out, photographing a plate of food can at once be interpreted as consumer individualism, or as an act of social caring. However, in our analysis we follow the systemic functional principles, where semiotic resources are structured as choices and where one choice within a given system precludes another. Accordingly, in our analysis of the Snaps, we differentiate between the positioning of individuals and groups, where an individual is a single person who is shown or heard on their own in the video and where collective identities are constructed in content where more than one person is seen or heard. This distinction is important as it cuts across the many different kinds of collective identities that can be construed in Snaps, and allows us to make comparisons between the kinds of multimodal positioning that can be negotiated for different types of collective

identities, such as protest identities (Rovisco and Veneti, 2017) or fandoms (Van den Bulke et al., 2015). Indeed, the ways in which individuals and groups are represented in Snaps can vary considerably, where the representation can infer the well-recognised distinctions between ‘me’, ‘us’ and ‘them’ used to suggest affiliation, approbation, legitimacy and distance as documented extensively in Critical Discourse Studies.

The distinction between an individual and a group is also an interpersonal matter. Snaps can be constructed so that the producers and recipients are positioned rhetorically as individuals, or groups. A key concept that bridges the ideational and interpersonal dimensions of meaning is intersubjectivity. Intersubjectivity is a term which ‘appears to express a core component of the social’, and has been used in various disciplines, although with different definitions (Reich, 2010: 40). Within the study of selfies, intersubjectivity has been used as the touchstone by which Zhao and Zappavigna (2018) have begun to tackle how interpersonal meaning is construed in selfies. In their recent work, they draw on the psychologist Gillespie’s (2009) definition, which conceptualises intersubjectivity as ‘the variety of relations between perspectives. These perspectives can belong to individuals, groups, or traditions or discourses’ (Gillespie and Cornish, 2009: 19). The key contribution of Zhao and Zappavigna’s (2018) work (see also Zappavigna, 2016; Zappavigna and Zhao 2017) has been to recognise that the analysis of perspective in social photography practices must incorporate the selfie-taker as the creator who is also represented in visual composition of the content (rather than in previous forms of photography where the image-producer does not appear in the image). In their approach to intersubjectivity, Zhao and Zappavigna thereby differentiate between types of selfies based on the levels through which the selfie-taker mediates their perspective on the viewed content. For example, they (2018: 1745) differentiate between the perspective of a personal selfie (as a mediated representation of the self) and a more complex layering of perspectives in other kinds of selfies, such as mirror selfies, where the viewer’s perspective of the ‘self’ in the image is further mediated by the mirror and camera phone (creating a meta-meta-perspective: ‘look at me, looking at me’).

In our paper we develop their framework for analysing intersubjectivity in two new directions. First, we argue that the analysis of how ‘perspective’ is shared is underpinned by a fundamental distinction based on the proxemic positioning of the camera relative to the represented content and which shapes the social distance projected between the producers and consumers of video and photo shared through social media. We argue that the multimodal resources of these videos function to construct different kinds of distance between the position of the camera and the represented content which the audience can view. In certain

cases, the audience is invited to imagine that they are co-present with the social video or photograph creator (henceforth, Snap creator) and thereby to share their spatial and social orientation towards the content in the video. This allows us to go beyond the analysis of visual resources in intersubjectivity, for sound and language can also be used to construct the perspective and position of participants, not just in social video but also in other kinds of audio-visual interactions. Second, we propose that perspectives are shared in ways that intersect with scales of sociality, and therefore the number of persons positioned relative to each other must also be taken into account. For example, the content of a Snap may construct a shared perspective between a small group of two people (a single recipient and a single producer of a selfie, for example), or between larger groups of people. This distinction is crucial because in other semiotic systems, such as verbal language, deictic person categories that index the position of self and others include singular and plural forms that can be used variously to include or exclude persons within groups.

Following van Leeuwen (1999) we understand perspective as intimately bound up with the meanings related to distance. In his discussion, van Leeuwen draws on Hall's (1969) work on proxemics, in which Hall set out four types of interpersonal distance:

- Intimate distance (touch to 18 inches)
- Personal distance (18 inches to 4 feet)
- Social distance (4 feet to 12 feet)
- Public distance (12 feet to 25 feet)

These proxemic categories are particularly important for understanding how sociality is constructed, for as Scollon (2003: 54) points out, the obligation to engage socially with another person depends on their relative position within each of these spaces, and within personal distance, the expectation for social engagement is particularly strong. Whilst we recognise that the measurements of Hall's proxemic zones may not be as precise as his initial outline suggested, the development of selfie-taking practices gives further support to proxemics as a framework for making sense of how perspective is made meaningful in social photography and video-sharing practices. In neuroscience, the peripersonal space (Rizzolatti et al., 1997), that is, the space within an arm's length of a person is recognised as a significant boundary in human interaction and coincides approximately with the boundary between personal and other types of space. Selfies, which typically are taken at arms' length, appear to place the camera on the boundary between intimate and personal distance. Given that this

also coincides with the approximate boundaries suggested for Hall's personal space, this may explain why selfies are considered to be 'conversational' and carry 'sociality' (Andreallo, 2017), all the more because selfies such as those found in Snaps are typically consumed via a mobile device which is positioned within the personal distance of the recipient viewing the image or video.

The perception of the personal space as constructed through the multimodal resources of video-sharing is a crucial within our framework. When the semiotic resources index the position of the camera relative to the personal space of the person(s) in the video, this strengthens the construction of sociality, especially where the person(s) include the video-creator. In contrast, where the position of the person(s) in the video is indexed as within the public distance of the camera (and which cannot, by virtue of practicality, be interpreted as the video-creator), the invitation to imagine co-presence with the snap-creator within their personal space is comparatively weak. Furthermore, the sociality as constructed through interpersonal distance can be further scaled up or down, depending on whether the viewer is being invited to imagine they are sharing the personal space with an individual or with a group.

Categories of intersubjectivity

In our framework, the categories of intersubjectivity are distinguished from each other depending on three factors:

- The construction of interpersonal distance
- The extent to which the presence of the photo or video creator is shown
- The number of people shown or heard in the video

Each factor can be constructed through the visual and aural resources in a video, as follows.

