

# ComFeel: Productivity is a Matter of the Senses Too

MARIOS CONSTANTINIDES, Nokia Bell Labs, UK

SANJA ŠĆEPANOVIĆ, Nokia Bell Labs, UK

DANIELE QUERCIA, Nokia Bell Labs, UK

HONGWEI LI, Nokia Bell Labs, UK

UGO SASSI, Nokia Bell Labs, UK

MICHAEL EGGLESTON, Nokia Bell Labs, US

Indoor environmental quality has been found to impact employees' productivity in the long run, yet it is unclear its meeting-level impact in the short term. We studied the relationship between sensorial pleasantness of a meeting's room and the meeting's productivity. By administering a 28-item questionnaire to 363 online participants, we indeed found that three factors captured 62% of people's experience of meetings: (a) *productivity*; (b) *psychological safety*; and (c) *room pleasantness*. To measure room pleasantness, we developed and deployed ComFeel, an indoor environmental sensing infrastructure, which captures light, temperature, and gas resistance readings through miniaturized and unobtrusive devices we built and named 'Geckos'. Across 29 real-world meetings, using *ComFeel*, we collected 1373 minutes of readings. For each of these meetings, we also collected whether each participant felt the meeting to have been productive, the setting to be psychologically safe, and the meeting room to be pleasant. As one expects, we found that, on average, the probability of a meeting being productive increased by 35% for each standard deviation increase in the psychological safety participants experienced. Importantly, that probability increased by as much as 25% for each increase in room pleasantness, confirming the significant short-term impact of the indoor environment on meetings' productivity.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile devices**; **Computer supported cooperative work**; **Empirical studies in ubiquitous and mobile computing**; *User studies*; **Empirical studies in ubiquitous and mobile computing**; *User studies*.

Additional Key Words and Phrases: productivity, meetings, environmental conditions, perceptions

## ACM Reference Format:

Marios Constantinides, Sanja Šćepanović, Daniele Quercia, Hongwei Li, Ugo Sassi, and Michael Eggleston. 2020. ComFeel: Productivity is a Matter of the Senses Too. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 123 (December 2020), 21 pages. <https://doi.org/10.1145/3432234>

## 1 INTRODUCTION

A prolonged exposure to poor indoor environmental conditions has been found to impact employees' cognitive functions [6] and decision-making [45], have a direct effect on workers' performance and productivity [23, 53], and eventually lead to increased levels of absenteeism [28, 71, 78]. Health-related problems such as asthma,

---

Authors' addresses: Marios Constantinides, Nokia Bell Labs, Cambridge, UK, [marios.constantinides@nokia-bell-labs.com](mailto:marios.constantinides@nokia-bell-labs.com); Sanja Šćepanović, Nokia Bell Labs, Cambridge, UK, [sanja.scepanovic@nokia-bell-labs.com](mailto:sanja.scepanovic@nokia-bell-labs.com); Daniele Quercia, Nokia Bell Labs, Cambridge, UK, [daniele.quercia@nokia-bell-labs.com](mailto:daniele.quercia@nokia-bell-labs.com); Hongwei Li, Nokia Bell Labs, Cambridge, UK, [hongwei.3.li@nokia-bell-labs.com](mailto:hongwei.3.li@nokia-bell-labs.com); Ugo Sassi, Nokia Bell Labs, Cambridge, UK, [ugo.sassi@nokia-bell-labs.com](mailto:ugo.sassi@nokia-bell-labs.com); Michael Eggleston, Nokia Bell Labs, Murray Hill, US, [michael.eggleston@nokia-bell-labs.com](mailto:michael.eggleston@nokia-bell-labs.com).

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2020 Association for Computing Machinery.

2474-9567/2020/12-ART123 \$15.00

<https://doi.org/10.1145/3432234>

headaches, throat irritation, respiration, and allergies [48, 51] have reportedly been associated with a phenomenon often referred to as the Sick Building Syndrome (SBS) [14, 60].

Traditionally, companies and organizations have resorted to sensors through which the environmental conditions could be sensed [3, 28, 78, 79], and even adapted accordingly [11, 29] to meet recommended standards<sup>1</sup>, thus increasing their employees' productivity and well-being. Sensing devices are often being deployed in spaces to sense indoor environmental quality (IEQ). To circumvent shortcomings of commercial devices such as costs or flexibility of platforms in terms of data acquisition, IEQ devices based on open-source electronic designs were made available [37]. For example, Ali et al. [4] proposed Elemental, an open-source device that combines IEQ readings with energy usage and HVAC operation sensors. A recent review of studies on indoor comfort [72] identified the application of machine learning to IEQ factors other than thermal comfort as one of the research gaps. Our work partly fills this gap. Additionally, as we shall see in §2, while previous studies highlighted the impact of the environmental conditions on people's productivity in the long run, it is unclear to which extent that impact translates in the short term at a meeting-level.

To explore the short term impact of indoor environmental quality, we looked at one of the most common daily activities at work: *meetings*. More specifically, we studied the relationship between the sensorial pleasantness of a meeting's room and the meeting's productivity. In so doing, we made four main contributions:

- We operationalized the concept of 'sensorial pleasantness' and 'productivity' by developing an online 28-item questionnaire, grounded on literature in the fields of Management and Organizational Science, and administering it to 363 participants (§3). Based on factor analysis, we indeed found that three factors captured 62% of people's experience of meetings: (a) *productivity* (i.e., whether a meeting has been productive); (b) *psychological safety* (i.e., whether participants felt listened to); and (c) *room pleasantness* (i.e., whether the room was felt to be pleasing to the senses).
- We developed an indoor environmental sensing infrastructure, called *ComFeel*, which captures light, temperature, and gas resistance readings through miniaturized and unobtrusive devices that we built and named 'Geckos' (§4).
- To measure room pleasantness, we deployed ComFeel across 29 real-world meetings held at a corporate setting, and collected 1373 minutes of environmental readings (§5). For each of these meeting, we also collected whether each participant felt the meeting to have been productive, the setting to be psychologically safe, and the meeting room to be pleasant.
- Using the collected data, we studied the interplay between the sensorial pleasantness and productivity (§6, §7). We found that, on average, the probability of a meeting being productive increased by 35% for each standard deviation increase in the psychological safety participants experienced. Importantly, the very same probability increased by as much as 25% for each standard deviation increase in room pleasantness. These results suggest that indoor environmental conditions do matter not only in the long run (as previous work found), but also in the short term so much so that significant differences in productivity were observed even within the constrained temporal space of individual meetings.

## 2 RELATED WORK

Next, we surveyed various lines of research that our work draws upon, and we grouped them into three main areas: *i*) indoor environmental sensing, *ii*) the effects of IEQ on productivity and performance, and *iii*) productivity and meetings.

<sup>1</sup><https://www.ashrae.org/technical-resources/bookstore/standards-62-1-62-2>

## 2.1 Indoor Environmental Sensing: Reality and Perception

A prolonged exposure to poor indoor environmental conditions has been found to impact people’s health [73], cognitive functions [6] and decision-making [45], and eventually lead to increased levels of absenteeism [28, 71, 78]. SBS is a term often used in Occupational Safety and Health<sup>2</sup> to describe health-related problems associated with poor “indoor air quality”. Symptoms such as watering eyes, headaches, dizziness, mental fogging, respiration, and allergies, just to name a few, make up the SBS symptoms list. Report after report<sup>3</sup> has shown that the amount of ‘good’ air and, more broadly, good indoor environmental conditions can reduce SBS symptoms, cut absenteeism, and even put on hold the transmission of infectious diseases [5]. The most prominent indoor environmental properties identified in previous works include the ventilation rate, the particle concentration, the amount of carbon dioxide and volatile organic compounds, the temperature, the floor-coverings and used supply air filters, the computational equipment, the lighting conditions, and the noise distraction and occupancy rate [6, 33, 34, 45, 67, 78]. Links between the actual indoor environmental conditions and people’s perceptions about it have also been reported in studies [18, 36]. For example, studies have shown differences between doctors’ and patients’ perceptions of the environmental conditions at the same hospital [19], and gender differences when reporting perceived environmental quality (female employees consistently report lower levels [40]).

## 2.2 Long and Short-term Effects of IEQ on Productivity

The effects of IEQ factors on productivity, performance, and absenteeism were studied at real-world workplaces [16, 28, 56, 71, 78], or in controlled laboratory environments [12, 31, 77]. Neidell [54] highlighted the importance of IEQ effects: *‘Although the damage per individual is small when compared to more extreme events, such as mortality and hospitalizations, the effects are more widespread and may thus represent an important cost to society.’* In addition to IEQ, another important concept is the perceived environmental quality (PEQ) [22]: in some cases, even when employees could not consciously perceive poor IEQ conditions their productivity and performance were negatively impacted [70, 78]. While the short-term effect of environmental quality on productivity in both experimental [78] and real workplace [56] settings was studied, it is yet unclear the short-term effect of indoor environments on a recurring activity within a constrained temporal space such as a *work meeting*.

## 2.3 Productivity and Meetings

In the workplace, performance is a multi-faceted construct mostly characterized by ‘how well workers and employees perform their tasks, the initiative they take and the resourcefulness they show in solving problems’ [64, 75]. Simply put, how productive they are. Traditionally, companies and organizations assess their workforce’s productivity using various methods including self-reports, and peer- and supervisors-ratings [13, 30]. However, these methods are not always administered frequently, thus making it challenging for organizations to obtain a snapshot of their workforce’s productivity in the short term, and are often debatable due to potential biases that self-assessments may entail [21, 66]. More recently, it has been shown that job performance and productivity might well be assessed using passive data collected from workers’ smartphones and wearable devices [50], and workers’ cognitive states can be successfully determined [68]. Generally speaking, measuring productivity, however, can be a complex, challenging task depending on the work activity one looks into.

