Robin I. M. Dunbar* and Emma Stirling-Middleton

Why cartoons make (some of) us smile

https://doi.org/10.1515/humor-2023-0111
Received January 27, 2023; accepted September 23, 2023; published online January 5, 2024

Abstract: Pocket cartoons are a regular feature of most contemporary newspapers and magazines. As such, they represent a way of conveying complex social and political commentary in a simple visual form. How well we enjoy verbal (oral) jokes depends on the number of mindstates in the joke, and here we ask whether this is also true of visual cartoons. We use survey data from a sample of 3,380 participants attending a public exhibition of published print media cartoons by well-known cartoonists to determine the extent to which viewers' ratings of cartoons are determined by the mentalizing content of cartoons, the participants' gender and age, and the publication date of the cartoon. We show that the number of mindstates involved in the cartoon affects its appreciation, just as in verbal jokes. In addition, we show that preferred topics vary by age and gender. While both genders strongly prefer cartoons that explore the complexities of romantic relationships, men rate visual jokes more highly than women do, whereas women prefer jokes that involve political commentary or the dynamics of close relationships. These differences seem to reflect differences in the way the social worlds of the two genders are organized.

Keywords: cartoons; mentalizing; gender differences; social commentary

1 Introduction

Humor and laughter have been the subject of considerable interest and research for at least two millennia (albeit, historically, mainly by philosophers). Much of this interest has, however, focused on verbal rather than pictorial forms of humor: what is it about a joke that makes it humorous? In part, this may reflect the fact that, as a form of joke, cartoons (not to be confused with artists' preparatory sketches or drawn caricatures) are of relatively recent origin. Although they first emerged as editorial commentary in eighteenth century broadsheets, with the first cartoon in the modern sense published in 1843, it was not until 1939 that the familiar 'pocket cartoon' (a single column, often single panel, drawing) first appeared

***Corresponding author: Robin I. M. Dunbar**, Department of Experimental Psychology, Radcliffe Quarter, University of Oxford, Oxford OX2 6GG, UK, E-mail: robin.dunbar@psy.ox.ac.uk
Emma Stirling-Middleton, The Cartoon Museum, 63 Wells Street, London W1A 3AE, UK

Open Access. © 2023 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

(Whitworth 2020). These kinds of cartoons typically focus on social or political commentary, usually by drawing attention to the foibles of a particular individual or social class. In this respect, they share a great deal in common with literary satire.

Cartoons in this sense (as opposed to idle doodling or artistic sketching – or, for that matter, storytelling forms such as comic books and Disney-type films) are a very distinct form of humor in that the joke takes a strongly visual form, with the verbal component (if any) often only providing contextualizing cues, usually in the form of brief dialogue or as captions that identify the individual(s) involved, the context or an historical allusion. These condensed forms of joke encapsulate a storyline in very compressed form that requires minimal need for reflection by the viewer – who usually (but not always) 'gets' the joke straight away. As with verbal jokes, if you don't 'get' the point at once, it is rarely amusing. This is because the viewer has to be familiar with the butt of the joke, and to be aware of associated contemporary social or political concerns. This requirement is especially intrusive in political cartoons, where the viewer may require considerable background knowledge not just of the political context of the day but also of historical events that are being exploited as metaphors (Pinar-Sanz 2020).

Aside from philosophical deliberations on the nature of humor, much of the focus of research has been on the pragmatics of how jokes work (Carbajal-Carrera and Sanchez-Castro 2020). The best jokes are thought to build up a set of expectations in the listener, with laughter triggered by the way the punchline violates these expectations (Canestrari and Bianchi 2012; Fortabosco 2008; Hurley et al. 2011; Ramachandran 1998; Suls 1972). An important feature of the ability to appreciate humor, and indeed any other kind of fiction, is the capacity to mentalize (i.e. understand others' mindstates) (Dunbar 2005; Dunbar et al. 2016; Howe 2002). However, mentalizing has not typically featured prominently in studies of humor.

Mentalizing (otherwise known as mindreading or 'theory of mind') is the capacity to understand the intentional states of other individuals, and is associated with the capacity to use 'intentional' terms (words like *believe, suppose, wonder, imagine, think* or *intend*) (Dennett 1988). Formal theory of mind (second order intentionality: I *suppose* that you *think* that …) appears in five-year-old children. However, the typical adult can manage as many of five orders of intentionality (Kinderman et al. 1998; Powell et al. 2010; Stiller and Dunbar 2007). A typical fifth order intentional statement would be: I *wonder* whether Peter *intends* that Jane *understands* why Rosie *wants* James to *believe* that …. (where the intentional terms are italicized and the dots at the end specify some factual statement). Both language production and comprehension depend on the ability to master high order mentalizing skills, with individual differences in achievable mentalizing level correlating with the complexity of the sentences the individual can parse correctly (Oesch and

Dunbar 2017). Mentalizing also limits the number of people that can take part in an informal conversation (Krems et al. 2016). Mentalizing, and especially high-order mentalizing, is cognitively very demanding, with the number of mental states an individual can manage at any one time being determined by the volumetric size of key brain regions in the prefrontal cortex and associated areas in the parietal and temporal lobes that form the mentalizing (or default mode) neural network (Lewis et al. 2011; Powell et al. 2010).

