Walusinski O (ed): The Mystery of Yawning in Physiology and Disease.
Front Neurol Neurosci. Basel, Karger, 2010, vol 28, pp 120–127

Methodological Problems in the Study of Contagious Yawning

Matthew W. Campbell · Frans B.M. de Waal

Living Links Center, Yerkes National Primate Research Center, Emory University, Atlanta, Ga., USA

Abstract

The recent interest in contagious yawning has raised several challenges as the varied methods of testing have left some unresolved issues. We do not know how differences in key variables affect the observed rates of yawning, and we highlight these as being in need of direct testing. Different researchers analyze their results differently, and we make some recommendations for more rigorous, thorough and informative analyses. Ultimately, problems arise when authors compare studies that used different methods and different analyses without acknowledging how these differences may have affected the results. In these cases, authors make inappropriate comparisons, which lead to conclusions that add confusion to the literature. Our goal in raising awareness of these issues is to generate new experiments and improve the discussion of existing research. With its link to empathy, a more standardized study of contagious yawning may be a useful tool for a variety of disciplines.

Recently there has been renewed interest in the study of contagious yawning. More research on contagious yawning has been published in the span of 2000–2009 than in the previous 2 decades (and even longer) combined. The cause of this renewed interest is the theoretical link between contagious yawning and empathy [1, 2], which has been supported empirically [3, 4]. Since the early experiments on contagious yawning [5–8], the study of this phenomenon has employed developmental [9], comparative [10–14], neurological [15–18] and mental health [19–21] approaches in a pursuit that is beginning to formulate a varied body of research. However, the now numerous studies employed different methods of experimentation and analysis with different strengths and weaknesses, which complicates efforts to compare results across studies. Focusing on experimental manipulations of contagious yawning, which have been far more numerous than naturalistic observations [4], our goal of this chapter is to discuss some of these problems, propose some solutions, and highlight some unanswered questions to facilitate future experiments and discourse.

Problems in Experimental Design

When trying to compare results across studies, the first challenge is the different ways people test contagious yawning. The most fundamental aspects of design, what is presented and how, vary so much that no two studies by different sets of authors have employed the same methods (table 1).

Three variables in particular differ more frequently than others:

(1) The duration of the yawn shown ranges from 3 to 9 s (table 1), with some studies not reporting this detail. This magnitude of difference is small, and there is ambiguity in the length of the actual yawn within the stated length of a clip (i.e. how much neutral-expression lead-time and lag-time there was). As all researchers were attempting to show more or less 'typical' yawns, relatively small differences in duration may not influence the results greatly. Nonetheless, no one has expressly studied whether the duration of a stimulus yawn influences rates of contagion.

(2) Much more challenging is that the number of yawns displayed to the subjects ranges from 1 to 140. Of the 18 studies that presented yawns, there are 13 different total number of yawns shown. This is an enormous magnitude of variation, with the added wrinkle that in some experiments all of the clips were shown in 1 session, whereas in others they were shown in multiple sessions. We do not know if the number of yawns viewed by the subjects influences the rate of contagion. Does watching more yawns induce more individuals to yawn? Does watching more yawns induce more yawns from the same individuals? The obvious co-variable that may also contribute to yawning rates is the amount of time that the subjects are observed. The longer subjects are observed, the more likely one will observe yawning. These variables, time and the number of yawns presented, can be manipulated independently, and this awaits testing. In the meantime, we need to be cautious when comparing relative rates of contagious yawning between studies with even moderately different durations of exposure. Many studies make these comparisons without acknowledging that the methods differ in potentially important ways.

(3) Another critical aspect of design is the control. There is no consensus about what makes for the ideal control. Studies have used smiles, coughs, laughs, mouths opening and closing (also called gaping), still faces and species-specific expressions (for 3 comparative studies). The merits of different controls may be argued. Some may support gaping as it mimics much of the motor pattern of a yawn, yet is a meaningless expression (note: Nahab et al. [17] had a reason to use this expression specific to using fMRI). Conversely, it can be argued that gaping does not include, and therefore control for, motor activity in the eye region the way smiles, laughs and species-specific expressions do, which Provine [7] demonstrated was important for the perception of yawns. Since there has been no comparison of yawn rates in response to these different expressions, we have no data on whether any are better or worse controls than the others. However, significant differences in yawning rates when comparing yawns and controls have been detected using all of these expressions except coughs