One of the most common daily activities at work is meetings. It is one of the most common practices in which organizations mobilize their individual members to work together, share ideas, brainstorm solutions, and solve problems. While people allocate significant amount of their work time in attending meetings, and organizations devote notable large amounts of resources to facilitate them [45], people still tend to feel disconnected and perceive them as unproductive [63]. Reports suggest that one in every three employees, across cities and industry

<sup>2</sup>[www.cdc.gov/niosh](http://www.cdc.gov/niosh)

<sup>3</sup><https://buildingevidence.forhealth.org/>

Table 1. Demographics of the Online Survey Participants (*Note: Five participants did not disclose their age.*)

|           | Range                   | #answers |
|-----------|-------------------------|----------|
| Age       | 18-24 years old         | 3        |
|           | 25-34 years old         | 122      |
|           | 35-44 years old         | 84       |
|           | 45-54 years old         | 50       |
|           | 55-64 years old         | 19       |
|           | 65+ years old           | 5        |
| Gender    | Female                  | 154      |
|           | Male                    | 132      |
|           | Prefer not to say       | 2        |
| Education | High school graduate    | 45       |
|           | Bachelor's degree       | 151      |
|           | Master's degree         | 63       |
|           | Ph.D. or medical degree | 29       |

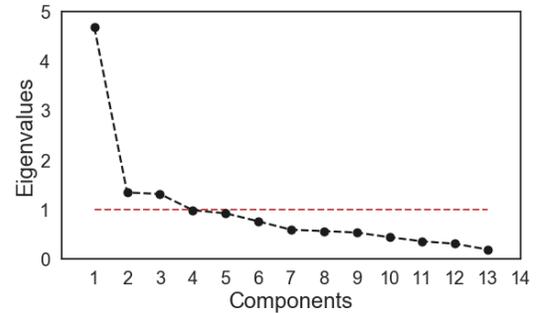


Fig. 1. Scree plot showing our PCA components and their eigenvalues in descending order.

sectors, considers meetings ineffective and unproductive [65]. However, ineffective meetings are bad for morale and productivity, negatively impact employees' health [43], and are directly linked to organizations' wasted time and resources [38, 61]. Therefore, it is paramount for organizations to understand how their employees experience meetings as an attempt to boost their productivity, and design tools to support it [7].

Traditionally, a meeting's performance and perceived experience has been evaluated through standardized questionnaires, or ad-hoc post meeting surveys. For example, the Meeting Effectiveness Index (MEI)<sup>4</sup> is a six-item questionnaire that focuses on a meeting's overall performance results. In the fields of Management and Organizational Science, meetings' productivity has been assessed by probing its agenda, structure, and purpose [17, 42, 69]. Inclusiveness [8, 9], dominance [63], peripheral activities [55], and psychological safety [10, 26] are aspects that have also been found to influence meetings' experience and productivity. The physical comfort of the workplace environment is an additional one that makes up the list [3]. For example, a recent survey<sup>5</sup> found that light and outdoor views were among the most popular perks employees craved for [47]. Furthermore, studies have shown that stuffy and stale air offices reduce employees' productivity [5, 6, 74].

### 3 ONLINE SURVEY

As stated in the previous section, existing questionnaires used to evaluate meetings focus primarily on performance, whereas the perceived experience of attendees has received little attention. We reviewed previous works from HCI communities, and articles in the fields of Management and Organizational Science relating to 'meetings' [5, 8–10, 26, 47, 55, 63, 74], and we identified themes concerning people such as their psychological experience (i.e., hierarchy and cultural balance, personality and gender diversity, and contribution and broad/participation), and the meetings' structure and content (i.e., agenda, environment, technology used, action points, time, and multitasking). Based on this review, we developed a questionnaire<sup>6</sup> consisting of 28 questions that cover the wide

<sup>4</sup>MEI: <http://www.meetingmetrics.com/EPI-Report-demo-data.pdf>

<sup>5</sup><https://futureworkplace.com/>

<sup>6</sup><http://bit.ly/psycho-meetings>

Table 2. Rotated Component Matrix for the 3 derived principal components of the psychological experience during meetings: *Psychological safety*; *Productivity*; and *Pleasantness*.

|   | <b>Psychological safety</b> | <b>Productivity</b> | <b>Pleasantness</b> |
|---|-----------------------------|---------------------|---------------------|
| 1. I received the appropriate attention when sharing my thoughts.                         | .835                        |                     |                     |
| 2. I felt comfortable sharing my thoughts and making contributions.                       | .814                        |                     |                     |
| 3. I felt that my peers were comfortable sharing their thoughts and making contributions. | .804                        |                     |                     |
| 4. I found the meeting to be useful for all attendees.                                    | .774                        |                     |                     |
| 5. I found the meeting to be useful for me.   | .747                        |                     |                     |
| 6. I felt overall satisfied during the meeting.   | .528                        |                     |                     |
| 7. Summary points were sent out before the meeting ended.                                 |                             | .773                |                     |
| 8. Summary points were sent out after the meeting has ended.                              |                             | .647                |                     |
| 9. The meeting had a clear purpose.   |                             | .611                |                     |
| 10. Amount of space in the meeting room.  |                             |                     | .865                |
| 11. Air quality in the meeting room.  |                             |                     | .790                |
| <i>component loadings below .5 are suppressed.</i>  |                             |                     |                     |

range of themes we identified, and an additional set of three questions relating to demographics (age, gender, and education levels).

### 3.1 Participants and Recruitment

We distributed the survey via social media (i.e., Facebook and Twitter on targeted lists of knowledge workers. We specified ‘job meeting’ and ‘project management’ as interests, and targeted a range of sectors from administrative to sale to legal services to business and finance), and as an Amazon Mechanical Turk (AMT) task. We ensured highly reputable AMT workers by targeting workers with 95% HIT approval rate, and at least 100 approved HITs. We received a total of 363 responses from both sources. We applied quality checks in the form of trap questions. For example, one question prompted participants to report whether they attended the meeting remotely or in-person, while a subsequent question asked them to state their perceived air quality of the room where the meeting took place. We rejected those who reported contradicting answers (e.g., attended the meeting remotely and answered the air quality question, while having the option to skip the air quality question), thus ensuring data quality. As the survey questions were not mandatory, we carried out further filtering of empty responses. We also eliminated those whose last meeting was held ‘more than a week ago’ to remove potential recall bias. The final dataset we analyzed consisted of 288 responses (Table 1).

### 3.2 Principal Component Analysis

We performed a Principal Component Analysis (PCA) on the survey responses to identify the main orthogonal dimensions that explain the responses. PCA “determine[s] the linear combinations of the measured variables that retain as much information from the original measured variables as possible.” [20]. The filtered dataset contained 288 responses and a ratio of 10.2 cases per variable, well above the recommended ratio of 1:10 for PCA analysis [58].

To ascertain the validity of our PCA results, we conducted three tests. First, we tested the correlation matrix between the 28 questions of the survey (excluding demographics). It showed significant correlations of above .3 with at least one other item for the 20 out of 28, suggesting reasonable factorability. Secondly, we used Bartlett's Test of Sphericity which was found to be significant ( $\chi^2(378) = 3092.43$ ,  $p = .000$ ). Thirdly, the Kaiser-Meyer-Olkin measure of sampling adequacy (MSA) was .716, which was above the recommended threshold of .7 [44]. Subsequently, fifteen items with MSA below .7 in the anti-image correlation matrix were excluded from the analysis [57]. These items included questions related to a respondent's current position (e.g., years of employment, employer); and the meeting's diversity (e.g., number of attendees, gender), interactions/conversation style, and time of day. This exclusion resulted in improved values of ( $\chi^2(78) = 1302.54$ ,  $p = .000$ ) in Bartlett's Test of Sphericity, and in an overall MSA score of .859. This made the PCA more robust with the remaining 13 items.

To determine the number of components that explain the most variance in our survey, we inspected the eigenvalues of the principal components (Figure 1). The eigenvalue of a principal component reflects the variance in all the variables explained by this component, and a higher value indicates higher variance in the variables loaded together in this component. The Kaiser's rule [35] suggests eigenvalues above 1.0 to be retained (resulting in 4 principal components), while the Cattell [15] proposed a visual inspection of the Scree Plot (Figure 1) to determine the 'elbow' point in which a leveling effect is observed. This resulted in 3 principal components that explained 56.12% of the total variance in the data. We considered the loadings of the 13 questionnaire items of the 3 principal components. The loadings describe the correlations between a single item and the principal component. We removed 2 items from further analysis due to their low component loading below .5. This resulted in 11 remaining items, which accounted for 62.31% of the total variance; a solution of 60% variance explainability is often considered as satisfactory in social sciences ([27] p.107). The additional two items left out concerned the respondent's meeting role (e.g., organizer, attendee), and the level of excitement during their last meeting, either of their loading components was below .5.

Table 2 shows the resulting components and their items. The first component is about *psychological safety*, which refers to personal and group's safety of contributing to the meeting, personal and groups' views on usefulness, attention received, and personal overall satisfaction. The second component is about *productivity*, which refers to whether the meeting had clear purpose and structure, and resulted in a list of actionable items (i.e., whether a meeting has been productive). The third component is about *pleasantness*, which refers to the meeting's room overall pleasantness (e.g., good indoor air quality and physical space utilization).

### 3.3 Three Factors of Meetings' Psychological Experience

To make use of these three PCA components in our user study (§5), we paraphrased them into three questions our participants are likely to understand without considerable effort:

*Q<sub>psychological</sub>*: Did you feel *listened to* during the meeting or *motivated to be involved* in it?

*Q<sub>productive</sub>*: Did the meeting have a *clear purpose and structure*, and did it result in a list of *actionable points*?

*Q<sub>pleasantness</sub>*: Did the *room feel pleasant* (in regards to air quality and crowdedness)?

## 4 COMFEEL FRAMEWORK

As our study aims to investigate the short-term effects of sensory layers on people's productivity in the context of meetings, we resorted to sensors to obtain objective measurements about the indoor environmental conditions. In what follows, we describe our data collection framework, namely ComFeel. It is a three-tier framework (Figure 2) composed by an application layer (a Gecko device and a Raspberry Pi), a web service layer, and a data access layer.