Theory of mind has been shown to play an important role in the understanding of cartoons (Griffin et al. 2006), with appreciation of cartoons disrupted by lesions in brain regions associated with mindreading (notably the right prefrontal cortex: Shammi and Stuss 1999; Griffin et al. 2006). More recently, studies have suggested that verbal jokes that involve more mindstates (in effect, characters) are perceived as being funnier than jokes that involve fewer mindstates up to a limit at fifth order (Dunbar et al. 2016). Sixth and seventh order verbal jokes are not only very rare but are viewed as being much less funny than fifth order jokes (Dunbar et al. 2016), mainly because it is difficult to integrate all the different mindstates involved. A similar effect has been noted in respect of fictional stories: stories that have higher mentalizing content (but do not exceed the mentalizing capacity of the reader) are perceived as being more engaging and more enjoyable (Carney et al. 2014; Dunbar 2005, 2017; Zunshine 2022). When there are too many mindstates involved, the story becomes too confusing. As a result, most great novelists and dramatists are very careful to keep their stories within the limits of their typical reader/audience (Dunbar 2005; Krems and Dunbar 2013; Krems et al. 2016; Stiller et al. 2003).

In order to allow the listener to recognize and resolve incongruency at the punchline, verbal humor involves manipulating the expectations of the listener via the mindstates of the characters in the joke (Canestrari and Bianchi 2012; Fortabosco 2008; Suls 1972). Oral jokes are thus an implicit two-way interaction between a storyteller (the comedian) and the audience, and hence necessarily involve two 'structural' mindstates (the listener and the speaker) before we get to the joke itself (Dunbar et al. 2016). In most cases, the speaker's presence is implicit in that we listen to a real person delivering the joke; however, many comedians explicitly interpolate their own mindstates into their jokes as though they are themselves part of the joke. In contrast, the artist is conspicuously absent as an intermediary in cartoons – indeed, we may not even know who the cartoonist is, other than as a signature at the bottom of the picture that provides little guidance about the individual concerned. Since this allows the viewer to enter directly into the story on the page, thereby freeing up one of the 'structural' mentalizing levels, a cartoon ought to be able to include one more mentalizing level in the story than is possible with oral jokes. However, the need to place the complete storyline in a single visual image with no time depth may impose a structural constraint on cartoons that is not present in oral

jokes (where, in principle, an infinite number of characters and events can be introduced, given enough time). This may limit how many mentalizing states can be incorporated into a cartoon. Indeed, an attempt to produce a cartoon version of a high order mentalizing task without using text found that it was only possible to do this as a fourth order task (the reader/observer plus three characters in the cartoon) (Barrett et al. 2002, Figure 11.1, p. 296).

To explore responses to cartoons, we took advantage of an exhibition mounted by a national museum to obtain data on how people rate published cartoons by well-known cartoonists. Our primary hypotheses are: (1) that cartoons, like oral jokes, are perceived as being funnier the more demanding they are of the viewer's mentalizing competences; and (2) that there is an upper limit on mentalizing content beyond which the joke becomes too complex to parse, with this limit being (2a) higher than in oral jokes if the constraint is simply the number of mentalizing levels available to the observer or (2b) lower than in oral jokes if the pictorial context imposes a constraint on the cartoonist's ability to include more characters. Since political cartoons, in particular, are likely to require knowledge of the time and context they refer to, we test (3a) whether cartoons with older dates of publication will be appreciated less than ones published more recently and (3b) whether preferences are influenced by participant age and gender. Finally, (4) since there are well known gender differences in social and relationship characteristics (Benenson 2014; Dunbar and Machin 2014; Dunbar 2018; Machin and Dunbar 2013), we test whether there are gender differences in preferences for social versus non-social (e.g., slapstick) topics: women typically have better mentalizing skills than men, so should be more likely to appreciate cartoons that involve social commentary.

2 Methods

The study formed part of an exhibition curated by The Cartoon Museum, London, to share with the public its collection of original cartoons published in mainstream UK newspapers and magazines between 1930 and 2010. The fact that this was an exhibition that members of the public paid a fee to view placed some constraints on the design of the study since it was important to minimize the burden on participants who were attending for pleasure rather than to take part in an experiment. Selection of cartoons was, therefore, necessarily dictated by the demands of the exhibition and not the demands of the experiment as such, and reflected the Museum's requirement that cartoons be both engaging for the casual viewer and representative of different time periods.

The exhibition consisted of 38 cartoons arranged in pairs, selected from an initial shortlist of 47. The selection was made by one author (ESM) in the order in which the cartoons were organized in the Museum's archive so as to reduce the influence of

personal preferences influencing the choice. Copies of all the cartoons are shown in the online Supplementary Information (file *Cartoons_SI.xlsx*). Most cartoons are multimodal (i.e., contain text as well as artwork), with the textual element (e.g., speech bubbles or caption text at the foot of the picture) usually providing the key information for the joke to work (Agüero Guerra 2016). However, five of the cartoons in our selection (13 %) contained no text at all, relying entirely on the visual artwork for their effect, although two of these (cartoons 7A and 13B) had signposting (labels) within the cartoon that provided context.

The mentalizing level for each cartoon was assessed, and agreed, by both authors prior to the exhibition, following the rubric applied in the previous study of verbal jokes (Dunbar et al. 2016). For present purposes, the mentalizing level of a cartoon is equivalent to the number of actively engaged characters in the cartoon (a crowd paying attention to the focal point of the cartoon was counted as a single mentalizing level providing it was deemed to be part of, or essential to, the joke itself), plus an additional level for the viewer (participant). Thus, a cartoon with two characters engaged in conversation would count as a third order mentalizing event (two mindstates in the cartoon plus the viewer's own mindstate). Cartoon 1A (see online Supplementary Information), for example, was assigned fourth order mentalizing because the speaking character (a rather bossy and forthright woman) is addressing the mermaid but at the same time signaling to her companion (her henpecked husband) that he will be in deep trouble should he show even a flicker of interest in the mermaid's coquettish advances. So, in this case, there are three active mindstates within the cartoon in addition to the observer. Without the companion, the cartoon would have much less "bite" as social commentary; the companion creates the context for virtue signaling by the speaker. Cartoon 1B, in contrast, is assigned only second order (the man and the viewer): the children are merely providing the context (and, like all children, are not really paying attention), while the second adult is not engaged with the speaker. It is a commentary on how we tend to exaggerate the "golden" past (hence the over-sized pot). Cartoons were arranged in pairs such that they were either of the same or different mentalizing levels. Previous experience indicates that ratings by independent observers are usually extremely high (~98 % agreement, and never more than one mindstate different: Dunbar et al. 2016).