Table 1. Studies on yawning

First author and reference	Subjects	Control expression	Duration of yawn clip	Number of exposures	Apparatus size
Anderson [9]	87 children (3–11 years old)	smile	unreported	17	unreported
Anderson [10]	6 adult chimpanzees	open mouth movements	unreported	40	35-cm monitor
Arnott [18]	10 adult humans	breath, scrambled yawn (auditory)	mean ± SD: 6 ± 1.16 s	40	NA, auditory only
Baenninger [8]	40 adult humans	none	unreported for video and live presentation	1? unclear	live presentation, and unreported for video
Campbell [14]	24 adult chimpanzees	play face, hoot, tooth clack	9 s	90	48-cm monitor
Giganti [21]	children: 7 high-functioning ASD, 10 low-functioning ASD, 10 TD	smile	5 s	20	unreported
Harr [13]	15 juvenile and adult dogs	gape	3–5s	10	unreported
Joly-Mascheroni [12]	29 adult dogs	gape	live presentation	10–19	live presentation
Moore [5]	36 adult humans	none	live presentation	5	unreported
Nahab [17]	18 adult humans	gape, cough, no movement	4 s	20	unreported
Paukner [11]	22 stumptail macaques (18 adult, 3 sub-adult, 1 infant)	open mouth movements	4.5 s	140	43-cm monitor
Platek [3]	65 adult humans	laughing	7 s	8	12.1 × 10 cm window on a monitor
Platek [15]	10 adult humans	laughing	7 s	15	fMRI goggles
Provine [6]	66 adult humans	smile	5 s	30	53-cm monitor
Provine [7]	360 adult humans	smile	5 s	30	43-cm monitor
Schurmann [16]	30 adult humans	mouth and tongue movement	24–27 s for 2 clips	6? unclear	unreported
Senju [19]	children: 24 ASD, 25 TD	gape	7 s	6	30.5-cm monitor
Senju [20]	children: 31 ASD, 31 TD	gape	7 s	6	30.5-cm monitor

The 4 studies with numerous NA entries are the 4 fMRI studies that could not observe actual yawns from the subjects due to the restrictions on head movement necessary for brain imaging. TD = Typical developers.

Post-exposure observation (for build-up effect)	Yawn recording method	Asked about feeling like yawning	Yawn vs. control comparison, Population-level	Yawn vs. control comparison, Individual-level
5 min	experimenter in room	Yes	Not tested	Not tested
3 min	videotaped	Nonhuman subjects	*t*-test, NS	binomial, 2 of 6 subjects significant
NA	NA	Yes	NA	NA
none	experimenter in room	Yes	Not tested	Not tested
5 min	videotaped	Nonhuman subjects	*t*-test, $p = 0.003$	binomial, 6 of 23 subjects significant
none	videotaped	No	Wilcoxon, $p = 0.01$	Not tested
3 min	experimenter in room	Nonhuman subjects	*t*-test, NS	binomial, 1 of 15 subjects significant
5 min	experimenter in room and videotaped	Nonhuman subjects	McNemar, $p < 0.001$	Not tested
none	experimenter in room and recorded by subject	Yes	Not tested	Not tested
NA	NA	Yes	NA	NA
3 min	experimenter out of room	Nonhuman subjects	Wilcoxon, $p = 0.02$	Not tested (but some appear significant on the graph)
none	experimenter out of room	No	Not tested (a significant Wilcoxon can be inferred from the results)	Not tested
NA	NA	No	NA	NA
none	recorded by subject	No	Chi-square, $p < 0.01$	Not tested
none	recorded by subject	No	Chi-square, $p < 0.02$	Not tested
NA	NA	Yes	NA	NA
1 min	videotaped	No	Wilcoxon, $p = 0.038$ (TD condition only)	Not tested
1 min	videotaped	No	Wilcoxon, NS	Not tested

and still faces, which have only been used once in an fMRI study [17]. Therefore, the specific control expression selected may not be important, as multiple expressions seem to turn up baseline levels of yawning. It would be interesting to know if yawning rates in response to any of these expressions differ from watching a neutral face, or even a blank screen. When selecting a control, it stands to reason that one should not use an extremely arousing expression, such as fear, but we reiterate that there has been no research comparing yawning rates to different controls.

There are many more methodological details that differ between studies, as can be seen in table 1, and the same comment applies: these details have not been studied for their effect on contagious yawning. We chose to highlight 3 aspects of methodology that we feel have the biggest potential to impact rates of contagion. The duration of the yawn clip may not be important, ultimately, and many different control expressions may do an equally good job of eliciting a baseline rate of yawning. However, the number of the yawn clips shown, with the co-variable of the amount of time the subjects are observed, has the potential to drastically impact the results and therefore interpretation. Research in this area is gravely needed, both for future experiments, and to more accurately compare existing studies that vary greatly in this domain.