Interpersonal distance

The interpersonal distance between the camera and the content of the video can be constructed visually through the composition of the image, for example by the size of the objects or people shown in the frame and their position in the foreground or background of the image. Interpersonal distance can also be constructed aurally (van Leeuwen, 1999: 14). In the Snaps, volume, timbre and acoustic precision can be affected by the distance between the source of the sound and the microphone of the recording device (Collins and Dockwray,

2015). Sounds can thus appear to be nearby (as in the intimate or personal space of the camera) when they are loud and precise or more distant (as in the social or public space relative to the camera) when they are quieter and less acoustically precise. For both the visual and the aural resources, we analyse the perspective relating to interpersonal distance using Halls' (1969) proxemic categories.

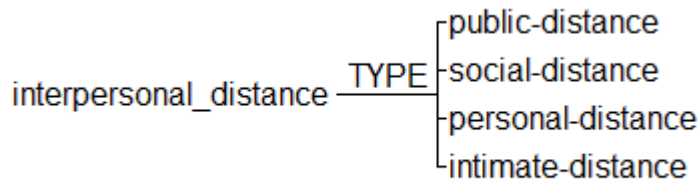


Figure 3. Choices within the system of interpersonal distance.

Number of persons seen or heard

The number of persons with whom the viewer is invited to share a perspective can be shown visually, in the subject matter which may or may not include people, and where people can be shown on their own as individuals, or in groups of various sizes. The aural resources can also indicate the number of people with whom the audience might share a perspective. In our framework we distinguish between whether the sound is produced by a person (such as speech, music, singing, clapping) or not produced by a person (e.g. background noise such as the sounds of weather, nature or traffic), and whether the sound is produced by an individual person or more than one person.

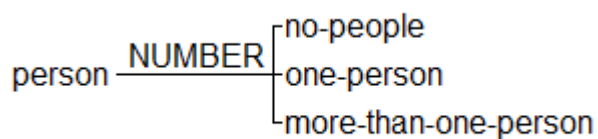


Figure 4. Choices within the system of person.

Presence of the Snap creator

Where a person is represented in the image, there may be additional compositional cues that the viewer is to interpret the person as the creator of the selfie.¹ This can include the person's position relative to the visual frame, for example relative to the angle of an arm or shoulder (assumed to be holding the camera), or the distorting effect of the camera's proximity to a

person's face. The effects of the reverse camera technology also reverse text which signals that the visual content is produced by as well as presenting the person in question. Aurally, the presence of the Snap creator can also be heard, for example when they speak or sing into the camera.

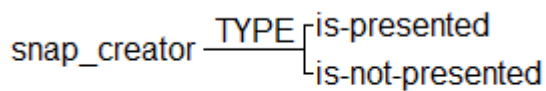


Figure 5. Choices of representation for the Snap creator.

The types of interpersonal distance, the number of people and the extent to which the Snap creator's presence is seen and heard combine to form three categories of intersubjectivity. Each category can be used to position individuals or groups relative to each other by the ways in which they invite the viewer, more or less strongly, to imagine that they are sharing the same personal space as the Snap creator.

‘Zero’ intersubjectivity

‘Zero’ intersubjectivity occurs when there is minimal invitation for the audience to share the perspective of the snap creator, as located within their personal space. Zhao and Zappavigna (2018: 1744) describe this as a direct perspective, which is constructed when the visual content is not restricted to the spatial perspective of a particular individual. Visually, this perspective tends to show the represented objects or scene so that it appears at a public distance from the camera. The Snap creator is not represented and the viewer ‘looks at’ the content in question, but has no visual cues as to the immediate physical space in which the creator is positioned as part of the interaction order. This is analogous to the concept of zero focalization (Genette, 1980), where the story is not told from the point of view of a character in the story, but from an external, omniscient narrator.

Aurally, a ‘zero’ invitation to share perspective does not appear to be restricted by the immediate spatial position of the source of the sound relative to the camera and hence also the Snap creator. Typically, the sounds heard in the video are at a volume to suggest they are at public distance from the camera, and thus the exact location of the Snap creator whose

perspective we might imagine is imprecise. In our data, examples of the sounds that create this effect include the use of weather, background traffic and other forms of ambient noise.

‘Zero’ intersubjectivity is weak in terms of the extent to which the audience is invited to ‘share the same perspective’ as the Snap creator. Whilst the very act of selecting the visual content with the frame by necessity entails sharing the same broad perspective, it is one where the Snap creator’s presence is suppressed (van Leeuwen, 2008). The choice to index the camera at public space from the content in the video also reduces the projected sociality, for the interpersonal meanings of this type of space do not usually presuppose the obligation for interaction. Lastly, the suppression of the Snap creator in the visual and aural resources means that for videos constructed with zero intersubjectivity, the Snap creator cannot be interpreted as an individual or as part of a group. Where individuals and groups are represented, they are at a distance and separate from the Snap creator. We interpret the invitation to the viewing audience for ‘zero’ intersubjectivity to be unmediated by the Snap creator and to ‘Look at X’ or ‘Listen to X’.

Presented intersubjectivity

In contrast, other kinds of perspectives that are shared (typically in personal and group selfies) use visual and aural resources to *present* the position of the video-creator in ways that depict them and their immediate location relative to the camera. The visual composition indexes the position of the person in the video relative to the personal or intimate space of the camera, and contains cues that the person shown in the image is to be interpreted as the same as the person creating the video. The Snap creator can be shown as a single person in the image, or appear with others in a group. The perspective that the viewing audience is invited to take is to ‘look at me, with me’ or ‘look at us, with us’.

The viewer can be invited to share a perspective which, through the aural resources of the video, also ‘presents’ the Snap creator. This occurs when the Snap creator’s voice is heard and assumed or shown to be produced within the intimate or personal space of the camera recording the video. This can be singular, when one person’s voice is heard, or plural when more than one voice is heard. The types of resources used to construct presented perspectives through sound in our data included speaking, singing and chanting. We interpret this as an

invitation to the audience to share the Snap creator's perspective and 'listen to me' or 'listen to us'.

Indirect intersubjectivity

In between the 'presented' and 'zero' options for intersubjectivity, there are other options whereby the Snap-creator can index their position relative to the camera indirectly, that is, without being fully seen or heard in the image or the sound. In Zhao and Zappavigna's (2018: 1745) visual categorisation, they label this kind of intersubjectivity as occurring with two types of selfie which they describe as 'inferred' and 'implied'. However, in terms of the interpersonal distance that is constructed in the so-called 'inferred' and 'implied' selfies, both index the position an object in the foreground of the image as within the intimate space of the camera. Therefore, based on our criteria, we regard the 'inferred' and 'implied' selfies as subtypes within a broader category of indirect intersubjectivity. With indirect intersubjectivity, the visual perspective is usually restricted in some way. For example, by framing the object in close focus and in line with the implied position of the Snap creator, the viewer is invited to imagine that they too are in that same personal space, taking the same restricted visual perspective towards the object in question, which we gloss as 'look at X with me'. Because the affordances of a smart phone camera are such that it is usually operated by one person alone, the Snap creator may be assumed to be an individual with whom the audience to share the perspective towards the content shown in the Snap.