The cartoons were also classified by content into six topic categories (Table 1). These categories partially overlap with those given by Pedrazzini and Scheuer (2018), who used a factor analysis of a sample of cartoons to develop a four-category classification; however, for present purposes, four categories was felt to be too limiting to do justice to the topics in our sample. Table 2 lists the pairings by topic, date of original publication and assigned mentalizing level.

Table 1: Classification of cartoon topics.

Type	Definition
Domestic dynamics	Jokes that involve domestic, marital or romantic relationships (e.g., lawyer reading out will to family after death: 'To my wife whom I promised to mention in my will: Hello Edna': cartoon 15B)
Social commentary	Poking fun at the social foibles, attitudes or behavior of particular classes of individual (e.g., men's fashion for aftershave: Edvard Munch's "The Scream" with head surrounded by white light redolent of pain and text reading "Aftershave – the truth": cartoon 11A)
Political commentary	Comments on political leaders or individuals in the public eye, usually with respect to specific political events (e.g., Nuremberg Rally with 'smiley face' kite being flown from rear of crowd, and text "A member of the Hitler Youth is determined to enjoy his day out": cartoon 10B)
Verbal jokes	Jokes that depend on word meanings or metaphors, puns or wordplay (e.g., passer-by comments that Diogenes sitting in his barrel is "a barrel of laughs": cartoon 9A)
Visual jokes	Jokes (often with no text) where the humor lies in the metaphor created by a juxtaposition of events or phenomena that wouldn't normally go together (e.g., a Chinese parade 'dragon' enters a restaurant where a long table is laid out in bends ready for it: cartoon 9B)
Situational jokes/ slapstick	Jokes that depend on unexpected events or contexts (e.g., painter and paint tins falling from roof onto well dressed characters walking underneath as one says "Don't you sometimes just yearn for a bit of honest, obvious slapstick?": cartoon 18B)

The cartoons, and their pairings, are shown in the online Supplementary Information, and gives the wording on each cartoon, its author and date of publication, its assigned mentalizing level and its topic.

The data were collected between mid-November 2021 and the end of April 2022. Viewers of the exhibition were invited to take part in the study on a voluntary basis. A total of 3,670 did so. So as to minimize the load on viewers of the exhibition, participants were simply asked to state, on a paper questionnaire, which cartoon of each pair they thought was the funnier as they viewed it. Before viewing the exhibition, participants completed a written consent form and an anonymous background questionnaire that requested information only on age and gender. Ages were converted to decadal categories (18–29 = 20s; 30–39 = 30s; 40–49 = 40s, etc.). Participants who declared their age to be <18 years or did not give their age, who did not declare their gender as either male or female, or did not rate all 19 cartoon pairs were excluded from the analysis, leaving a final sample of 3,380 adults (age range 18–88 years; 1,740 female), representing 92 % of the original sample.

Because of the constraints of the museum exhibit, the data are in the form of summed preferences of one cartoon (always referred to as 'A,' defined by the listed

Table 2: Cartoon pairings and their characteristics.

Pairing	Cartoon 'A'			Cartoon 'B'		
	Topic	Date[b]	Mentalizing level[c]	Topic	Date[b]	Mentalizing level[c]
1	Social	c.1930	4	Social	1940	3
2	Political	2007	2	Political	1961	4
3	Political	1958	3	Verbal	2003	2
4	Verbal	1950	2	Situational	1960	3
5	Social	1985	3	Political	1990	3
6	Domestic	1967	3	Verbal	1951	4
7	Verbal	1979	1	Verbal	1950	3
8	Political	1970	3	Situational	1980	2
9	Verbal	1969	3	[a]Visual	1988	2
10	Verbal	1990	3	Political	1954	2
11	Social	1967	1	Situational	1952	3
12	Social	1969	3	Verbal	1930	3
13	Domestic	1970	3	Verbal	1976	2
14	Political	1964	3	[a]Situational	1976	2
15	Social	1950	3	Domestic	2000	4
16	Domestic	1986	3	Social	2000	4
17	Verbal	1965	4	Social	1980	3
18	[a]Political	1984	2	[a]Situational	1942	3
19	Political	1950	3	Visual	1988	3

The cartoons themselves are shown in the online Supplementary Information. [a]Purely visual cartoons with no accompanying text. [b]Publication date. [c]Mentalizing level (including the viewer's own mindstate): the mentalizing structure of each cartoon is given in the online Supplementary Information.

order in their pairing) over the other cartoon ("B") for each decadal age group, for each pair of cartoons. We first compare the difference in the relative preference with the difference in mentalizing levels for the two cartoons (Hypothesis 1) using a GLM with proportion of preferences for cartoon A as the dependent variable, and gender and the difference in mentalizing score (cartoon A minus cartoon B) as predictor variables. We next test for a difference in mentalizing level between cartoons and verbal jokes (Hypothesis 2) using a Mann-Whitney non-parametric test because the distribution of verbal jokes is bimodal and hence not normally distributed. We then test for potential confounds in respect of the decadal age and gender of the participants, and the cartoon's date of publication (Hypothesis 3) by regression analysis of the summed preferences for cartoon A relative to B. Finally, we compare gender differences in preference for jokes of different type (Hypothesis 4) by plotting the summed votes for specific cartoon types for the two genders against each other (i.e. mean male rating against mean female rating for each pair of cartoons and each topic), with the main diagonal defining the null hypothesis (where both genders vote

equally for the same type of joke). Here, we seek to identify those pairings or topics that are significantly offline (i.e. lie outside the 95 % CIs for the regression line).