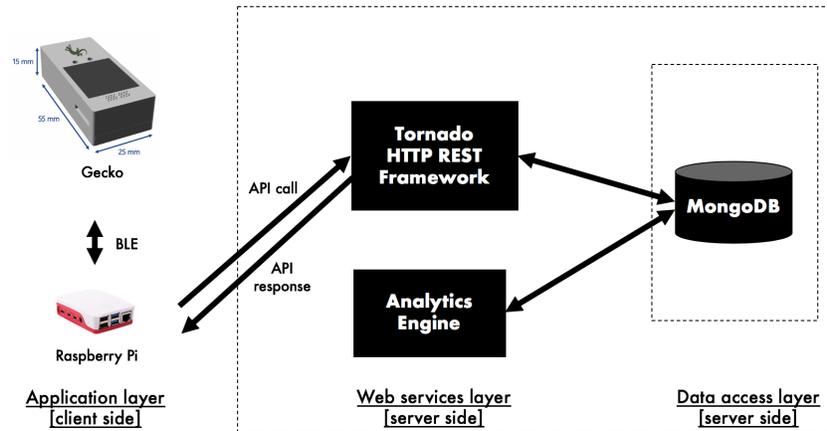


Fig. 2. ComFeel sensing infrastructure. Our Gecko device continuously monitors indoor environmental properties (i.e., light, temperature, gas resistance), which are transmitted to an intermediate agent implemented using a Raspberry Pi. Upon data acquisition, the Raspberry Pi transmits the data to our backend infrastructure for further processing.

## 4.1 Application Layer

**4.1.1 Gecko Device.** We assembled Gecko, a small (55x25x15), wearable environmental device, using off-the-shelf commercial sensors. By building our own device with low-manufacturing costs, we had full control of the firmware and we owned the raw data, which is often not the case in commercial devices. Additionally, its flexible design and plug&play functionality make it easy to deploy, and attractive for large-scale deployments. Its total costs is 88 USD, which includes a solar cell (2 USD), a battery (2 USD), a plastic case (8 USD), and a PCB board with all the electronic components (76 USD). This estimation is based on production of 50 units, whereas, in case of mass production (~1000 units), the costs could further be reduced to 50 USD. Assembly costs are not included in the aforementioned costs, as they are deemed negligible due to a Gecko's design being assemblable in less than 10 minutes.

We used Bosch's environmental sensor BME680<sup>7</sup> for monitoring indoor temperature levels (C°) and gas resistance (ohm), and a BH1745NUC digital 16-bit serial output type color sensor IC for obtaining light luminosity<sup>8</sup> (fc: foot-candle). We configured both sensors to sample every 60 seconds, and sensors readings were obtained through a micro-controller using an I2C interface on the predefined interval. Upon reading, Gecko packs up each sensor's raw readings and sends it over using a Bluetooth Low Energy (BLE) protocol. Given that the payload of each BLE packet cannot exceed 244 bytes, as specified in the Bluetooth standard 4.2 onwards, each sensor reading is wrapped up in a packet. Therefore, a packet contains up to 244 bytes of information, and has the following signature: [packet identifier, timestamp, sensor id, sensor data length], where *packet identifier* indicates that a packet carries sensor data payload, the *timestamp* contains the time in which the reading occurred, the *sensor id* is a unique identifier for each sensor, and the *sensor data length* is the actual reading in bytes. When the packet is ready, we used a Nordic UART service<sup>9</sup> to transmit it to an intermediate agent (i.e., Raspberry Pi), which we describe next. The choice of the Nordic UART service was primarily reinforced by the fact that it is an easy to implement and adapt it to our streaming requirements, which are independent of the type of reading and the

<sup>7</sup><https://www.bosch-sensortec.com/products/environmental-sensors/gas-sensors-bme680/>

<sup>8</sup>Color sensor: <https://www.mouser.co.uk/datasheet/2/348/bh1745nuc-e-519994.pdf>

<sup>9</sup>[https://developer.nordicsemi.com/nRF\\_Connect\\_SDK/doc/latest/nrf/include/bluetooth/services/nus.html](https://developer.nordicsemi.com/nRF_Connect_SDK/doc/latest/nrf/include/bluetooth/services/nus.html)

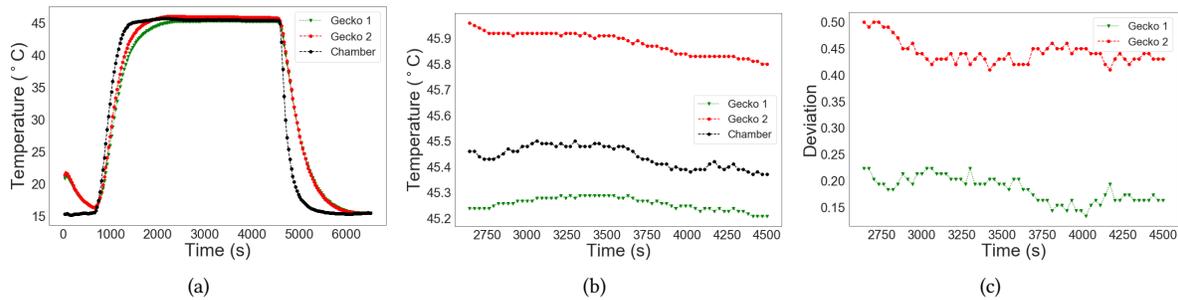


Fig. 3. (a) Geckos' vs. environmental chamber temperature measurements at 15°C and 45°C, (b) zoomed-in temperature measurements at 45°C for which the accuracy was estimated, and (c) deviations of the two Gecko devices against the environmental chamber's measurements.

sampling rate. Additionally, the BLE protocol itself ensures that packet loss or data errors during transmission would not happen. At the end of every link layer packet there is a 24-bit CRC; a mechanism that BLE uses, which contains a sequence number and the next expected one. Every packet is then acknowledged at the link layer, thus if no acknowledgment is received the packet will be retransmitted. To further validate any potential errors due to packet drops, in a subsequent section (§4.1.3), we describe the reliability of the Raspberry Pi gateway.

**4.1.2 Gecko Sensors Reliability.** To ascertain our device's sensors reliability, we carried out experiments that compared our devices' readings against the measurements operated in an environmental chamber; a typical instrument for reliability testing of different devices at different temperatures [41]. The accuracy of the temperature sensor inside the chamber is of 0.1°C.

According to the manufacturer<sup>10</sup>, the temperature sensor of our Gecko has an operating range from -40°C to 85°C with an accuracy of 1°C; the gas sensor has a response time of 1s, and an accuracy of 0.08%; and the light sensor is capable to measure a wide dynamic lux range up to 40K lux with an accuracy of 0.005<sup>11</sup>.

To test the temperature sensor readings, we placed two Gecko devices in the environmental chamber. We set the chamber's temperature to 15°C and waited until it stabilized in order to place the devices inside. After 10 minutes of stabilization, we increased the temperature to 45°C and, in turn, after roughly 50 minutes, we decreased it again to 15°C for another 10 minutes (Figure 3a). Using the values obtained at 15°C, we calibrated our sensors, and computed the accuracy of our temperature sensor from the measurements at 45°C (Figure 3b). We then estimated the accuracy as the mean value of all deviations between the Gecko measurements and the chamber temperature values at 45°C (Figure 3c). We observed worst-case precision of 0.5°C, which is in the range of the accuracy stated by the manufacturer<sup>12</sup>. Additionally, we observed that the two Gecko devices showed very similar response, which suggests low sensor-to-sensor variability.

Taking into account typical illumination levels of office spaces (i.e., in the range of 500 lux [52]), and given the capability of our light sensor to measure a wide dynamic lux range from 0.005 to 40K lux with an accuracy of 0.005 lux<sup>13</sup>, even an error of one order of magnitude would still provide accurate measurements. The same applies to the high accuracy of the gas resistance measurement as stated by the manufacturer (i.e., 0.08%<sup>14</sup>).

<sup>10</sup><https://www.bosch-sensortec.com/products/environmental-sensors/gas-sensors-bme680/>

<sup>11</sup><https://www.mouser.co.uk/datasheet/2/348/bh1745nuc-e-519994.pdf>

<sup>12</sup><https://www.bosch-sensortec.com/products/environmental-sensors/gas-sensors-bme680/>

<sup>13</sup><https://www.mouser.co.uk/datasheet/2/348/bh1745nuc-e-519994.pdf>

<sup>14</sup><https://www.bosch-sensortec.com/products/environmental-sensors/gas-sensors-bme680/>



Fig. 4. The two meetings rooms utilized in our user studies. (a) the *small* room has a maximum capacity of six people (16.25 m<sup>2</sup>), while (b) the *large* room has a maximum capacity of ten people (24.70 m<sup>2</sup>).

**4.1.3 Raspberry Pi.** As Gecko is designed to be a cheap and flexible indoor environmental device, it has no capabilities of streaming its data to a web service. To that end, we implemented an intermediate agent using Raspberry Pi which acts as a gateway between a Gecko device and our backend server. The Raspberry Pi collects data from one or more Gecko devices and forwards it to a web server. The gateway is implemented in such a way to actively listen for connections over BLE. Once a Gecko device establishes a BLE connection with the Raspberry Pi, the gateway decodes the bytes array sent from Gecko, formats it in as a JSON, and transmits it to our backend for further processing and analysis. To further validate whether any potential errors occurred at the Raspberry Pi side, we transmitted data to the gateway from two Gecko devices. (Figure 2). In total, both devices transmitted 1178 packets, and we observed 5 misses out of 587 packets sent from the first device, and 8 misses out of 591 packets sent from the second device. On average, this accounts for no more than 1.1% error rate at the gateway level.