In certain cases, the content shown within the intimate space of the camera is a group of people looking away from the camera towards a further focus beyond them in the picture (such as a sports pitch or a concert stage). This form of indirect intersubjectivity similarly indexes the proximity of the content within the personal or intimate space of the Snap creator, but does not present them visibly. However, because the creator is assumed to be in such close ecological proximity to the group in the foreground, she or he may be considered part of that same group, who as 'reacters' direct their eye line to a further 'phenomenon' (Kress and van Leeuwen, 1996). The perspective to be taken is no longer shared with that of an individual. Instead, the invitation to the viewer is to share the same space as a larger collective group, to 'look at X with *us*'.

One further option to construct indirect intersubjectivity by showing part of the Snap creator's body (their hand or feet) in the foreground of the image, so that the position of the

camera is interpreted within the intimate distance of that body part (typically hands or feet) and any other objects or items nearby. The type of indirect perspective in these ‘inferred selfies’ (Zhao and Zappavigna, 2018: 1745) can invite the viewer to share the same perspective of the Snap creator as an individual, or where the metonymic representation of the Snap creator suggests the presence of more than one person, to share the perspective of a group. Hence we interpret the perspectival invitation in these examples to ‘look at X with us’ or ‘look at X with me’.

The visual composition that invites the audience to share the perspective of the Snap creator in this indirect way often also includes objects or people in the background of the image, which may be in the social or public space relative to the camera. The aural resources heard as indirect forms of intersubjectivity are interpreted as being within this social or public space of the camera by means of their volume and acoustic precision. In our data this included sounds such as applause, music, chanting. The Snap creator’s voice is not heard directly in these sounds (by virtue of their distance from the camera) and therefore the perspective of the Snap creator cannot be easily inferred, except in cases where the sound appears to be made from the larger group of which the Snap creator is assumed to be a part. We gloss the interpretations to share the perspective arising from these aural resources as ‘Listen to X’ or ‘Listen to X with us’.

<insert Table 1 about here>

Table 1: Categories of intersubjectivity and their meanings, as constructed through visual and aural resources found in video-sharing.

A summary of the categories for intersubjectivity is given in Table 1. The multimodal construction of perspective can be used to compare and contrast group and personal selfies with other kinds of audio visual material, both in terms of the representational choices and also the extent to which the viewing audience is invited to imagine their co-presence alongside others within the personal space of the Snap-creator. Having set out this framework, we now consider two research questions. First, how are the choices for intersubjectivity distributed within Snapchat’s Featured Stories? Second, what ideological purposes do these categories of perspective-sharing serve? These questions are addressed through the analysis of the data collected for this paper.

Data

The data collected for this paper began with observation of the Featured Stories in Snapchat in several stages over a period of 15 months: 59 stories from 15 May through 11 July 2016, 10 stories from 12 through 21 September 2016; 3 stories from 20 January 21 through 22 January 2017, and 60 stories from 1 May through 20 July 2017. The Featured Stories we observed concerned events which were experienced by many people at the same time, often in the same locality (e.g. a sports stadium), although on some occasions the geographical distribution was more widespread (e.g. with Snaps sent from different towns or even countries). The type of events included sports events (rugby, football, baseball matches, tennis and racing tournaments all featured), music concerts and music festivals, conventions (such as fashion, gaming), national or religious festivals (Eid, Father's Day, Midsummer), political events and protests (UK referendum and election, the inauguration of the US President, protests at the G20 summit, and following the inauguration of President Donald Trump). Other stories were more light hearted, and were created to draw interaction from the Snapchat community around particular themes, such as the creation of a dancing hot dog filter which could be added to Snaps. Although our observation of the stories was ad hoc, and so no quantitative conclusions can be drawn about the topics that are covered, it is clear that the general themes in the Featured Stories stayed more or less stable over the course of a year, focusing on events where displays of being part of the audience might be anticipated (such as sports and festivals).

During the time that we observed this data, the terms and conditions of Snapchat precluded audio-visual recording the video content. Given that the data is highly ephemeral (available only for 24 hours), this posed a significant restriction on the way in which the data could be gathered and prepared for analysis. In order to comply with the site's terms and conditions, we transcribed a selection of 30 Featured Stories (897 Snaps in total), covering events with a range of collective experiences. For each of these stories, the terms and conditions of Snapchat mean that the audio and visual content of the stories had to be captured separately and in real time. Each Featured Story was watched repeatedly. Screenshots of each Snap were taken, including multiple screenshots of Snaps to show changing aspects of the representation in the visual stream (e.g. zoom in effects, changes in the participants who appeared in the frame). The audio content of each video clip was transcribed separately, using Jefferson's transcription system for the verbal spoken content and noting non-linguistic aural sounds such as clapping, music or road noise and the distance from which the sounds were interpreted relative to the camera as indicated by volume. Any

camera movements were also noted for each Snap. The material we collected was published in the publicly available space of Snapchat (that is, anyone with access to the app could view these Snaps whilst they were published in the Discover page). The collation of the Snaps in the Featured Stories meant that the Snap account of the individual creator was not available, meaning it was impossible to seek informed consent from the individual Snap creators of the content. In order to reduce the possibility of persons being identified from the Snaps, we blurred out faces in the examples reproduced in this paper.

From the 30 transcriptions, we down sampled ten examples that were from the most frequently occurring topics based on our wider observation of Featured Stories: concerts, national festivals, political events and sports. We selected Featured Stories so as to include different national contexts in Europe and North America, and then chose the Featured Stories with the largest number of Snaps in each category. The ten stories contained all the types of zero, presented and indirect intersubjectivity, but combined these in different ways, allowing us to compare and contrast the ends to which the visual and aural resources might be put. There were 435 Snaps in this subset of the data (summarised in Table 2). This subset was then analysed using a step-wise process of annotation, which we describe below.