All analyses were executed in SPSS v.29.

3 Results

We first test whether the participants' preference ratings for cartoons are influenced by the mentalizing content of the cartoon (Hypothesis 1). Figure 1 plots the preference for cartoon 'A' over cartoon 'B' as a function of the difference in mentalizing score between the two cartoons, for the two genders separately. The unit of analysis is the proportion of participants in each decadal age group who express a preference for cartoon 'A'. Cases where the mentalizing level is the same in the two cartoons provide us with a reality check since, if preference is dictated by mentalizing level, then the preference for one over the other should be random when the cartoons do not differ in mentalizing level. In other words, in this case, the preference for 'A' over 'B' should cluster around the dashed line at 50 %, as indeed is the case. In contrast, the 95 % CIs for the estimates lie well outside the random expectation when the cartoons differ in their mentalizing demands, with the magnitude of the difference being larger the greater the difference in mentalizing levels. These results are independent of gender. A GLM with proportion of preferences for cartoon A (for individual gender and

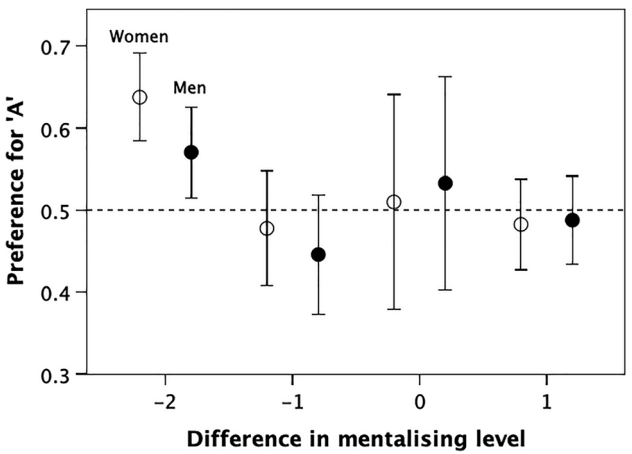

Figure 1: Mean (±95 % CIs) of preference for cartoon 'A' over cartoon 'B' within decadal age classes as a function of the difference in mentalizing level between the two cartoons ('A' minus 'B'). Dashed line indicates random choice. Sample sizes at each level are, L to R: 73, 71, 71, 12. Number of participants is the same for all four categories (1,740 women and 1,640 men).

decadal age categories) on each of the 19 pairs as the dependent variable, and gender and the difference in mentalizing score (cartoon A minus cartoon B) as predictor variables, gives a significant effect for mentalizing difference ($F_{3,220} = 9.95$, $p < 0.001$), but no effect for gender ($F_{1,220} = 0.80$, $p = 0.416$) and no interaction effect ($F_{3,220} = 0.46$, $p = 0.710$). Hypothesis 1 is thus confirmed.

We next test whether the distribution of mentalizing levels in this sample of cartoons is the same as that observed in a sample of verbal jokes (Hypothesis 2). The distribution of mentalizing levels in the 38 cartoons ranged between 2 and 4 (mean = 2.82 ± 0.77SD). The range in mentalizing level is lower than that observed in oral (verbal) jokes, where the mean for a standard sample was 4.27 ± 1.05SD (range 3–7, $N = 101$) (Figure 2). Verbal jokes include the mindstate of the comedian as well as the observer, whereas the mind of the cartoonist is assumed not to intervene between the mind of the observer and the characters in a cartoon. However, even allowing for this extra 'structural' level, the mentalizing levels for the cartoons are significantly lower than those in the sample of verbal jokes (Mann–Whitney test, $N_1 = 38$, $N_2 = 101$, $p = 0.002$). Hypothesis 2 is confirmed (there does seem to be a limit in how cognitively demanding cartoons can be, with this limit being at about third order intentionality), with hypothesis 2a being rejected in favor of Hypothesis 2b (the availability of a 'spare' mentalizing level does not allow the cartoonist cognitive space to increase the complexity of the joke). This suggests that the single-panel 2D format of the cartoon makes it difficult for the cartoonist to represent many mindstates at the same time.

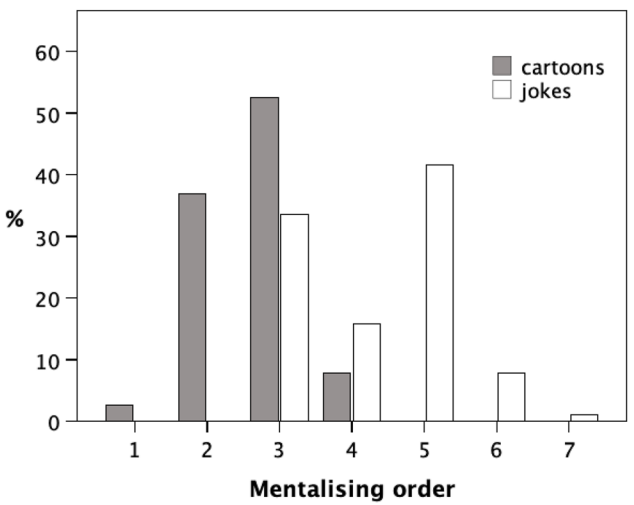

Figure 2: Proportional distribution of mentalizing levels in the sample of cartoons and the 100 Best Jokes sample (from Dunbar et al. 2016).