Problems in Analysis

What qualifies as yawn contagion? This simple question has no clear answer in the literature. Whereas all authors agree on what yawn contagion is: a yawn stimulated by another yawn, there is no agreement on how to measure it. The most common analytical methods assess population-level differences between yawn and control conditions using either parametric (e.g. Student's t test) or non-parametric (e.g. Wilcoxon signed-rank test, χ^2, McNemar's test) statistics. Buried within the parametric or non-parametric distinction is a subtle difference in what these tests measure. The t test and Wilcoxon examine whether there were more *yawns* in the yawn condition than the control; χ^2 and the McNemar variant examine whether more *subjects* yawned in the yawn than the control condition, regardless of the magnitude of response. The χ^2 and McNemar may be of more limited use as they require a lot of 0's in the control condition. Whether one type of test is preferable to the other is not clear, but experimenters and readers need to be aware of the differences.

While relatively rare, a few studies did not do a statistical test between their yawn and control responses, which is a significant oversight. Population-level tests tell us whether one group as a whole is different from another, but they do not inform us about any one individual. Less commonly used, a few studies (curiously, all studies of non-human animals) have also employed individual-level statistics in the form of binomials. Binomials compare one individual's rate of response between two conditions. This test can identify individuals who show a strong difference in yawning regardless of the population-level statistics. In short, they can identify high-performers. The limitation to binomials is that they are not sensitive to small differences that may be meaningful and consistent within a population, hence the need for population-level statistics. These two levels of analysis should be seen as complementary: population-level tests tell us whether one group as a group differs from another, whereas individual-level tests identify particularly strong performers within a population. One is not better than the other; rather, they test different questions. All experimental manipulations of contagious yawning should compare the response to the yawn stimulus to the control stimulus. This is mandatory. Our suggestion is that a thorough analysis should contain statistics at both population and individual levels.

One particular result that is frequently reported but difficult to interpret is the percent of subjects showing contagious yawning. The difficulty arises because first one must define operationally what qualifies a subject as showing contagious yawning. An operational definition of contagious yawning has not been made explicit in any study reporting this result. There are at least 3 different ways one could operationally define whether a subject shows contagious yawning: (1) if subject A yawns at all during the yawn condition, regardless of how many times or if the subjects yawns at all, or even more frequently, during the control; (2) if subject A yawns more in the yawn condition than the control; (3) if subject A yawns significantly more during the yawn than the control on a binomial test. Method 1 is the most liberal (and most commonly used), and method 3 is the most conservative. Using our own data on chimpanzees as an example [14], these three different methods yield percentages of contagiously yawning subjects of 65, 56.5 and 26%, respectively. Thus, when authors compare percentages calculated with different methods [calculation method not reported in: 10, 12, 13, 21] they end up drawing conclusions from heterogeneous comparisons.

Secondly, how many individuals are observed to yawn during the yawn condition will depend to some extent upon the duration of the experiment. The longer a subject is observed, the more likely one will see a yawn. This is not a problem when comparing results within a study since the yawn and control conditions will be the same, but when comparing results between studies, there are serious problems if one study watched subjects for 5 min, and another for 20 min. Comparing yawns per minute, as opposed to the absolute number of yawning individuals, is one way to control for this.

The last problem we raise with this comparison is that a population that naturally yawns at high rates will look like possessing more contagion when compared to a population that naturally yawns very little, regardless of how these populations perform between yawn and control conditions. One can imagine a population that yawns a lot spontaneously, with no significant difference between the yawn and the control condition, and another population that yawns very little, but does show a significant difference. Would it be accurate to conclude that the first population shows more contagious yawning than the second? It is the difference between the yawn and control conditions that indicates susceptibility to contagion. A better comparison would be to look at the magnitudes of the difference in response between yawn and control conditions. Therefore, we would look at which populations show strongly significant differences, moderately significant differences, or no differences at all. This allows each population to control for itself and its own natural rates of yawning, as well as controlling for different methodologies, as discussed above.

Conclusions

Our analysis yields the following broad recommendations:

- new controlled experiments on variables of methodology, specifically how the number of yawns shown, the time duration of the experiment, and different control expressions affect rates of yawning;
- demonstrating yawn contagion experimentally requires a statistically significant difference between the response to a yawn stimulus and a control;
- calculations of percent of subjects showing yawn contagion are not informative;
- when making comparisons between different studies, authors must acknowledge how different methods and analyses may influence the results.