Title of Story	Number of Snaps	Theme
Trump protests	55	Protest
Women's protests	71	Protest
Bye Bye football	26	Sport
Election 2017	30	Politics
Bank Holiday	25	National Festival
Ariana and Friends	60	Music
G20	28	Protest
Election results	37	Politics
Glastonbury	29	Festival
Wimbledon	22	Sport

Rolling Loud	52	Music
Total	435	

Table 2: Summary of the Stories and total Snaps in the down-sampled subset.

Methods

Whilst multimodal discourse analysis has opened up important lines of inquiry, it has also been subject to critique whereby the frameworks used in this area are understood as a post-hoc interpretation of the data (Bateman et al., 2004) and overly-reliant on the analyst’s contextual knowledge (Forceville, 1999). In this paper, we aim to increase the transparency of our analysis through the following steps. The first step is to document the choice of software used to annotate the data, as this can influence the ways in which analysts make sense of the data (Paulus et al., 2017). In our analysis, we used the UAM image annotation tool (O’Donnell, 2008) rather than other qualitative data analysis software, as this tool allowed us to develop our own annotation system but did so within a template derived from systemic functional principles of semiotic organisation. This choice is shaped by the origins of our framework in Zhao and Zappavigna’s (2018) work, which is likewise developed from systemic functional principles. We annotated the 435 Snaps at the level of the whole image, using an annotation manual prepared by the two authors of this paper, both of whom were familiar with the data in question. The annotation manual provided the descriptive criteria used to identify each type of visual perspective with sample images.

The interpretation of an annotation manual and its application to a particular dataset is still a subjective process. In order to test and refine the categories of analysis, we used inter coder reliability tests, using a team of three analysts. The analysts included author 1, and two further research assistants, neither of whom had been involved in the original data collection, but were familiar with Snapchat as a messaging app and were trained by author 1. The inter coder reliability tests were applied in two rounds. The agreement between the analysts for every decision in each Story was reviewed by author 1 and the level of agreement using Fleiss Kappa was calculated. In the first round, the mean level agreement across each Story was 0.8, which is considered the threshold between substantial and perfect levels of agreement for Kappa values (Landis and Koch, 1977). This rose to 0.9 in the second round of coding. Given the high Kappa values, no Snaps were excluded from the data. Following the visual analysis of the data, the aural aspects of the Snaps were analysed for the types of perspective that were constructed. The separate layers of annotation were then combined into

a composite bank of transcription, in order to examine the points of intersection between the different semiotic resources.

Quantitative results and discussion

The frequency of the types of visual and aural perspective as they appeared in the data was first calculated separately. The results are presented in Table 3 and Table 4.

Intersubjectivity	Count	%	Person	Count	%
Zero <i>Look at x</i>	268	61%	NA	-	
Indirect <i>Look at X with me/us</i>	53	13%	Single	14	3%
			Plural	39	9%
Presented <i>Look at me/us and with me/us</i>	114	26%	Single	69	16%
			Plural	45	10%
Total	435	100%		167	39%

Table 3: Quantitative comparison of the visual intersubjectivity found in Featured Story Snaps

Intersubjectivity	Count	%	Resource	Count	%	Person	Count	%
Zero <i>Listen to x</i>	84	20%	Background noise	61	15%	NA	-	-
			Silence	23	5%		-	-
Indirect <i>Listen to x (with us)</i>	139	32%	Instrumental music	52	12%	Single	0	0
						Plural	52	12%
			Clapping	22	5%	Single	0	0
						Plural	22	5%
			Chanting	39	9%	Single	0	0
						Plural	39	9%
			Singing	26	6%	Single	2	1%
						Plural	24	5%

Presented <i>Listen to me/us</i>	201	47%	Chanting	12	3%	Single	1	0.2%
						Plural	11	3%
			Speaking	171	40%	Single	133	31%
						Plural	38	9%
			Singing	18	4%	Single	4	1%
						Plural	14	3%
Total	424	100%		424	100%		340	79%

Table 4: Quantitative comparison of aural intersubjectivity found in Snaps from Featured Stories (still photographs removed).

The results prompt several observations. First, in terms of the visual resources, the dominant choice was zero intersubjectivity, accounting for 61 percent of the Snaps. Presented intersubjectivity was the next most frequent, suggesting that selfies continue to play an important part in the representational choices in Snapchat. In these selfies, the singular form (that is, selfies of an individual person alone) was more frequent than selfies of groups (16 percent as compared to 10 percent of all Snaps respectively). However, when the presented and indirect types of intersubjectivity were taken together, plural and single forms were equal in this data, accounting for 19 percent of the Snaps in each case. This suggests that the invitation to identify with a larger group is at least as important as imagining a shared position with another individual in Snapchat Featured Stories.

The invitation for the viewer to imagine herself as sharing the perspective with a group is reinforced through the way that aural perspectives were used. The analysis of the aural perspectives shows the importance of taking into account the sound alongside visual resources, for the perspective of the Snap creator was presented or indirectly constructed through sound in almost 80 percent of the Snaps. Within the aural intersubjectivity, there were many more invitations for the audience to imagine they were part of group (that is, where sounds were produced by more than one person), than sounds that were produced by an individual person, accounting for 48 percent as compared to 33 percent of the Snaps respectively. The majority of sounds produced by a group were within the indirect type of intersubjectivity (31 percent of all Snaps), such as clapping, chanting, singing and music performed by many people and heard within the social or public distance from the camera, whereas the presented types of aural intersubjectivity (speaking and singing heard within the

personal space of the camera) occurred less often (15 percent of all Snaps). No doubt, this is due in part to the kinds of events that were covered in the Featured Stories, which included concerts, sports events and protests, where the nearby sounds might be expected to be collectively produced. Nonetheless, the effect of the aural perspectives to invite a sense of co-presence with a larger group of others should not be underestimated, and understood as part of the rhetoric promoted by Snapchat in their construction of the Featured stories as co-produced with the Snapchat ‘community’.

We then compared how often the different types intersubjectivity as constructed through the visual and verbal resources were combined with each other, focusing on the perspectives where the viewer was invited to share an imagined co-presence within the Snap creator’s personal space (that is, the indirect and presented types of visual intersubjectivity). The results are summarised in Table 5.

Aural Intersubjectivity	Visual Intersubjectivity			
	Indirect		Presented	
	Single (n=5)	Plural (n=37)	Single (n=69)	Plural (n=45)
‘Zero’	40%	13%	4%	5%
Indirect (single)	20%	0%	0	1%
Indirect (plural)	40%	82%	18%	32%
Presented (single)	0	5%	72%	35%
Presented (plural)	0	0	6%	28%
Total	100%	100%	100%	100%

Table 5: Frequency with which the types of visual and aural perspectives were combined in Featured Stories.