To check for possible confounds, we ask whether the differences in preference are influenced by age, gender or date of publication (Hypothesis 3). To do this, we first plot preference for 'A' against decadal age for the two genders separately. Although these vary considerably across the 19 pairs, they can be reduced to three main patterns (illustrated in Figure 3; with all 19 individual comparisons shown in the Supplementary Information). In Figure 3a (pairing 2), the two genders map directly on top of each other. There were four other cases (pairings 11, 12, 13, 15) where this was also the case, but these all had a nonlinear form (an exponential decline in the first three cases, an exponential increase in the fourth case). Pairings 4, 7, 8, 10, 16 and 17 exhibited a pattern similar to that in Figure 3b (pairing 3), with one gender lying consistently above the other. There was no consistent overall tendency for one gender to score higher than the other. Four pairings (6, 14, 18 and 19) were similar to Figure 3c (pairing 9), where the two genders exhibited diametrically opposite quadratic relationships, again with no tendency for one gender to dominate in any way. In two additional cases (pairings 1 and 5), the two genders were indistinguishable, with no correlation with decadal age in either case. In general, there were no consistent age effects: eight of the pairings exhibited a general positive slope with age, and nine had a negative slope, with two exhibiting no change.

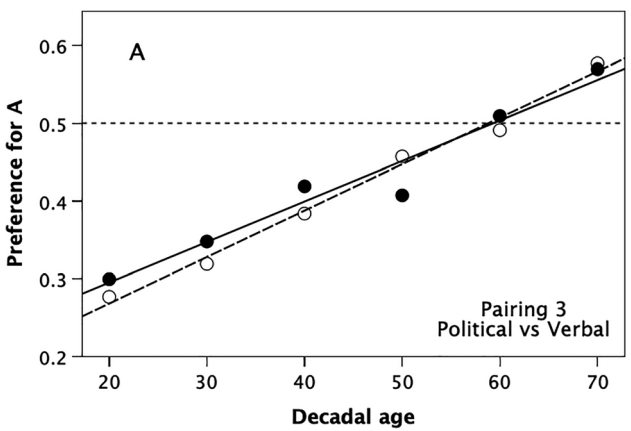

Figure 3a: Preference of cartoon 'A' in each pairing plotted by decadal age for each gender. The three graphs illustrate the three different patterns that emerge: (a) complete agreement between the genders (five pairings, with two other pairings that showed no correlation with age for either gender); (b) parallel patterns for the two genders, with one gender exhibiting a consistently higher preference than the other (seven pairings); (c) the two genders exhibit inverse quadratic preferences with age (five pairings). The pairing illustrated is indicated in the lower corner in each case, along with the topics for the two cartoons (listing 'A' first). Women: unfilled symbols; men: filled symbols. Sample sizes for decadal categories (both genders combined): 1,288, 671, 550, 450, 271 and 150.

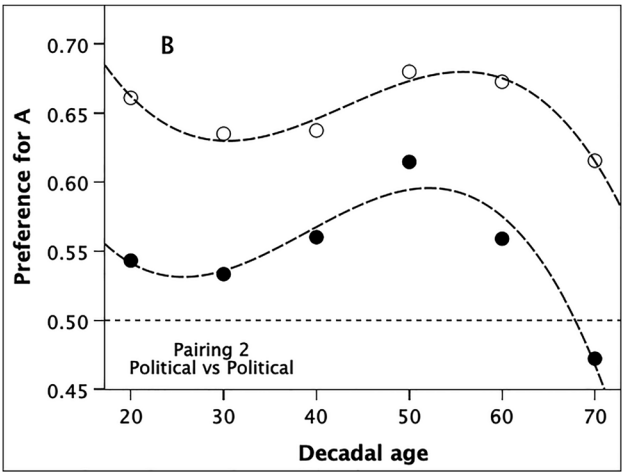

Figure 3b: Parallel patterns for the two genders, with one gender exhibiting a consistently higher preference than the other (seven pairings).

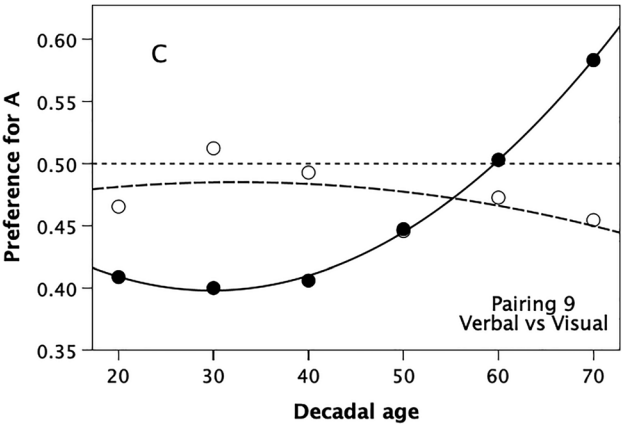

Figure 3c: The two genders exhibit inverse quadratic preferences with age (five pairings).

To check for possible confounds due to date of publication, we ran a multiple regression with preference for cartoon 'A' as the dependent variable, and the difference in date of publication ('A' minus 'B'), and decadal age and gender of participants as predictors. Only the difference in publication date was significant (full model: $F_{3,224} = 29.73$, $p << 0.0001$; difference in publication date: $t = 9.42$, $p < 0.0001$; participant decadal age: $t = -0.30$, $p = 0.765$; participant gender: $t = -0.58$, $p = 0.562$). In most cases, there is no strong effect due to age, but where there was a marked

preference it was the more recent of the two cartoons that was preferred (6/7 of the cases with a strong preference: $t_6 = 2.75$, $p = 0.033$). This appeared to be true as much for older participants as it was for younger ones. As a final check, we ran an ANOVA with cartoon mentalizing level as a function of topic: there were no significant differences in mentalizing levels between the six topics ($F_{5,32} = 0.51$, $p = 0.766$).