## 4.2 Web Services Layer

Using the Python Tornado framework, we developed a RESTful API that runs continuously on a server. It exposes an endpoint that handles the communication with the Raspberry Pi module and the data access layer. The endpoint is responsible for receiving the data from the gateway and flushing it to the database.

## 4.3 Data Access Layer

The data access layer is implemented by a MongoDB instance, and the raw data is stored in a database collection with the following generic schema: [deviceID, sensor\_type, sensor\_value, ts], where *deviceID* is a unique identifier that distinguishes Gecko devices, *sensor\_type* is a key identifier that corresponds to the type of environmental data (i.e., temperature, gas resistance, or light luminosity), *sensor\_value* is the sensor reading, and *ts* is the time that the sensor reading was sampled.

## 5 USER STUDY

Having developed a sensing infrastructure to obtain objective measurements of indoor environmental quality, we conducted a longitudinal study, spanned over a period of one month, to understand the sensorial pleasantness of a meeting's room and the meeting's productivity.

Table 3. The four-item survey which our participants completed at the end of each meeting.

| Name                | Question  | Responses   |
|---------------------|---|---|
| $Q_{type}$          | What was the type of the meeting you just had?  | (a) status update<br>(b) information sharing /presentation<br>(c) decision making/problem solving |
| $Q_{productive}$    | Did the meeting have a clear purpose and structure, and did it result in a list of actionable points? | Likert Scale<br>(1: Strongly disagree; 7: Strongly agree)   |
| $Q_{psychological}$ | Did you feel <i>listened to</i> during the meeting, or <i>motivated to be involved in it</i> ?        | Likert Scale<br>(1: Strongly disagree; 7: Strongly agree)   |
| $Q_{pleasantness}$  | Did the room feel pleasant?<br>(in regard to air quality and crowdedness)                             | Likert Scale<br>(1: Strongly disagree; 7: Strongly agree)   |

### 5.1 Study Setup

The study was conducted at the facilities of a research laboratory in a corporate setting during February-March. Two meeting rooms with varying capacities and dynamics were utilized (Figure 4). We refer to them as the *small* room and the *large* room. The small room's is 16.25 m<sup>2</sup>, which is ideal for small-size meetings (six seats), while the large room's ones is 24.70 m<sup>2</sup> (ten seats), which is more preferable for hosting larger-size meetings. The choice of these two particular rooms sizes was reinforced by their popularity. Therefore, we did not set up any 'experimental' rooms, rather we deployed our framework in rooms people regularly use. This also allowed us to understand people's behavior and perception under real-world corporate settings as opposed to controlled lab settings. Additionally, these two rooms, when empty, are similar in terms of light (same view), temperature, and gas resistance. We also kept both rooms' blinds open during the trial, thus ensuring similar exposure to outside views in both rooms. We deployed two Gecko and Raspberry Pi devices, one in each room, which were continuously monitoring and storing the indoor environmental information on our backend server (Figure 2).

### 5.2 Materials and Apparatus

To capture our participants' perceptions during a meeting, we asked them the three questions we identified in the online survey (§3), at the end of their meeting. We included an additional question to report the type of the meeting. All questions along with their response options are listed in Table 3. Finally, participants noted in the survey the start and end time of their meeting, which we made use in our analysis (§6). Participation in the questionnaires was anonymous. We opted-in for completely anonymized responses to ensure privacy compliance according to the EU General Data Protection Regulation (GDPR). As we did not have any particular requirements for participation, we neither adopted a recruitment protocol nor specified any eligibility criteria; instead, we simply installed an information sheet in both rooms, detailing the study's protocol.

### 5.3 Protocol, Procedure, and Participants

The protocol stated the aims of the study, the data collection, and made explicit to participants that they agree that their responses will be analyzed for research purposes. In accordance to GDPR, no researcher involved in the study could have tracked their identity by any means, and all their anonymous responses were analyzed at an aggregated level. In addition to the data handling and data analysis, the protocol stated the experimental procedure. Participants were informed that, for a period of one month, two Gecko devices were installed in the two meeting rooms, and these devices were monitoring both rooms' environmental conditions. In total,

Table 4. Statistics of the twenty nine meetings held in the two meeting rooms.

|                          | small room | large room |
|--------------------------|------------|------------|
| #Meetings                | 18         | 11         |
| #Minutes of sensing data | 843        | 530        |
| #responses               | 220        | 144        |
| #participants responded  | 55         | 36         |

we received 364 self-reported answers to the four questions (Table 3) from 91 participants who attended those meetings (Table 4).

#### 5.4 Dataset

In total, we collected 1373 minutes of environmental data across 29 unique meetings that were held in the two meeting rooms (Table 4). As there is no unified classification of meeting types, we grouped our meetings into three common meeting types found in previous literature [25, 39, 46, 49]. These are: (a) *status update*, (b) *information sharing/presentation*, and (c) *decision making/problem solving*. In total, our participants attended 58.2% status update (*type 0*), 28.6% information sharing/presentation (*type 1*), and 13.2% decision making/problem solving (*type 2*) meetings (Figure 6a), with varying duration (Figure 6e). To assess the quality of the responses to our three questions, we looked at the inter-rater agreement (IRA) among participants of the same meeting. We computed an IRA statistic called  $r_{wg}$  score [32] and the standard deviation across participants' answers, which are commonly applied IRA measures [59] for Likert-scale questionnaires. The  $r_{wg}$  statistic (1) is a function of the observed variance in participants' answers ( $S_x^2$ ), and the variance in the ratings, if they were random, i.e., null distribution ( $\sigma_{eu}^2$ ):

$$r_{wg} = 1 - (S_x^2 / \sigma_{eu}^2), \quad (1)$$

where  $x$  is the set of the participants' answers,  $S_x^2 = var(x)$ , and  $\sigma_{eu}^2$  is set by convention to a constant value of  $(A^2 - 1)/12$ , and  $A$  is the number of Likert-scale options (in our case,  $A = 7$ ). The values of  $r_{wg}$  range from 0 to 1, with those closer to 1 representing a better agreement. The distributions of scores across meetings shown in Figure 5 reveal a relatively high level of agreement among our participants for all the three questions (average  $r_{wg}$  for  $Q_{psychological}$  is .78, for  $Q_{productive}$  is .87, and for  $Q_{pleasantness}$  is .75). Unsurprisingly, our participants agreed more on the meeting's productivity than on the other two more subjective aspects.

## 6 ANALYSIS

Having collected objective environmental measurements and employees' self-reported experiences of their meetings in terms of their perceived psychological safety, productivity, and room pleasantness, we set out to investigate the interplay of these factors with productivity. Based on prior work, we formulated three hypotheses, and defined a set of metrics to test them.

### 6.1 Hypotheses and Metrics

As one expects, the outcome of a meeting (i.e., productive or not) is dependent on its type. In the fields of the Management and Organizational Science, the most common types include *status update*, *information sharing/presentation*, and *decision making/problem solving* [25, 39, 46, 49]. It is also known that the length of a meeting might influence its productivity. Often, meetings that do not start or end on time tend to leave attendees' unsatisfied [17, 69]. Therefore, we formulated our first hypothesis as:

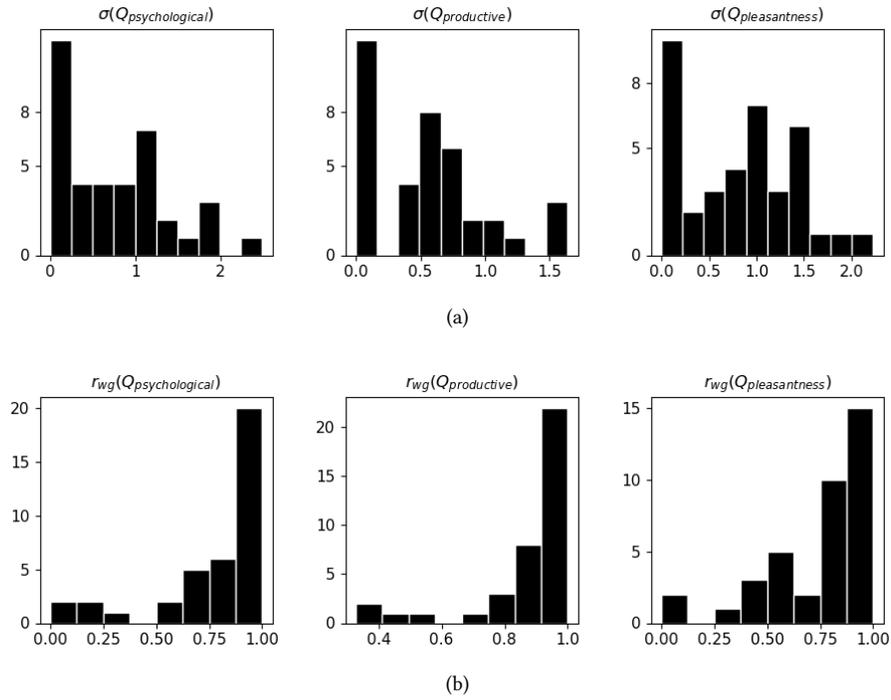


Fig. 5. Distributions of inter-rater agreement among meeting participants for the three questions in terms of: (a) standard deviation  $\sigma$  of their answer scores, and (b) the Likert-scale inter-rater agreement measure  $r_{wg}$  [32].

**H<sub>1</sub>**: Whether a meeting is productive depends on the meeting's *type* and *duration*.

To test it, we used a meeting's *type* by asking our participants the  $Q_{type}$  in Table 3, and *duration* computed from their timestamped responses for each meeting (Figure 6e). Since the *meeting type* question (Figure 6a) is a categorical variable, we converted it into three dummy variables to then make use of them in a logistic regression: *type 0*, *type 1*, and *type 2*.