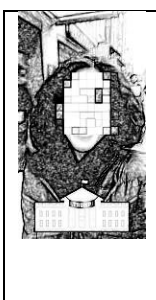
The quantitative comparison of the visual and aural perspectives suggests that these categories combine in ways that reinforce the individualising and collective aspects of the perspective which the viewer is invited to share. The presented visual perspective in singular forms (that is, in personal selfies) was most frequently combined with the Snap creator’s voice alone (72% of the Snaps in that category, n=50), resulting in an invitation to ‘look at

me, with me, and listen to me'. The group selfie occurred with a range of aural resources. However, the majority of the sounds invited the viewer to share the perspective with a group rather than an individual (60 percent as compared with 35 percent of the Snaps in that visual category). Hence the invitation to 'look at us, with us' is combined with 'listen to us' and 'listen to X' more often than to 'look at us and listen to me'. Lastly, the indirect visual perspective where the Snap creator was indexed as part of a group, most often is combined with the analogous auditory resources, and so invites the viewer to 'look with us at X' and 'listen with us to X'.


We do not suggest that this is necessarily a naturalistic reflection of the kinds of perspectives that are found in all types of video-sharing, but these results certainly suggest that Snapchat curates the Snaps for the Featured stories in ways that scale up the visual and aural invitation to sociality, in line with the large-scale sociality that the imagined audience of Snapchat constitutes. This emphasis on sociality does not mean that individualistic self-expression does not exist in the Featured stories; the discourses of 'us-ness' and 'me-ness' co-exist. Instead, it leads us to question further what ends to which these 'group perspectives' are put and how the discourses of sociality might be interwoven with other, more individualistic forms of positioning.

The Ideological uses of Perspective-sharing

Exactly how the combined visual and aural perspective are used in Featured Stories to position collective and individual identities requires qualitative analysis of particular Snaps and the stories in which they occur. The following qualitative analysis is used to answer our second research question: What ideological purposes might multimodal constructions of perspective serve? In our data, the combination of visual, auditory and verbal resources were used to emphasise the effects of personalisation and collectivisation which mapped broadly onto individual selfies and group selfies respectively. For individual selfies, the presented intersubjectivity invited the viewer to 'look at me, with me, and listen to me' using multimodal resources that emphasised personalisation of the speaker, personalisation which was reinforced by the verbal content spoken in these Snaps. This personalisation was put to well-recognised uses, for example, to position the speaker as an individual relative to others who might be positioned as an in-group or out-group. Example 1 illustrates this pattern, where the Snap creator uses the first person singular for herself, but uses plural forms for the non-present others, who are described as a functionalised (van Leeuwen, 2008) group, 'protestors', an out-group, 'they' from whom the speaker is differentiated and distanced.


	<p>Example 1:</p> <p>I'm sorry but the protestors are not (2.0) peaceful (.) they're violent (.) they're pushing people with disabilities just because they don't want anyone to cross to go to the inauguration</p>
---	---


In other cases, a speaker in an individual selfie might position his or herself within a larger, in-group from whom other individuals are differentiated. In example 2, the Snap creator starts by using the first person pronouns, but goes on to speak on behalf of a larger group, using the inclusive plural pronoun, 'our', differentiating herself from a non-present individual, 'this girl', who the speaker addresses using the second person pronoun 'you' and to whom the imperative in the caption, 'STAND UP FOR YOUR RIGHTS, GIRL' is also directed (although clearly visible to the wider viewing audience of Snapchat members).

	<p>Example 2:</p> <p>I just heard this girl complain about the protests and say 'oh I just don't understand how you could feel so passionate about something' (1.0) He's gonna be OUR president ↑ (.) How dare you</p>
--	---


In other cases, the individual shown in the selfie addressed the audience using the lexicogrammatical resources typical of synthetic personalisation (Fairclough, 1989), thereby allowing the viewer to feel that they are individually addressed whilst being part of a larger imagined community. In example 3 and example 4, the Snap creator addresses the audience using collective terms of address, 'you guys' and 'Snapchat', where the use of the second person pronoun 'you' is referentially ambiguous and can be interpreted as directed to the collective imagined audience of all those viewing the Story, but also to individual Snapchat members. After all, it is an individual who will 'swipe up' the screen on their phone and an individual who will participate by voting. These examples illustrate how synthetic personalisation (Fairclough, 1989) seeps into the discourse of the Featured Stories, and is used both by elite persons, such as the celebrity figure (Pharrell Williams) appealing for charity aid and lay persons, such as the citizen calling for others to exercise their rights to

vote. On the one hand, these examples might suggest that these calls to action within Snapchat can be made by anyone to the wider community, but the double call to action, that is, to identify also with the Snapchat community, is not free from ideological implications and is itself driven by commercial imperatives to consolidate and increase its membership in the face of competitor apps and services.

	<p>Example 3:</p> <p>So guys</p> <p>It's Pharrell here</p> <p>I'm going to be there with Ariana</p> <p>Supporting her in Manchester</p> <p>It's gonna be fun</p> <p>If you want to support her and the red cross all you got to do is swipe up</p> <p>See you guys there</p>
---	--

	<p>Example 4:</p> <p>Good morning Snapchat</p> <p>Man off to vote</p> <p>Make sure you vote</p> <p>Listen (.)</p> <p>Let your voice be heard man</p>
--	--

When we turn to the group selfies, the presented perspectives which invited the audience to share the social and spatial orientation of a group also emphasised the collective in-groups of which the Snap creator was part in the verbal content that was spoken. Thus in example 5, the speaker's repeated use of the plural pronoun, 'we' and the relational groups 'everyone' and 'all the families' emphasise the collective expression of grief and solidarity following the terrorist attack in Manchester in 2017.

	<p>Example 5:</p> <p>We're here tonight in Manchester to support all the families</p> <p>We're so sorry what happened</p> <p>and we just want to have a good time</p> <p>and remember everyone</p>
---	--

clapping, chanting, background noise and music, each of which serve to emphasise the response of the audience as a collective group, rather than differentiating individual social actors who are heard and seen. The aural perspectives found in these quasi-selfies thus positioned the viewing audience along with the Snap creator as co-consumers, sharing in a collective event along with others. This included watching a tennis match (background noise), watching a football match (clapping), watching a performance (instrumental music), and marking a memorial silence.