The gender differences in Figure 3 lead into our final question: what do the data tell us about gender differences in humor preference (Hypothesis 4)? Figure 4 plots the relative preferences for cartoon 'A' in each pairing for the two genders. Overall, there is very close agreement between the preferences of the two genders ($r = 0.948$, $N = 19$ pairings, $p < 0.0001$), with the regression line mapping very closely onto the main diagonal where men's preference for cartoon 'A' is the same as women's preference for cartoon 'A'. In two cases, both genders show a strong preference for cartoon 'B' (pairings 13, 15, 19 in the lower left quadrant), while in two other cases (pairings 1 and 12, in the upper right quadrant), but arguably four (pairings 6 and 11 in addition), they both exhibit strong preference for cartoon 'A'. In six cases (pairings 2, 8, 10, 16, 17 and 18), the datapoints lie outside the 95 % CIs for the regression,

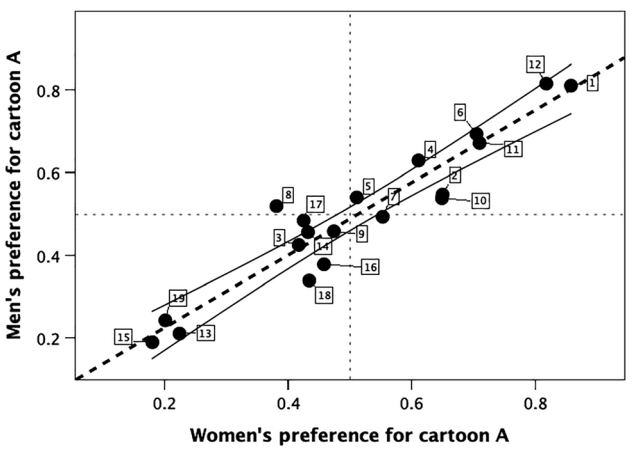

Figure 4: Preference for cartoon 'A' in each pairing for men plotted against that for women. The heavy dashed line is the least squares regression line (with 95 % CIs as thin solid lines either side); the regression line is virtually identical to the line of equality where the two genders agree. Values above the line indicate that, relative to cartoon 'B', men prefer cartoon 'A' more than women do; values below the line that men prefer cartoon 'A' less than women do. The fine dashed vertical and horizontal lines indicate the point of indecision in the preference of 'A' over 'B' for each gender (horizontal for men; vertical for women). The OLS regression passes through this point (as it should do if this is the point of indecision). In the lower left quadrant, both genders prefer 'B'; in the upper right quadrant, both genders prefer 'A'. The numbers against the datapoints identify the cartoon pairings (as given in the Supplementary Information).

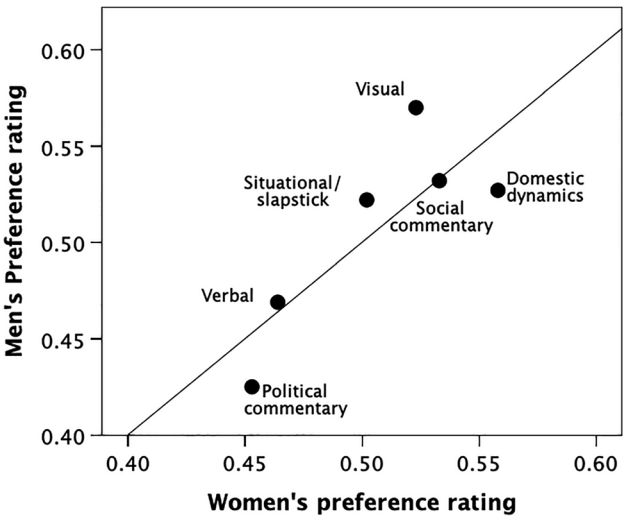

Figure 5: Mean preference ratings for the six topics for men plotted against the equivalent values for women. The diagonal line demarcates the line of equivalence (the two genders give the same rating).

indicating that the two genders differ significantly in their preferences. In these cases, women broadly to prefer social and verbal jokes whereas men tend to prefer situational jokes.

To explore the gender differences in more detail, we summed the votes for each topic, irrespective of pairing, and calculated an overall topic preference score for each gender for the six topics. Figure 5 plots men's preferences against the equivalent preference for women for each of the topics. The main diagonal (the line of equivalence, where both genders give equal weighting to the topic) is shown as the dashed line. When a datapoint lies off the line, this signifies a stronger preference for that joke type by the indicated gender. Both genders clearly consider visual jokes (puns or wordplay), social commentary and jokes about domestic situations (usually marital relationships) funnier than any other topic, and are least enthusiastic about political jokes. The first three are probably the most immune to the effects of fashion, whereas political jokes are most likely to reflect particular times and circumstances. Overall, the differences between genders are modest (only the preference for visual jokes differs significantly between the two genders), mainly because, with such a small sample, the CIs are very wide. Broadly speaking, women have a slightly greater preference for jokes about domestic situations and political commentary, whereas men exhibit a slightly greater preference for situational (slapstick) and visual jokes.

4 Discussion

Our main findings are (1) that pocket cartoons are perceived as being more amusing the more challenging they are in terms of mentalizing (mindreading); (2) that there is an upper limit on the mentalizing quality of jokes, with this limit being lower for visual cartoons than for verbal (i.e. oral) jokes; (3) that, although there is a significant effect on preferences of date of publication (more recent cartoons are preferred), there is no consistent effect due to the age of participants and, moreover, the joke topics themselves do not vary in mentalizing levels; and, finally, (4) that there are interesting gender differences in the preference for some types of joke. The significant effect of publication date confirms what is probably intuitively well appreciated, namely that fashions change as much in cartoon topics as they do in other aspects of our social, political and cultural life. Age of viewer had an effect, but not always in a simple way and there was often a strong interaction with gender.