The last 10 years has seen a resurgence in interest in contagious yawning. We are very excited about the varied directions research has taken, and we hope these trends continue. Studying key variables of presenting yawns will be beneficial for two reasons: (1) it will facilitate comparisons across studies, including clarifying which variables in past studies may be important; (2) it will bring us closer to a standardized methodology, which may be useful as a diagnostic test for some mental health conditions [3, 19–21]. Answering these basic questions and limiting confusion in the literature will facilitate using contagious yawning as a serious tool to better understand how empathy functions from developmental, comparative, neurological and mental health perspectives.

References

1 Lehmann HE: Yawning: a homeostatic reflex and its psychological significance. Bull Menninger Clin 1979;4:123–136.

2 Preston SD, de Waal FBM: Empathy: its ultimate and proximate bases. Behav Brain Sci 2002;25:1–72.

3 Platek SM, Critton SR, Myers TE, Gallup GG: Contagious yawning: the role of self-awareness and mental state attribution. Brain Res Cogn Brain Res 2003;17:223–227.

4 Palagi E, Leone A, Mancini G, Ferrari PF: Contagious yawning in gelada baboons as a possible expression of empathy. Proc Natl Acad Sci USA 2009;108:19262–19267.

5 Moore JE: Some psychological aspects of yawning. J Gen Psychol 1942;27:289–294.

6 Provine RR: Yawning as a stereotyped action pattern and releasing stimulus. Ethology 1986;72:109–122.

7 Provine RR: Faces as releasers of contagious yawning: an approach to face detection using normal humans subjects. Bull Psychon Soc 1989;27:211–214.

8 Baenninger R: Some comparative aspects of yawning in *Betta splendens, Homo sapiens, Panthera leo*, and *Papio sphinx*. J Comp Psychol 1987;101:349–354.

9 Anderson JR, Meno P: Psychological influences on yawning in children. Curr Psychol Lett 2003;11:1–7.

10 Anderson JR, Myowa-Yamakoshi M, Matsuzawa T: Contagious yawning in chimpanzees. Proc R Soc Lond B Biol Sci 2004;271(suppl):S468–S470.

11 Paukner A, Anderson JR: Video-induced yawning in stumptail macaques (*Macaca arctoides*). Biol Lett 2006;2:36–38.

12 Joly-Mascheroni RM, Senju A, Shepherd AJ: Dogs catch human yawns. Biol Lett 2008;4:446–448.

13 Harr AL, Gilbert VR, Philips KA: Do dogs (*Canis familiaris*) show contagious yawning? Anim Cogn 2009;12:833–837.

14 Campbell MW, Carter JD, Proctor D, Eisenberg ML, de Waal FBM: Computer animations stimulate contagious yawning in chimpanzees. Proc R Soc Lond B Biol Sci 2009;276:4255–4259.

15 Platek SM, Mohamed FB, Gallup JGG: Contagious yawning and the brain. Brain Res Cogn Brain Res 2005;23:448–452.

16 Schurmann M Hesse MD, Stephan KE, Saarela M, Zilles K, Hari R, Fink GR: Yearning to yawn: the neural basis of contagious yawning. Neuroimage 2005;24:1260–1264.

17 Nahab FB, Hattori N, Saad ZS, Hallett M: Contagious yawning and the frontal lobe: an fMRI study. Hum Brain Mapp 2009;30:1744–1751.

18 Arnott SR, Singhal A, Goodale MA: An investigation of auditory contagious yawning. Cogn Affect Behav Neurosci 2009;9:335–342.

19 Senju A, Maeda M, Kikuchi Y, Hasegawa T, Tojo Y, Osanai H: Absence of contagious yawning in children with autism spectrum disorder. Biol Lett 2007; 3:706–708.

20 Senju A, Kikuchi Y, Akechi H, Hasegawa T, Tojo Y, Osanai H: Brief report: does eye contact induce contagious yawning in children with autism spectrum disorder? J Autism Dev Disord 2009;39:1598–1602.

21 Giganti F, Esposito Ziello M: Contagious and spontaneous yawning in autistic and typically developing children. Curr Psychol Lett 2009;25:1–11.

Matthew W. Campbell
Living Links Center, Yerkes National Primate Research Center, Emory University
Atlanta, GA 30072 (USA)
Tel. +1 404 727 9071, Fax +1 404 727 9069
E-Mail matthew.campbell@emory.edu