We treated this problem as a classification task, and we used the binarized self-reported  $Q_{productive}$  as the dependent variable. To binarize it into a positive and negative class, we defined a set  $E$  as the set of all participants' responses to  $Q_{productive}$ , and we partitioned it into two groups such that  $E = E^> \cup E^<$ , where  $E^>$  is the set of participants who reported higher levels of meetings' productivity (above the median) defined as  $E^> = \{e \mid e > \bar{e}\}$ , where  $e$  is a participant's  $Q_{productive}$  response, and  $E^<$  is the set of participants who reported lower levels of meetings' productivity (below the median) defined as  $E^< = \{e \mid e < \bar{e}\}$ , where  $e$  is a participant's  $Q_{productive}$  response. Then, the  $E^>$  responses were assigned to the positive class, while the  $E^<$  ones to the negative (Figure 6c).

We built the  $M_1$  logistic regression model (2) to test  $H_1$  by including meetings' *duration* and the three dummy variables (i.e., *type 0*, *type 1*, and *type 2*), which reflect the *type* of meeting:

$$\Pr(Q_{\text{productive}} = 1 \mid \text{type 0, type 1, type 2, duration}) = \frac{1}{1 + \exp^{-(\alpha + \beta_1 \text{type0} + \beta_2 \text{type1} + \beta_3 \text{type2} + \beta_4 \text{duration})}}, \quad (2)$$

where Pr is the posterior probability of an instance belonging to the positive class of  $Q_{\text{productive}}$ , estimated through a linear combination of the input features (i.e., type 0, type 1, type 2, and duration) passed through a sigmoidal function, and  $\alpha, \beta_1, \beta_2, \beta_3,$  and  $\beta_4$  are the parameters to be learned by model  $M_1$ .

Furthermore, a productive meeting might also depend on the psychological safety experienced by participants (i.e., whether they felt listened to). As we previously illustrated (§3), psychological factors such as inclusiveness [8, 9], dominance [63], and comfort in sharing and contributing [10, 26] are linked with people’s overall experience of a meeting. To understand whether these factors also impact meetings’ productivity, we formulated our second hypothesis as:

**H<sub>2</sub>:** Whether a meeting is productive depends on the meeting’s *type, duration,* and on the *psychological safety* experienced by participants.

To investigate  $H_2$ , we used the z-score transformed ( $\frac{x-\mu_x}{\sigma_x}$ ) of  $Q_{\text{psychological}}$  reported by our participants (Figure 6b) along with meetings’ *duration,* and the three dummy variables, which reflect the meeting *type.*

We built the  $M_2$  logistic regression model (3) to investigate  $H_2$  by adding *psychological safety* to the  $M_1$  model. We used the binarized self-reported  $Q_{\text{productive}}$  as the dependent variable:

$$\Pr(Q_{\text{productive}} = 1 \mid \text{type 0, type 1, type 2, duration, } Q_{\text{psychological}}) = \frac{1}{1 + \exp^{-(\alpha + \beta_1 \text{type0} + \beta_2 \text{type1} + \beta_3 \text{type2} + \beta_4 \text{duration} + \beta_5 Q_{\text{psychological}})}}, \quad (3)$$

where Pr is the posterior probability of an instance belonging to the positive class of  $Q_{\text{productive}}$ , estimated through a linear combination of the input features (i.e., type 0, type 1, type 2, duration, and  $Q_{\text{psychological}}$ ) passed through a sigmoidal function, and  $\alpha, \beta_1, \beta_2, \beta_3, \beta_4,$  and  $\beta_5$  are the parameters to be learned by model  $M_2$ .

Finally, the role of indoor environmental conditions into workplace’s productivity has been the subject of several reports. For example, the lighting conditions were among the most popular perks<sup>15</sup> employees crave for at the workplace [47]. Additionally, studies have shown that stuffy and stale air offices reduce productivity [5, 6, 74]. Other studies have reported that the actual environmental conditions might not always match the subjective perception of the physical space though [18, 36]. Therefore, we formulated our third hypothesis as:

**H<sub>3</sub>:** Whether a meeting is productive depends on the meeting’s *environmental conditions,* both actual and perceived conditions.

To investigate  $H_3$ , we used the z-score transformed self-reported  $Q_{\text{psychological}}$  along with two metrics that capture the *sensed pleasantness* and the *not-sensed* one. The sensed pleasantness is a metric that captures the environmental conditions our ComFeel framework is able to capture. To built this metric, we resorted to the sensor readings obtained using our Gecko devices, and we did so in two steps. First, we used the indoor *temperature* level (C°), *gas resistance* (ohm), and *light luminosity* (fc: foot-candle) (Figure 6f-h). For the sake of brevity, we refer to them as *temp, gas,* and *light* respectively for the rest of our analysis. As the device was set to sample sensor readings every 60 seconds (§4.1.1), we aggregated the minutely obtained values across each meeting’s duration, took their median value, and z-score ( $\frac{x-\mu_x}{\sigma_x}$ ) transformed them to then make use of them in a logistic regression.

<sup>15</sup><https://futureworkplace.com/>

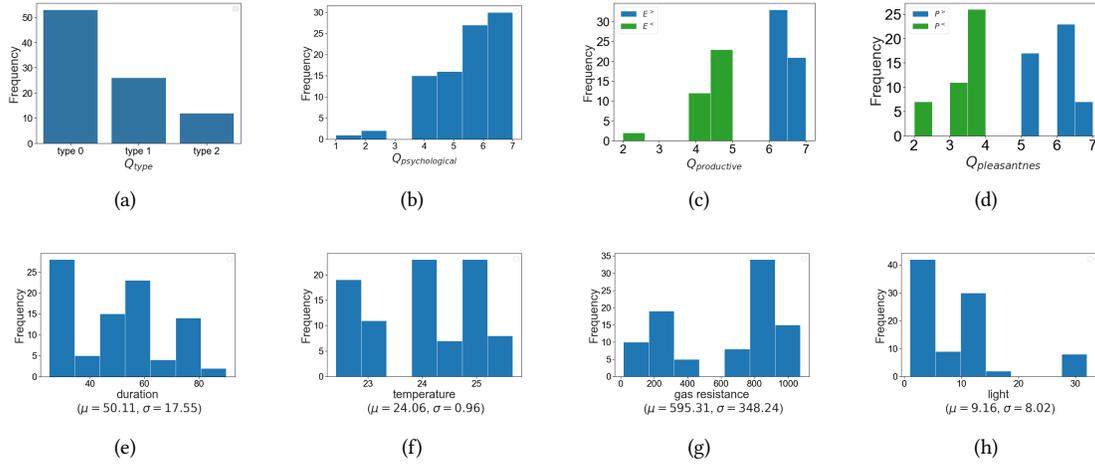


Fig. 6. Frequency distribution of variables used in the analysis. (a)  $Q_{type}$ ; (b)  $Q_{psychological}$ ; (c)  $Q_{productive}$ ; (d)  $Q_{pleasantness}$ ; (e) duration; (f) temperature; (g) gas resistance; (h) light.

We chose the median statistic as a more representative value due to its robustness against outliers, while the choice of using z-score transformations as opposed to raw values allowed us to incorporate the variance of the sensor-derived data. We then computed a logistic regression as:

$$\Pr(Q_{pleasantness} = 1 \mid temp, gas, light) = \frac{1}{1 + \exp^{-(\alpha + \beta_1 temp + \beta_2 gas + \beta_3 light)}}, \quad (4)$$

where  $\Pr$  is the posterior probability of being in the positive class of  $Q_{pleasantness}$ , and  $\alpha, \beta_1, \beta_2,$  and  $\beta_3$  are the parameters to be learned. To binarize  $Q_{pleasantness}$  (i.e., positive and negative class), we defined a set  $P$  as the set of all participants' responses to  $Q_{pleasantness}$ , and we partitioned it into two groups such that  $P = P^> \cup P^<$ , where  $P^>$  is the set of participants who reported higher levels of room pleasantness (above the median) defined as  $P^> = \{p \mid p > \bar{p}\}$ , where  $p$  is a participant's  $Q_{pleasantness}$  response, and  $P^<$  is the set of participants who reported lower levels of room pleasantness (below the median) defined as  $P^< = \{p \mid p < \bar{p}\}$ , where  $p$  is a participant's  $Q_{pleasantness}$  response. Then, the  $P^>$  responses were assigned to the positive class, while the  $P^<$  ones to the negative (Figure 6d).

Secondly, using the learned parameters (i.e.,  $\alpha, \beta_1, \beta_2,$  and  $\beta_3$ ) from (4), we linearly combined them and computed *sensed* as:

$$sensed = (\alpha + \beta_1 temp + \beta_2 gas + \beta_3 light), \quad (5)$$

where *temp*, *gas*, and *light* are the sensor readings obtained using our Gecko devices. We then computed *z-sensed* as the z-score ( $\frac{x - \mu_x}{\sigma_x}$ ) of *sensed*.

The *not-sensed* pleasantness metric captures the difference between the sensed and the self-reported pleasantness, and we computed it as:

$$z\text{-not-sensed} = (z\text{-sensed} - z\text{-}Q_{pleasantness}) \quad (6)$$

The higher its value, the more our sensors overestimated what people felt, while the lower its value, the more our sensors underestimated what people felt.

We build the  $M_3$  logistic regression model (7) to address  $H_3$  using the  $Q_{psychological}$ , *z-sensed*, and *z-not-sensed*.