The mentalizing effect parallels that found in both oral jokes (Dunbar et al. 2016) and fictional stories (Carney et al. 2014). In all three cases, appreciation increases with the amount of mentalizing work that has to be done to parse the storyline, but declines once the required mentalizing level exceeds the natural capacities of normal adults (fifth order intentionality). In both verbal jokes and cartoons, this is set at the mindstates of three characters (plus the observer's own mindstate and, in the case of oral jokes, that of the comedian). Cartoons differ from verbal jokes in one important respect, however: despite the fact that they obviate the need to model the comedian's mindstate, the cartoon format makes it difficult for the cartoonist to increase the level of mentalizing complexity. A likely explanation is that representing several characters engaging with each other in a 2-dimensional space with no time dimension is challenging. This was certainly our experience in trying to draw static (single panel) cartoon versions of standard mentalizing tasks: cartoons with four or more characters simply didn't work, and we fell back on simpler fourth order versions (Barrett et al. 2002, Figure 11.1, p. 296). Even so, some verbal labelling was required in order to point the observer to the right interpretation – unlabeled drawings are open to too many alternative interpretations.

Of all the results, it is perhaps the gender differences that are the most surprising: we did not anticipate that these would be as large as they are. They seem to reflect differences in the way the two genders engage with the social world (the one more reflectively, the other in a more superficial humor-based way). The present results thus add to a growing body of work in social psychology that points to pervasive gender differences in the way the social world is organized

(for reviews, see Benenson 2014; Dunbar 2018). That said, both genders have a strong preference for cartoons about the dynamics of social relationships. This fascination with relationships is reflected in the fact that fiction (which is almost always about social relationships) and biographies account for the vast bulk of book sales. Beyond this point of convergence, however, the two genders differ. Figure 5 suggests that women give higher ratings to a wider range of topics than men. Men rate visual humor (slapstick or situational) more highly than women do, whereas women rate jokes about political issues and domestic dynamics (i.e., intimate relationships) more highly than men do.

This contrast may reflect differences in the way the two genders manage their relationships. Women's relationships are generally more complex than men's, both in terms of the criteria used in mate choice (Dunbar 2012; Waynforth and Dunbar 1995) and the way they process socially relevant information (Benenson 2014; Pearce et al. 2019). In addition, the dynamics of relationships are very different: women's friendships are created and maintained principally through conversation (where the focus is often on intimacies and the discussion of social and emotional issues: Benenson 2014; Dahmardeh and Dunbar 2017; Dunbar et al. 1997) whereas men's are more activity-based, where conversation is used more to trigger laughter than discuss emotional issues (Dàvid-Barrett et al. 2015; Roberts and Dunbar 2015). These differences may reflect women's greater facility with language (Coates 2015; Newman et al. 2008), itself reflected in their more intense reading habits (Chiu and McBride-Chang 2006; Logan and Johnston 2010; Summers 2013; Uusen and Müürsepp 2012). We suggest that gender differences would be a particularly promising direction for future research.

Cartoons have become a regular feature of our cultural life, both as a vehicle to amuse and as a way of making political and social comments in the form of satire. As such, they represent an intersection of our psychological and social interests with both our ability to convey complex ideas in simple form and our capacity to entertain. In this respect, they provide us with novel insights into aspects of our psychology that are often difficult to study. What is clear, however, is that, at least in the context of pocket cartoons, the two-dimensional medium imposes severe constraints on the artist's capacity to create socially complex vignettes.

Acknowledgments: We thank The Cartoon Museum for making the project possible, and the Museum's staff for implementing it and creating the electronic database.
Ethics approval: Ethics approval was provided by the Cartoon Museum in accordance with the principles of the Helsinki Declaration. Participants were (unpaid) volunteers and completed a written consent form prior to undertaking the study.

References

Agüero Guerra, Marta. 2016. Beyond verbal incongruity: A genre-specific model for the interpretation of humour in political cartoons. In: Leonor Ruiz-Gurillo (ed.), Metapragmatics of humour: Current research trends, 57–77. Amsterdam: Benjamin.

Barrett, Louise, Robin I. M. Dunbar & John Lycett. 2002. *Human evolutionary psychology*. London: Macmillan/Palgrave and Princeton NJ: Princeton University Press.

Benenson, Joyce F. 2014. *Warriors and worriers: The survival of the sexes*. Oxford: Oxford University Press.

Canestrari, Carla & Ivana Bianchi. 2012. Perception of contrariety in jokes. *Discourse Processes* 49. 539–564.

Carbajal-Carrera, Beatriz & Olga Sanchez-Castro. 2020. The role of secondary incongruities in cartoon appreciation. *European Journal of Humour Research* 8. 25–48.

Carney, James, Rafael Wlodarski & Robin I. M. Dunbar. 2014. Inference or enaction? The influence of genre on the narrative processing of other minds. *PLoS One* 9. e114172.

Chiu, Ming M. & Catherine McBride-Chang. 2006. Gender, context, and reading: A comparison of students in 43 countries. *Scientific Studies of Reading* 10. 331–362.

Coates, Jennifer. 2015. *Women, men and language: A sociolinguistic account of gender differences in language*. London: Routledge.

Dahmardeh, Mahdi & Robin I. M Dunbar. 2017. What shall we talk about in Farsi? Content of everyday conversations in Iran. *Human Nature* 28. 423–433.

Dàvid-Barrett, Tamas, Anna Rotkirch, James Carney, Isabel Behncke Izquierdo, Jaimie Krems, Dylan Townley, Elinor McDaniell, Anna Byrne-Smith & Robin I. M. Dunbar. 2015. Women favour dyadic relationships, but men prefer clubs. *PLoS One* 10. e0118329.

Dennett, Daniel C. 1988. Précis of the intentional stance. *Behavioral and Brain Sciences* 11. 495–505.

Dunbar, Robin I. M. 2005. Why are good writers so rare? An evolutionary perspective on literature. *Journal of Cultural and Evolutionary Psychology* 3. 7–22.

Dunbar, Robin I. M. 2012. *The science of love and betrayal*. London: Faber & Faber.