We have coined the term *synthetic collectivisation* to describe this emerging audio-visual form of positioning found in video-sharing. The term is a deliberate echo of Fairclough's (1989) concept, synthetic personalisation. 'Synthetic' is understood to imply the convergence which characterises the medium (Thurlow, 2013), where Snapchat blurs the distinction between public and private, and juxtaposes Snaps from 'ordinary people' alongside content created by mainstream media channels and celebrities. We retain the criticality implicit in the term, 'synthetic', for we do not interpret this form of representation to be neutral or authentic, but rather we argue that it is constructed strategically and put to ideological uses that serve socio-economic ends, typically for those in positions of power. In particular, we propose that these are focused on the strategic use of a discourse of 'us-ness'. We point to this through our term, *synthetic collectivisation* to indicate the construction of a group identity as an alternative to individualisation (cf van Leeuwen, 1996: 48).

Synthetic collectivisation is similar to synthetic personalisation, in that it uses semiotic resources to minimize social distance and effect solidarity between the producer and receiver of the Snap, in a discourse that is intended for a large audience. However, unlike synthetic personalisation which contrives personal communication directed at a mass audience, synthetic collectivisation contrives to position the individuals within the audience as if they were part of a larger group. For, as with other kinds of broadcast talk, the audience of Snapchat's Featured Stories are made up of individuals who access the content through their individual devices, and may well be geographically disparate and interact with the Story asynchronously (although within the 24 hour restriction in which the Featured Story is publicly available).

Synthetic collectivisation is also distinct from synthetic personalisation in its use of semiotic resources. Instead of the direct gaze of the visual address which typifies virtual visual synthetic personalization (Thompson 2012), the indirect perspectives constructed multimodally do not represent the Snap creator either visually or through aural resources. Instead, a group is shown in the foreground of the image, within the personal space of the

camera, and so by implication, the Snap creator. The aural perspective similarly invites the viewer to ‘listen with’ the Snap producer to resources which emphasise the social meanings of collectivisation, such as applause (which is typically a response of many persons). Synthetic collectivisation is thus used to counter the individualised experience of the Snapchat member and to promote the illusion that they are part of a group, together consuming a shared experience, both related to an individual Snap within a Featured Story, but also more generally of being part of the ‘Snapchat community’.

The emergence of synthetic collectivisation is by no means neutral, but is constrained and reproduces the power relations that are held in tension within other kinds of synthetic media. For example, on the one hand, the performance of audiencing enacted through the multimodal construction of sociality appears to give greater visibility and value to those spectating at an event alongside the elite persons performing at the event in question (singers, sports persons, speakers). However, this should not be taken as an empowering gesture, for there are commercial imperatives to interact with Snapchat, most obviously in terms of the advertising also included within some Featured stories, but perhaps also on the part of Snapchat itself to commodify its own community and thereby secure its commercial value in a competitive market of similar video and photo sharing apps and platforms. It is no surprise that the events at which the audience are celebrated were, for the most part (although not always), requiring commercial entry, such as concerts, sports events or fashion shows. The synthetic collectivisation is part of this wider discourse which produces and reinforces the citizen consumer.

More generally, the synthetic collectivisation shows the ongoing, ambiguous value of large scale sociality within apps like Snapchat and common to many other types of social media, (such as Friend lists in Facebook, Follower lists in Twitter and so on). The imagined presence of the Snap creator, participating with others at sociable events, is clearly a construction of pseudo-sociability (for no interaction with the person who created the Snap is possible via the Featured Story itself). Synthetic collectivisation is a subtle but seductive form of self-regulation that has replaced the selfie, oft-critiqued for its narcissism (Rettberg, 2014; Burns 2015). The demand to present an idealised self seems to have been replaced by the pro-social norms to present oneself as part of a legitimised group, a group whose perspective the imagined audience is also invited to share. Synthetic collectivisation is not devoid of the same micro-celebrity mind set associated with the individualism of personal selfies, but avoids the overt display of the individual whilst still enabling the Snap creator the gratification of visibility as indicated through the meta-data they receive, showing how many

times someone has viewed their Snap. In terms of the wider socio-cultural context, it is notable that synthetic collectivisation as found in the Snapchat Featured Stories excludes dissent, homogenizes the audience and endorses the in-group. Even in sports matches where you might expect to find one set of supporters opposing another, there is no such content included in the Featured Stories: The celebration of success is shown for both teams and there are equal invitations for the audience to share the perspective of each group. Likewise, synthetic collectivisation occurred frequently in the Featured Stories where group selfies also occurred: The Ariana and Friends Concert and The Women's March, both of which occurred immediately after socially disruptive events. The potential for synthetic collectivisation to construe solidarity in the face of wider social unrest might thus be regarded as a panacea that avoided documentation of the troubles in question.

Conclusions

In this study, we have provided the first account of how groups and individuals are positioned relative to one another in the videos and photographs that are produced in the Featured Stories of Snapchat as personal selfies, group selfies and quasi-selfies. The quantitative comparisons show that in this data, the invitation to share the perspective of a group is much stronger than the individualistic interpretation of selfies might suggest and that the ways in which sociality are constructed therefore deserve careful analytical and critical scrutiny. We have argued that in the case of Snapchat, the importance of these group perspectives is not necessarily a naturalistic representation but points to the highly strategic use of a discourse of 'us-ness' that promotes an idealised construction of the 'Snapchat community'. Our analysis suggests that the 'scales of sociality' that operate more widely across all forms of social media should not be taken as neutral or treated naïvely as a solely optimistic development. Instead, as discourse analysts, our empirical scrutiny of the construction of group perspective allows us to speak with confidence to wider macro-social concerns about the ends to which these 'scales of sociality' might be put.