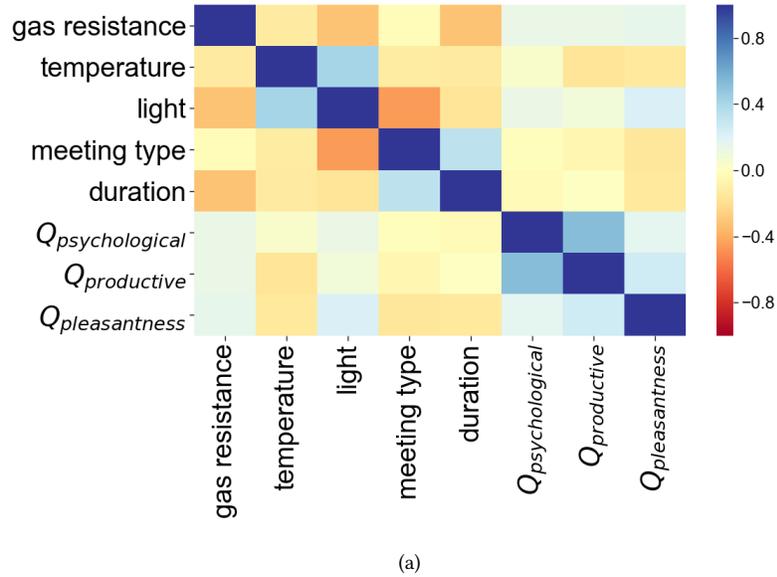


Fig. 7. Cross-correlation matrix of the variables in our analyses. Overall, the independent variables are fairly orthogonal, while some pairs of variables such as light and gas resistance show weak to moderate correlation.

$$\Pr(Q_{\text{productive}} = 1 \mid Q_{\text{psychological}}, z\text{-sensed}, z\text{-not-sensed}) = \frac{1}{1 + \exp^{-(\alpha + \beta_1 Q_{\text{psychological}} + \beta_2 z\text{-sensed} + \beta_3 z\text{-not-sensed})}} \quad (7)$$

where  $\Pr$  is the posterior probability of being in the positive class of  $Q_{\text{productive}}$ , estimated through a linear combination of the input features (i.e.,  $Q_{\text{psychological}}$ ,  $z\text{-sensed}$ , and  $z\text{-not-sensed}$ ) passed through a sigmoidal function, and  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  are the parameters to be learned by model  $M_3$ .

## 7 RESULTS

To evaluate the performance of our models, we used the standard classification performance metrics. Specifically, we calculated the *Area Under the Receiver Operating Characteristics (ROC) curve (AUC)* for our predictions. The ROC is a function of the true positive rate (TPR) against the false positive rate (FPR) of a classifier. Hence AUC will be larger for models that are good at predicting the positive class while not, at the same time, misclassifying the negative class.

All three models' AUC scores and their significant predictors are shown in Table 5. We observed that  $M_1$  predicts meetings' productivity slightly better than tossing a coin (AUC = 56%), which led us to *reject  $H_1$*  based on the finding that whether a meeting was productive or not did not depend on the meeting's type or duration only. When adding the  $Q_{\text{psychological}}$  into  $M_2$  model, however, the performance drastically improved, reaching an AUC of 82%. This suggests that participants' perception on whether a meeting was productive is mediated by the psychological safety that they experienced, in line with previous findings on the importance of psychological factors that allows for effective coordination and cooperation in a meeting, leading us to *accept  $H_2$* . Finally, we found that the performance of our  $M_3$  model further improves reaching up to 86%. This translates into saying that

Table 5. Logistic regression coefficients.  $M_0$  model’s coefficients were used to predict *sensed* pleasantness;  $M_1$ ,  $M_2$  and  $M_3$  tested hypotheses  $H_1$ ,  $H_2$ , and  $H_3$  respectively. Statistically significant predictors at  $< .05$  are marked in bold. (-) indicates that a model was fit without that predictor.

| <b><math>M_0</math></b> |               | <b><math>M_1</math>   <math>M_2</math>   <math>M_3</math></b> |        |              |              |
|-------------------------|---------------|---|--------|--------------|--------------|
| <b>Predictors</b>       |               | <b>Predictors</b>   |        |              |              |
| Intercept ( $\alpha$ )  | -0.576        | Intercept ( $\alpha$ )  | -0.089 | -0.613       | -0.646       |
| temp                    | <b>-0.630</b> | type 0  | -0.165 | 0.051        | -            |
| gas                     | <b>1.121</b>  | type 1  | -0.089 | -0.635       | -            |
| light                   | <b>0.752</b>  | type 2  | 0.064  | -0.028       | -            |
| <b>AUC score</b>        | 77.08%        | duration  | -0.009 | -0.001       | -            |
|                         |               | $Q_{psychological}$   | -      | <b>1.526</b> | <b>1.389</b> |
|                         |               | z-sensed  | -      | -            | <b>1.005</b> |
|                         |               | z-not-sensed  | -      | -            | -0.256       |
|                         |               | <b>AUC score</b>  | 55.72% | 82.56%       | 85.84%       |

the environmental conditions, both actual and perceived ones, further mediate people’s perception on whether a meeting was productive. Therefore, we *accept*  $H_3$ , and we attempt to explain these observations by interpreting  $M_3$ ’s model coefficients.

Additionally, we run a fourth model by including all the seven predictors and observed an AUC of 86.6%. This yielded very little gain in AUC compared to  $M_3$ . Further, to explore whether there are any interactions between the variables in our analyses, we measured whether pairs of the independent variables correlated with each other. To do so, we computed the Spearman’s rank cross-correlation matrix among all pairs (Figure 7). Some pairs such as light and gas resistance, and light and meeting type (as certain types of meeting, such as status updates, tended to happen more often in the morning) show weak to moderate correlations, but overall our independent variables are fairly orthogonal.

To get an upper bound of the predictive difference corresponding to a unit difference in each predictor, we use the ‘divide by 4’ rule [24] (p. 82). This says that, in a logistic regression, by dividing each predictor’s coefficient by 4 one can draw conclusions for 1 unit increase in each predictor. Simply put, we can quantify the percentage of increase or decrease in the probability of the dependent variable to be in the positive class.

### 7.1 The Known Unknowns of Productivity

By applying this rule, we inspected  $M_3$ ’s model predictors’ coefficients and found that the probability of a meeting perceived as a productive one is increased by 35% every one standard deviation increase in the  $Q_{psychological}$ . This corroborates previous findings, which highlighted the importance of inclusiveness [8, 9], dominance [63], and comfort in sharing and contributing [10, 26] in meetings. Furthermore, we found that the probability of a meeting perceived as a productive one is increased by 25% every one standard deviation increase in the *z-sensed* pleasantness. This suggests that environmental conditions also impact meetings’ productivity, supporting this work’s overall hypothesis that environmental conditions impact not only long-term productivity (as suggested by previous work [5, 6, 47, 74]) but also short-term and immediate one. We also found that the probability of perceiving meetings to be productive is increased by 6% every one standard deviation decrease in the part of people’s self-reported  $Q_{pleasantness}$  that our sensors could not capture (encoded into *z-not-sensed*), confirming that ComFeel captured most of the variability in the environmental conditions people experienced (adding external

validity to our framework's evaluation) and, at the same time, suggesting the orthogonal importance of asking a simple question even in the presence of an effective sensing infrastructure.

## 8 DISCUSSION

Our operationalization of the concept of sensorial pleasantness and productivity allowed us to study, for the first time, the short term impact of environmental conditions on meetings' productivity. By analyzing an online survey, which we administered to 363 participants, we unveiled three orthogonal dimensions that capture meetings' psychological experience: (a) *psychological comfort*, (b) *productivity*, and (c) *room pleasantness*.

Using the collected data obtained from our ComFeel framework and employee's self-reported experiences during their meetings, we investigated the interplay of environmental conditions with productivity. We first captured the sensed pleasantness by employing a machine learning method on three environmental conditions (i.e., temperature, light, and gas resistance), moving beyond thermal comfort previous work focused on [72]. We built a model that learned the linear combination of these factors, and we found that gas resistance was the most predictive one, followed by light and temperature. We then built three models to test our three hypotheses by incrementally introducing variables. In our final  $M_3$  model, we found that, on average, the probability of a meeting perceived as a productive one increased by 35% for each standard deviation increase in the psychological safety participants experience. This result confirms previous findings reported in literature [8–10, 26, 63], and suggests that meetings' productivity is mediated by whether participants felt listened to and were able to freely contribute. Additionally, we found that the very same probability increased by as much as 25% for each standard deviation increase in room pleasantness, suggesting the short-term impact of environmental conditions in meetings' productivity.

### 8.1 Implications

The theoretical implications of our work concern the fields of Management and Organizational Science. By analyzing our online questionnaire, we found that productivity, psychological safety, and room pleasantness capture, to a great extent, people's experience of meetings. This finding corroborates previous ones, which highlighted the importance of inclusiveness [8, 9], dominance [63], and comfort in sharing and contributing in meetings [10, 26]. All these aspects could be potentially captured by simply administering only three questions, making it possible for a large organization to understand their employees' overall experience during meetings.

Practically, our work speaks to the broader Ubicomp community. We found that particular environmental properties were more significant than others to predict perceived environmental conditions, and this could drive the focus of the research on sensing. Specifically, among the three environmental properties used to capture room pleasantness, we found that gas resistance was the most significant one, while lighting conditions and temperature had a second-order, yet significant and considerable effect. A finding also aligned with previous work that showed that high exposures to volatile compounds (which gas resistance captures) were associated with a decline in cognitive functions [6]. As of immediate practical use, our ComFeel framework is an easy-to-deploy infrastructure that could be used across any meeting room within an organization. Its flexible design, data reliability (off-the-shelf environmental sensors), and low manufacturing costs make it an appealing device for future deployments across any organization, independently to its size. More broadly, our ComFeel framework could further facilitate research on Human-Building Interaction (HBI); an emerging area aiming at unifying HCI research in the built environment [1]. Building on HBI's mission, we foresee to be able to answer questions related to human comfort (consider situations in shared offices where comfort needs to be agreed upon, even with basic interventions such as opening a window to more automated ones like adjusting automatically the thermostat setting) and the broader social dimensions of the workplace [2, 3]. Beyond the workplace, ComFeel could also be utilized in smart factories, which are considered an integral part of the fourth industrial revolution [76]. The very

same technology used here might well be used in an industrial setting to allow companies to optimize for their workers' productivity and well-being in real time.