Dunbar, Robin I. M. 2017. Cognitive and network constraints in real life and literature. In Ralph Kenna, Máirín MacCarron & Padraic MacCarron (eds.), *Maths meets myths: Quantitative approaches to ancient narratives*, 7–20. Berlin: Springer.

Dunbar, Robin I. M. 2018. The anatomy of friendship. *Trends in Cognitive Sciences* 22. 32–51.

Dunbar, Robin I. M. & Anna Machin. 2014. Sex differences in relationship conflict and reconciliation. *Journal of Evolutionary Psychology* 12. 109–133.

Dunbar, Robin I. M., Neil Duncan & Anna Marriot. 1997. Human conversational behaviour. *Human Nature* 8. 231–246.

Dunbar, Robin I. M., Jacques Launay & Oliver Curry. 2016. The complexity of jokes is limited by cognitive constraints on mentalizing. *Human Nature* 27. 130–140.

Fortabosco, Giovannantonio. 2008. Is the concept of incongruity still a useful construct for the advancement of humor research? *Lodz Papers in Pragmatics* 4. 45–62.

Griffin, Richard, Ori Friedman, Jon Ween, Ellen Winner, Francesca Happé & Hiram Brownell. 2006. Theory of mind and the right cerebral hemisphere: Refining the scope of impairment. *Laterality* 11. 195–225.

Howe, Norman E. 2002. The origin of humor. *Medical Hypotheses* 59. 252–254.

Hurley, Mathew M., Daniel C. Dennett & Reginald B. Adams. 2011. *Inside jokes*. Cambridge, MA: MIT Press.

Kinderman, Peter, Robin I. M. Dunbar & Richard P. Bentall. 1998. Theory-of-mind deficits and causal attributions. *British Journal of Psychology* 89. 191–204.

Krems, Jaimie & Robin I. M. Dunbar. 2013. Clique size and network characteristics in hyperlink cinema: Constraints of evolved psychology. *Human Nature* 24. 414–429.

Krems, Jaimie, Steven Neuberg & Robin I. M. Dunbar. 2016. Something to talk about: Are conversation sizes constrained by mental modeling abilities? *Evolution and Human Behavior* 37. 423–428.
Lewis, Penelope A., Roozbeh Rezaie, Rachel Browne, Neil Roberts & Robin I. M. Dunbar. 2011. Ventromedial prefrontal volume predicts understanding of others and social network size. *NeuroImage* 57. 1624–1629.
Logan, Sarah & Rhona Johnston. 2010. Investigating gender differences in reading. *Educational Review* 62. 175–187.
Machin, Anna & Robin I. M. Dunbar. 2013. Sex and gender in romantic partnerships and best friendships. *Journal of Relationships Research* 4. e8.
Newman, Matthew L., Carla J. Groom, Lori D. Handelman & James W. Pennebaker. 2008. Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes* 45. 211–236.
Oesch, Nathan & Robin I. M. Dunbar. 2017. The emergence of recursion in human language: Mentalizing predicts recursive syntax task performance. *Journal of Neurolinguistics* 43. 95–106.
Pearce, Eluned, Rafeal Wlodarski, Machin Anna & Robin I. M. Dunbar. 2019. Exploring the links between dispositions, romantic relationships, support networks and community inclusion in men and women. *PLoS One* 14. e0216210.
Pedrazzini, Anna & Nora Scheuer. 2018. Distinguishing cartoon subgenres based on a multicultural contemporary corpus. *European Journal of Humour Research* 6. 100–123.
Pinar, Maria J. 2020. Humour and intertextuality in Steve Bell's political cartoons. *European Journal of Humour Research* 8. 16–39.
Powell, Joanne L., Penelope A. Lewis, Robin I. M. Dunbar, Marta García-Fiñana & Neil Roberts. 2010. Orbital prefrontal cortex volume correlates with social cognitive competence. *Neuropsychologia* 48. 3554–3562.
Ramachandran, Vilayanur S. 1998. The neurology and evolution of humor, laughter, and smiling: The false alarm theory. *Medical Hypotheses* 51. 351–354.
Roberts, Samuel B. G. & Robin I. M. Dunbar. 2015. Managing relationship decay: Network, gender, and contextual effects. *Human Nature* 26. 426–450.
Shammi, Prathiba & Donald T. Stuss. 1999. Humour appreciation: A role of the right frontal lobe. *Brain* 122. 657–666.
Stiller, James & Robin I. M. Dunbar. 2007. Perspective-taking and memory capacity predict social network size. *Social Networks* 29. 93–104.
Stiller, James, Daniel Nettle & Robin I. M. Dunbar. 2003. The small world of Shakespeare's plays. *Human Nature* 14. 397–408.
Suls, Jerry M. 1972. A two-stage model for the appreciation of jokes and cartoons. In Jeffrey H. Goldstein & Paul E. McGhee (eds.), *The psychology of humor*, 81–100. New York, NY: Academic Press.
Summers, Kate. 2013. Adult reading habits and preferences in relation to gender differences. *Reference and User Services Quarterly* 52. 243–249.
Uusen, Anne & Mare Müürsepp. 2012. Gender differences in reading habits among boys and girls of basic school in Estonia. *Procedia – Social and Behavioral Sciences* 69. 1795–1804.
Waynforth, David & Robin I. M. Dunbar. 1995. Conditional mate choice strategies in humans: Evidence from 'Lonely Hearts' advertisements. *Behaviour* 132. 755–779.
Whitworth, James. 2020. Visual humour and the pocket cartoon: Osbert Lancaster and a paradigm shift in the British Press in the interwar years. *Journal of European Periodical Studies* 5. 71–83.
Zunshine, Lisa. 2022. *The secret Life of literature*. Cambridge MA: MIT Press.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/humor-2023-0111).