The framework we have developed and applied goes beyond earlier monomodal accounts of selfies and demonstrates the importance for researchers to analyse how visual, aural and verbal resources are combined to create meaning in video-sharing. Although there are differences in the resources that can be used to create meaning in visual, aural and verbal semiotic systems, our framework shows that there is an underpinning use of interpersonal distance that allows us to make comparisons across and between the uses of different

multimodal resources. We adopt Hall's (1969) proxemic categories to point to the critical boundary between personal and other kinds of distance as a meaning-making resource in social video-sharing. When the position of the camera is constructed as within the personal space of the content in the video, this creates a strong invitation for the audience to share that same, filtered perspective on the content. This is not just a matter of spatial perspective but can be used to also examine how ideological perspectives are constructed. The framework is thus of particular value within Critical Discourse Studies as a tool to explore the ways in which interpersonal distance in multimodal texts can position in-groups and out-groups, legitimizing one rather than another depending on the context. It opens up questions about how multimodal positioning is created in other audio-visual genres that are commonly examined in discourse analysis and media studies, such as television broadcasts, photojournalism and citizen journalism. Further research is now possible about how the multimodal construction of in-groups and out-groups can be used to explore the aesthetics of specific kinds of collective activities, such as protest and participation. This research is all the more important given the increasing role of social media within activism of various kinds and across global contexts. The interdisciplinary potential for the framework also extends into fields such as psychology, where the framework and concepts in this paper could be used to explore a range of social identities, such as team identities or protest identities and the effect that this might have on a person's subsequent actions, such as their political engagement or commitment to a particular sports team. In this field, our text-based analysis would usefully stimulate other, more participant-centred forms of research, such as experimental design that tests the actual perception of interpersonal distance and its effect on sharing the group's perspective.

By systematically mapping the ways in which the visual, aural and verbal resources are used in video-sharing, our study has brought to light an emerging discursive strategy, which we call *synthetic collectivisation*. We have identified the visual and aural characteristics of synthetic collectivisation as indirectly positioning the video creator within a larger group, with whom the viewer of the video is invited to share the same visual and aural perspective. This was used to depict groups of many different kinds, such as protestors, mourners, fans at sports and music events, and occurred in different Featured Stories of Snapchat. However, this kind of positioning does not just occur in Featured Stories, but is also seen frequently within other areas of Snapchat, such as the local Stories and personal Stories that people create. Moreover, synthetic collectivisation is also beginning to be found in other kinds of videos and images shared via social media, for example used to depict

groups such as families and friends in adverts for wedding photography. In online picture galleries, similar visual constructions are used to depict crowds at festivals, public holidays that include spectacular displays (e.g. fireworks), demonstrations, and religious events such as Christian worship and Muslim prayer services. Given that these image banks are one means by which particular ideologies gain influence across many forms of media (including but not limited to social media), we should not underestimate just how potent and versatile the multimodal resources that encourage a viewer to share the perspective of a group can be. The subtle similarities and differences within synthetic collectivisation as it is used for different socio-cultural groups in particular deserves further scrutiny, especially as a counter-balance to the positioning of religious identities as out-groups of different kinds.

The concept of synthetic collectivisation is an important provocation to explore further how participatory culture is constructed across different social and cultural contexts. In particular, it calls for a nuanced approach to the audiencing that takes place through technologies like smart phones and suggests that there are emerging modes of addressing the ‘imagined audience’ (Marwick and boyd, 2011) that need further examination. In the case of the Snapchat stories, there may be more than one imagined audience which includes both the wider Snapchat community and the actual audiences who are participating in a particular event. The implications of synthetic collectivisation as it complicates and reframes the relationship between that individuals and larger social groups – self and society – remain to be seen.

The framework and concepts we have set out in this paper provide tools for discourse analysts, media and communication scholars and the visual arts, tools that are needed more than ever, given the increasing multimodal complexity of the interactions in which people engage through social media sites and services. Snapchat Featured stories are by no means the only context in which video is shared between individuals and groups. Other messaging services also include the option to share video clips, such as Whatsapp’s ‘moments’ (introduced in 2015) and Wechat’s ‘moments’. Instagram launched its ‘stories’ in 2016. Video-sharing sites continue to proliferate as do other video based genres such as gifs and video memes that are shared through other social network sites, which construct sociality in many ways that are shaped by their affordances and socio-cultural contexts. The scope and richness of the data that emerges from the vast number of interactions that take place through video-sharing is expanding the horizons for discourse studies in exciting, but challenging, new directions, for earlier frameworks that concentrate only on how group and individual identity are created verbally are insufficient for such a task. Our framework and methods

demonstrate how discourse analysts can make sense of this kind of data in a way that is transparent, systematic and rigorous, moving from micro-analytic observations (about types of intersubjectivity) to develop new theoretical concepts (namely, synthetic collectivisation). We believe that this will open up many rewarding lines of inquiry for discourse analysts and for media and communications scholars as we begin to explore the wide range and uses to which group identities are put, as these are constructed through selfies and other kinds of audio-visual genres.

Acknowledgements

The research in this paper began as a collaborative venture between Dr Ruth Page, Professor Jill Walker Rettberg and Dr Anders Fagerjord. Although as the paper developed, I became sole author, I am grateful to Professor Walker Rettberg and Dr Fagerjord for their assistance in developing the annotation manual, for creating Figures 1 and 2, drafting initial versions of the description of Snapchat and commenting on earlier versions of the paper. I am also grateful to Dr Matteo Fuoli who assisted me with the design of the inter-coder reliability calculations, and to the anonymous reviewers from the journal for their constructive feedback.

References

- Abidin C (2016) “Aren’t these just young, rich women doing vain things online?” Influencer selfies as subversive frivolity. *Social Media and Society*, April-June 2016: 1-17.
- Andreollo F (2017) The semeful sociability of digital memes: Visual communication as active and interactive conversation. Unpublished Doctoral thesis, University of Technology, Sydney.
- Arpaci I, Yalcin, S B, Baloglu M, Kesici, S (2018) The moderating effect of gender in the relationship between narcissism and selfie-posting behaviour. *Personality and Individual Differences* 134: 71-74.
- Bateman J, Delin JL, Henschel R (2004) Multimodality and empiricism: preparing for a corpus-based approach to the study of multimodal meaning-making. In: Ventola E, Cassily C, Kaltenbacher M, *Perspectives on Multimodality*, Amsterdam: John Benjamins, pp.65-87.
- Bucholtz M and Hall K (2005) Identity and interaction: a sociocultural linguistic approach. *Discourse Studies*, 7(4-5): 585 – 614.