## 8.2 Limitations and Future Work

The current work is subject to three main limitations that call for further work. The first limitation concerns the online survey. While we adopted a rigorous statistical approach to derive the three main principal components, we eliminated seventeen questions from the initial set that account for 37.69% of the variance. As previously stated though (§3), a set of three questions that accounts for 60 percent of the total variance can be deemed as satisfactory [27]. Additionally, the choice of the main three principal components was also reinforced by the practicalities of asking three questions only, making it attractive for deployments in a real application. Although a questionnaire's length and its response rate is often debatable [62], a shorter set of questions makes our application practical to use. The second limitation is about the sample size used in the analyses. Despite the fact that we obtained a fairly large amount of raw environmental readings, and hundreds of self-reported answers, a larger scale deployment would further benefit our analyses beyond data obtained at the facilities of one building. Immediate future work includes the deployment of our ComFeel framework in a broader set of meetings within our office's environment and then across the entire organization, thus enabling us to investigate other factors such as people's role within the company (e.g., individual contributors, managers, executives). The third limitation concerns the data collection and, particularly, the off-the-shelf environmental sensors used to obtain the readings. Despite these sensors are commercial ones, and were validated separately, future directions include the integration of additional and alternative vendors into our framework. Additionally, in future, we plan to extend our devices by incorporating additional sensors that capture a wider set of environmental factors such as noise or humidity. Along these lines, our models make use of the perceived productivity levels that our participants reported, whereas alternative ways of obtaining objective measurements of meetings' productivity (e.g., number of completed agenda items) could also be investigated in future work.

## 9 CONCLUSION

We studied the short term impact of indoor environmental quality on productivity at one of the most common daily activity at work: *meetings*. We operationalized the concept of sensorial pleasantness and productivity in meetings by analyzing an online questionnaire, and identified three factors that capture people's experience of meetings, namely psychological comfort, productivity, and room pleasantness. Building on these insights, we measured room pleasantness using ComFeel, an indoor environmental sensing infrastructure we developed that captures light, temperature, and gas resistance readings through our Gecko devices. We deployed our ComFeel across 29 real-world meetings, which spanned over a period of one month, and collected 1373 minutes of environmental sensor readings. For each of these meetings, we also collected whether each participant felt the meeting to have been productive, the setting to be psychologically safe, and the meeting room to be pleasant. As one expects, we found that, on average, the probability of a meeting being productive increased by 35% for each standard deviation increase in the psychological safety participants experienced. Importantly, the very same probability increased by as much as 25% for each standard deviation increase in room pleasantness. These results suggest that environmental conditions do matter not only in the long term but also in the short term so much so that significant differences in productivity were observed even within the constrained space of individual meetings.

## ACKNOWLEDGMENTS

We thank Philip Derrick for his contributions in the development of the Gecko devices. Thanks also to our participants and anonymous reviewers for their comments and feedback on this manuscript.

## REFERENCES

- [1] Hamed S Alavi, Elizabeth F Churchill, Mikael Wiberg, Denis Lalanne, Peter Dalsgaard, Ava Fatah gen Schieck, and Yvonne Rogers. 2019. Introduction to Human-Building Interaction (HBI) Interfacing HCI with Architecture and Urban Design. *ACM Transactions on Computer-Human Interaction (TOCHI)* 26, 2 (2019).
- [2] Hamed S Alavi, Denis Lalanne, Julien Nembrini, Elizabeth Churchill, David Kirk, and Wendy Moncur. 2016. Future of Human-Building Interaction. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 3408–3414.
- [3] Hamed S Alavi, Himanshu Verma, Michael Papinutto, and Denis Lalanne. 2017. Comfort: A Coordinate of User Experience in Interactive Built Environments. In *IFIP Conference on Human-Computer Interaction*. Springer, 247–257.
- [4] Akram Syed Ali, Christopher Coté, Mohammad Heidarinejad, and Brent Stephens. 2019. Elemental: An Open-Source Wireless Hardware and Software Platform for Building Energy and Indoor Environmental Monitoring and Control. *Sensors* 19, 18 (2019), 4017.
- [5] Joseph G. Allen. 2017. Stale Office Air Is Making You Less Productive. *Harvard Business Review* (2017).
- [6] Joseph G Allen, Piers MacNaughton, Usha Satish, Suresh Santanam, Jose Vallarino, and John D Spengler. 2016. Associations of Cognitive Function Scores with Carbon Dioxide, Ventilation, and Volatile Organic Compound Exposures in Office Workers: A Controlled Exposure Study of Green and Conventional Office Environments. *Environmental Health Perspectives* 124, 6 (2016), 805–812.
- [7] Bon Adriel Aseniero, Marios Constantinides, Sagar Joglekar, Ke Zhou, and Daniele Quercia. 2020. MeetCues: Supporting Online Meetings Experience. In *Proc. of the IEEE Visualization Conference (VIS)*. IEEE.
- [8] Paul Axtell. 2016. 6 Reasons to Get Better at Leading Meetings. *Harvard Business Review* (2016).
- [9] Paul Axtell. 2017. How to Design Meetings Your Team Will Want to Attend. *Harvard Business Review* (2017).
- [10] Paul Axtell. 2018. How to Respond When You're Put on the Spot in a Meeting. *Harvard Business Review* (2018).
- [11] Patrick Bader, Alexandra Voit, Huy Viet Le, Paweł W Woźniak, Niels Henze, and Albrecht Schmidt. 2019. WindowWall: Towards Adaptive Buildings with Interactive Windows as Ubiquitous Displays. *ACM Transactions on Computer-Human Interaction (TOCHI)* 26, 2 (2019), 1–42.
- [12] Zsolt Bako-Biro, Paweł Wargocki, Charles J Weschler, and Povl Ole Fanger. 2004. Effects of pollution from personal computers on perceived air quality, SBS symptoms and productivity in offices. *Indoor air* 14, 3 (2004), 178–187.
- [13] Murray R Barrick and Michael K Mount. 1991. The Big Five Personality Dimensions and Job Performance: A Meta-analysis. *Personnel psychology* 44, 1 (1991), 1–26.
- [14] P Sherwood Burge. 2004. Sick building syndrome. *Occupational and Environmental Medicine* 61, 2 (2004), 185–190.
- [15] Raymond B Cattell. 1966. The Scree Test for the Number of Factors. *Multivariate Behavioral Research* 1, 2 (1966), 245–276.
- [16] Derek Clements-Croome and Li Baizhan. 2000. Productivity and indoor environment. In *Proceedings of Healthy Buildings*, Vol. 1. 629–634.
- [17] Jordan Cohen. 2016. Use Subtle Cues to Encourage Better Meetings. *Harvard Business Review* (2016).
- [18] Valeria De Giuli, Roberto Zecchin, Livio Corain, and Luigi Salmaso. 2014. Measured and Perceived Environmental Comfort: Field Monitoring in an Italian School. *Applied Ergonomics* 45, 4 (2014), 1035–1047.
- [19] Valeria De Giuli, Roberto Zecchin, Luigi Salmaso, Livio Corain, and Michele De Carli. 2013. Measured and Perceived Indoor Environmental Quality: Padua Hospital Case Study. *Building and Environment* 59 (2013), 211–226.
- [20] Leandre R Fabrigar, Duane T Wegener, Robert C MacCallum, and Erin J Strahan. 1999. Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Methods* 4, 3 (1999), 272.
- [21] Xitao Fan, Brent C Miller, Kyung-Eun Park, Bryan W Winward, Mathew Christensen, Harold D Grotevant, and Robert H Tai. 2006. An Exploratory Study About Inaccuracy and Invalidity in Adolescent Self-report Surveys. *Field Methods* 18, 3 (2006), 223–244.
- [22] P Ole Fanger. 1988. Introduction of the Olf and the Decipol Units to Quantify Air Pollution Perceived by Humans Indoors and Outdoors. *Energy and Buildings* 12, 1 (1988), 1–6.
- [23] William J Fisk. 2000. Health and Productivity Gains from Better Indoor Environments And Their Relationship With Building Energy Efficiency. *Annual Review of Energy and the Environment* 25, 1 (2000), 537–566.
- [24] A Gelman and J Hill. 2007. Analytical Methods for Social Research. *Data analysis using regression and multilevel/hierarchical models* (2007).
- [25] Sarah Goff-Dupont. 2018. 6 types of meetings that are actually worthwhile. <https://thinkgrowth.org/6-types-of-meetings-that-are-actually-worthwhile-432877707493>
- [26] Joseph Grenny. 2017. How to Save a Meeting That's Gotten Tense. *Harvard Business Review* (2017).
- [27] Joseph F Hair, William C Black, Barry J Babin, Rolph E Anderson, Ronald L Tatham, et al. 1998. *Multivariate Data Analysis*. Vol. 5. Prentice Hall.
- [28] Eriko Hantani, Toshiharu Ikaga, Shuzo Murakami, and Ken Ichi Kameda. 2009. Effect of Indoor Environmental Quality on Performance and Satisfaction in Call-center. In *9th International Healthy Buildings Conference and Exhibition*.
- [29] Charlie Huizenga, Sahar Abbaszadeh, Leah Zagreus, and Edward A Arens. 2006. Air Quality and Thermal Comfort in Office Buildings: Results of a Large Indoor Environmental Quality Survey. *Proceeding of Healthy Buildings 2006* 3 (2006), 393–397.