- Burns A (2015) Self(ie)-discipline: Social regulation as enacted through the discussion of photographic practice. *International Journal of Communication* 9: 1716–1733.
- Chang L (2016) Snapchat users are watching 10 billion videos per day. *Digital Trends Blog*. Available at: <https://www.digitaltrends.com/social-media/snapchat-10-billion/>.
- Coates J (2003) *Men Talk: Stories in the Making of Masculinities*. Oxford: Blackwell.
- Collins K and Dockwray R (2015) Sonic proxemics and the art of persuasion. *Leonardo Music Journal* 25: 53-56.
- Culpeper J (2011) *Impoliteness: Using Language to Cause Offence*. Cambridge: CUP.
- Deller R and Tilton S (2015) Selfies as charitable meme: Charity and national identity in the #nomakeupselfie and #thumbsupforstephen campaigns. *International Journal of Communication* 9: 1788-1805.
- Dodson C (2015) Why Snapchat's Live Stories are the most powerful new social media. *Fast Company Blog*. Available at: <https://www.fastcompany.com/3052322/why-snapchats-live-stories-are-the-most-powerful-new-social-media>.
- Eagar T and Dann S (2016). Classifying the narrated# selfie: genre typing human-branding activity. *European Journal of Marketing* 50(9-10): 1835-1857.
- Fairclough N (1989) *Language and Power*. Longman: London and New York.
- Forceville C (1999) Educating the eye? Kress and Van Leeuwen's Reading Images: The Grammar of Visual Design (1996) *Language and Literature* 8(2): 163 – 178.
- Genette G (1980) *Narrative Discourse: An Essay in Method*. Translated by Jane E. Lewin. Ithaca New York: Cornell University Press.
- Georgakopoulou A (2016) From narrating the self to posting selfies: A small stories approach to selfies. *Open Linguistics* 2: 300-317.
- Gillespie A and Cornish F (2009) Intersubjectivity: towards a dialogical analysis. *Journal for the Theory of Social Behaviour* 40(1): 19-46.
- Gillespie T (2010) The politics of 'platforms'. *New Media and Society* 12 (3): 347-364.
- Guardian Newspaper (2017) Don't Look Back in Anger becomes symbol of Manchester's spirit. Available at: <https://www.theguardian.com/uk-news/2017/may/25/dont-look-back-in-anger-becomes-symbol-of-manchester-spirit>
- Halliday MAK (1994) *An Introduction to Functional Grammar*, 2nd edition. London: Edward Arnold.
- Hall ET (1969) *The Hidden Dimension*. Garden City, New York: Random House.

- Herring S (2007) A faceted classification scheme for computer-mediated discourse. *Language@internet* 4. Available at <http://www.languageatinternet.org/articles/2007/761>.
- Klandermans PG (2014) Identity politics and politicized identities: Identity processes and the dynamics of protest. *Political Psychology* 35(1): 1-22.
- Kress G and van Leeuwen T (1996) *Reading Images: The Grammar of Visual Design*. London: Routledge.
- Kress G and van Leeuwen T (2001) *Multimodal Discourse*. London: Bloomsbury.
- Kuntsman A (2017) *Selfie Citizenship*. Basingstoke: Palgrave Macmillan.
- Lambert Graham S (2015) Relationality, friendship, and identity in digital communication. In: Georgakopoulou A and Spilioti S (eds.) *Handbook of Language and Digital Communication*. London: Routledge, pp.305-320.
- Landis, JR and Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33(1): 159-174.
- Lewis T (2018) Digital food: from paddock to platform. *Communication Research and Practice* 4(3): 212-228.
- Marquardt N and Greenberg S (2012) Informing the design of proxemic interactions. *Pervasive Computing* 11 (2): 14-23.
- Marwick A and boyd d (2010) I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society* 13 (1): 114-133.
- Miller D, Costa E, Haynes B, McDonald T, Nicolescu R, Sinanan J, Spyer J, Venkratraman S, Wang X (2016) *How the World Changed Social Media*. London: UCL Press.
- O'Donnell M (2008) Demonstration of the UAM Corpus Tool for text and image annotation. Proceedings of the ACL-08: HLT Demo Session, Columbus Ohio, June 2008. Association for Computational Linguistics, pp.13-16.
- Paulus T, Woods M, Atkins D and Macklin R. (2017). The discourse of QDAS: Reporting practices of ATLAS.ti and NVivo users with implications for best practices. *International Journal of Social Research Methodology* 20(1): 35-47.
- Reich W (2010) Three problems of intersubjectivity and one solution. *Sociological Theory* 28(1): 40-63.
- Rettberg JW (2014) *Seeing Ourselves through Technology: How we use Selfies, Blogs and Wearable Devices to See and Shape Ourselves*. Basingstoke: Palgrave.
- Rizzolatti G, Fadiga L, Fogassi L, Gallese V (1997) The space around us. *Science* 277(5323): 190-191.

- Rovisco M and Veneti A (2017), Picturing protest: visibility, visibility and the public sphere, *Visual Communication* 16(3): 271 – 277.
- Scollon R and Scollon, SW (2003) *Discourses in Place: Language in the Material World*. London and New York: Routledge.
- Simon B (2004) *Identity in Modern Society. A Social Psychological Perspective*. Blackwell, Oxford.
- Smith, A and Anderson M (2018) Social media use in 2018. Pew Research Center, March 2018. Available at: <http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/>
- Spencer-Oatey H (2005) (Im)Politeness, face and perceptions of rapport: Unpackaging their bases and interrelationships. *Journal of Politeness Research* 1(1): 95-119.
- Steinfeld C, Ellison NB, Lampe C (2008) Social capital, self esteem and use of online social network sites: A longitudinal analysis. *Journal of Applied Developmental Psychology* 29(6): 434-445.
- Thompson R (2012) Looking healthy: visualizing mental health and illness online. *Visual Communication* 11(4): 395 – 420.
- Van Dijck J (2013) *The Culture of Connectivity: A Critical History of Social Media*. Oxford: OUP.
- Van Leeuwen T (1999) *Speech, Music, Sound*. Basingstoke: Macmillan.
- Van Leeuwen T (2008) *Discourse and Practice: New Tools for Critical Discourse Analysis*. Oxford: OUP.
- Veum A and Undrum LVM (2018) The selfie as a global discourse. *Discourse & Society* 29(1): 86 – 103.
- Zappavigna M (2016) Social media photography: construing subjectivity in Instagram images. *Visual Communication* 15 (3): 271 – 292.
- Zappavigna M and Zhao S (2017) Selfies in ‘mommyblogging’: An emerging visual genre. *Discourse, Context and Media* 20: 239 – 247.
- Zhao S and Zappavigna Z (2018) Beyond the self: Intersubjectivity and the social semiotic interpretation of the selfie. *New Media & Society* 20(5): 1735-1754.

¹ We acknowledge that it is the construction of the selfie that leads to the interpretation that the person shown is the same person as the selfie-creator. In some cases, such as 'fake selfies' the image is not created by the person actually taking the image. However, the visual conventions of the selfie as a digital self portrait are so strong that we retain the definition of a selfie as a 'digital self portrait', but point to the construction of identities through this emerging genre.