- [30] Michelle T Iaffaldano and Paul M Muchinsky. 1985. Job Satisfaction and Job Performance: A Meta-analysis. *Psychological bulletin* 97, 2 (1985), 251.
- [31] Hirotake Ishii, Hidehiro Kanagawa, Yuta Shimamura, Kosuke Uchiyama, Kazune Miyagi, Fumiaki Obayashi, and Hiroshi Shimoda. 2018. Intellectual productivity under task ambient lighting. *Lighting Research & Technology* 50, 2 (2018), 237–252.
- [32] Lawrence R James, Robert G Demaree, and Gerrit Wolf. 1984. Estimating within-group interrater reliability with and without response bias. *Journal of applied psychology* 69, 1 (1984), 85.
- [33] Ming Jin, Nikolaos Bekiaris-Liberis, Kevin Weekly, Costas J Spanos, and Alexandre M Bayen. 2016. Occupancy Detection via Environmental Sensing. *IEEE Transactions on Automation Science and Engineering* 15, 2 (2016), 443–455.
- [34] Ming Jin, Han Zou, Kevin Weekly, Ruoxi Jia, Alexandre M Bayen, and Costas J Spanos. 2014. Environmental Sensing by Wearable Device for Indoor Activity and Location Estimation. In *IECON 2014-40th Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 5369–5375.
- [35] Henry F Kaiser. 1960. The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement* 20, 1 (1960), 141–151.
- [36] Shengxian Kang, Dayi Ou, and Cheuk Ming Mak. 2017. The Impact of Indoor Environmental Quality on Work Productivity in University Open-plan Research Offices. *Building and Environment* 124 (2017), 78–89.
- [37] Majid Karami, Gabrielle Viola McMorrow, and Liping Wang. 2018. Continuous monitoring of indoor environmental quality using an Arduino-based data acquisition system. *Journal of Building Engineering* 19 (2018), 412–419.
- [38] Simone Kauffeld and Nale Lehmann-Willenbrock. 2012. Meetings Matter: Effects of Team Meetings on Team and Organizational Success. *Small Group Research* 43, 2 (2012), 130–158.
- [39] Jasmine Kim. 2019. *6 Most Common Types of Business Meetings*. <https://www.lifesize.com/en/video-conferencing-blog/types-of-business-meetings>
- [40] Jungsoo Kim, Richard de Dear, Christhina Candido, Hui Zhang, and Edward Arens. 2013. Gender Differences in Office Occupant Perception of Indoor Environmental Quality (IEQ). *Building and Environment* 70 (2013), 245–256.
- [41] SC Lee, Sanches Lam, and Ho Kin Fai. 2001. Characterization of VOCs, Ozone, and PM10 Emissions from Office Equipment in an Environmental Chamber. *Building and Environment* 36, 7 (2001), 837–842.
- [42] Richard M Lent. 2015. *Leading Great Meetings: How to Structure Yours for Success*. Meeting for Results.
- [43] Eun Leslie A. Perlow, Constance Noonan Hadley. 2017. *Stop the Meeting Madness*. <https://hbr.org/2017/07/stop-the-meeting-madness>
- [44] Susana Lloret, Adoración Ferreres, Ana Hernández, and Inés Tomás. 2017. The Exploratory Factor Analysis of Items: Guided Analysis Based on Empirical Data and Software. *Anales de Psicología* 33, 2 (2017), 417–432.
- [45] Piers MacNaughton, James Pegues, Usha Satish, Suresh Santanam, John Spengler, and Joseph Allen. 2015. Economic, Environmental and Health Implications of Enhanced Ventilation in Office Buildings. *International Journal of Environmental Research and Public Health* 12, 11 (2015), 14709–14722.
- [46] MeetingSift. 2020. *The Six Most Common Types of Meetings*. <http://meetingsift.com/the-six-types-of-meetings/>
- [47] Jeanne C. Meister. 2018. The #1 Office Perk? Natural Light. *Harvard Business Review* (2018).
- [48] Mark J Mendell, Gina M Naco, Thomas G Wilcox, and W Karl Sieber. 2003. Environmental Risk Factors and Work-related Lower Respiratory Symptoms in 80 Office Buildings: An Exploratory Analysis of NIOSH data. *American Journal of Industrial Medicine* 43, 6 (2003), 630–641.
- [49] Minute. 2020. *5 Types of Meetings: All You Need to Know for Successful Meetings*. <https://www.getminute.com/types-of-meetings/>
- [50] Shayan Mirjafari, Kizito Masaba, Ted Grover, Weichen Wang, Pino Audia, Andrew T Campbell, Nitesh V Chawla, Vedant Das Swain, Munmun De Choudhury, Anind K Dey, et al. 2019. Differentiating Higher and Lower Job Performers in the Workplace Using Mobile Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 2 (2019), 1–24.
- [51] Clifford S Mitchell, Junfeng Zhang, Torben Sigsgaard, Matti Jantunen, Paul J Liroy, Robert Samson, and Meryl H Karol. 2007. Current State of the Science: Health Effects and Indoor Environmental Quality. *Environmental Health Perspectives* 115, 6 (2007), 958–964.
- [52] KW Mui and LT Wong. 2006. Acceptable Illumination Levels for Office occupants. *Architectural Science Review* 49, 2 (2006), 116–119.
- [53] Igor Mujan, Aleksandar S Andelković, Vladimir Munćan, Miroslav Kljajić, and Dragan Ružić. 2019. Influence of indoor environmental quality on human health and productivity-A review. *Journal of cleaner production* 217 (2019), 646–657.
- [54] Matthew Neidell. 2017. Air pollution and worker productivity. *IZA World of Labor* (2017).
- [55] Karin Niemantsverdriet and Thomas Erickson. 2017. Recurring Meetings: An Experiential Account of Repeating Meetings in a Large Organization. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 84.
- [56] Raimo Niemelä, Mika Hannula, Sari Rautio, Kari Reijula, and Jorma Railio. 2002. The effect of air temperature on labour productivity in call centres—a case study. *Energy and Buildings* 34, 8 (2002), 759–764.
- [57] Geoffrey R Norman and David L Streiner. 2008. *Biostatistics: The Bare Essentials*. PMPH.
- [58] Jum C Nunally and Ira H Bernstein. 1978. Psychometric Theory.

- [59] Thomas A O'Neill. 2017. An overview of interrater agreement on Likert scales for researchers and practitioners. *Frontiers in psychology* 8 (2017), 777.
- [60] Carrie A Redlich, Judy Sparer, and Mark R Cullen. 1997. Sick-Building Syndrome. *The Lancet* 349, 9057 (1997), 1013–1016.
- [61] Steven G Rogelberg, Linda Rhoades Shanock, and Cliff W Scott. 2012. Wasted Time and Money in Meetings: Increasing Return on Investment. *Small Group Research* 43, 2 (2012), 236–245.
- [62] Sindre Rolstad, John Adler, and Anna Rydén. 2011. Response Burden and Questionnaire Length: Is Shorter Better? A Review and Meta-Analysis. *Value in Health* 14, 8 (2011), 1101–1108.
- [63] Nicholas C Romano and Jay F Nunamaker. 2001. Meeting Analysis: Findings from Research and Practice. In *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*. IEEE, 13–pp.
- [64] Sebastiaan Rothmann and Elize P Coetzer. 2003. The Big Five Personality Dimensions and Job Performance. *SA Journal of Industrial Psychology* 29, 1 (2003), 68–74.
- [65] Daniel Russel. 2015. *America Meets a lot. An Analysis of Meeting Length, Frequency and Cost*. <http://attentiv.com/america-meets-a-lot/>
- [66] Riccardo Sartori, Arianna Costantini, Andrea Ceschi, and Andrea Scalco. 2017. Not Only Correlations: A Different Approach for Investigating the Relationship Between the Big Five Personality Traits and Job Performance Based on Workers and Employees' Perception. *Quality & Quantity* 51, 6 (2017), 2507–2519.
- [67] Neil L Sbar, Lou Podbelski, Hong Mo Yang, and Brad Pease. 2012. Electrochromic Dynamic Windows for Office Buildings. *International Journal of Sustainable Built Environment* 1, 1 (2012), 125–139.
- [68] Florian Schaule, Jan Ole Johanssen, Bernd Bruegge, and Vivian Loftness. 2018. Employing Consumer Wearables to Detect Office Workers' Cognitive Load for Interruption Management. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 1–20.
- [69] Roger Schwarz. 2016. 8 Ground Rules for Great Meetings. *Harvard Business Review* (2016).
- [70] Olli A Seppänen and William Fisk. 2006. Some Quantitative Relations Between Indoor Environmental Quality and Work Performance or Health. *Hvac&R Research* 12, 4 (2006), 957–973.
- [71] Stephen Snow. 2018. Indoor Air Quality: Opportunities for Behaviour Change Towards Healthier Offices, a Two-part Report. (2018).
- [72] Ying Song, Fubing Mao, and Qing Liu. 2019. Human Comfort in Indoor Environment: A Review on Assessment Criteria, Data Collection and Data Analysis Methods. *IEEE Access* 7 (2019), 119774–119786.
- [73] Jan Sundell. 2004. On the History of Indoor Air Quality and Health. *Indoor Air* 14, s 7 (2004), 51–58.
- [74] Tal Gross Tom Y. Chang, Joshua Graff Zivin and Matthew Neidell. 2016. Air Pollution Is Making Office Workers Less Productive. *Harvard Business Review* (2016).
- [75] Chockalingam Viswesvaran and Deniz S Ones. 2000. Perspectives on Models of Job Performance. *International Journal of Selection and Assessment* 8, 4 (2000), 216–226.
- [76] Shiyong Wang, Jiafu Wan, Di Li, and Chunhua Zhang. 2016. Implementing Smart Factory of Industrie 4.0: An Outlook. *International journal of distributed sensor networks* 12, 1 (2016), 3159805.
- [77] Pawel Wargocki, Jan Sundell, W Bischof, G Brundrett, Povl Ole Fanger, F Gyntelberg, SO Hanssen, P Harrison, A Pickering, O Seppänen, et al. 2002. Ventilation and health in non-industrial indoor environments: report from a European Multidisciplinary Scientific Consensus Meeting (EUROVEN). *Indoor air* 12, 2 (2002), 113–128.
- [78] David P Wyon. 2004. The Effects of Indoor Air Quality on Performance and Productivity. *Indoor Air* 14 (2004), 92–101.
- [79] Sailin Zhong, Hamed S Alavi, and Denis Lalanne. 2020. Hilo-wear: Exploring Wearable Interaction with Indoor Air Quality Forecast. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–8.