COVID-19 Living Evidence Synthesis 14.1b:
Unintended consequences/outcomes of masking in response to COVID-19 in non-health care community-based settings

Executive Summary

Question
What is the best-available evidence about the unintended health and social consequences/outcomes (positive and negative) of masking in response to COVID-19 in non-health care community-based settings?

Sub-questions:
1. What are the unintended physiological health consequences/outcomes of wearing a mask? (e.g., gas exchange/cardio-respiratory function, smothering feeling, skin impact, exercise impact, toxicity, pathogenicity)
2. What are the unintended psychological and/or developmental consequences/outcomes of mask wearing? (e.g., anxiety, discrimination/social pressures, facial recognition, reading faces and emotions, inability to lip read, harder to hear speech)

Background

• The aim of this review is to analyze and summarize the evidence from randomized controlled trials (RCTs) and quasi-experimental studies with comparison groups about the unintended consequences/outcomes of masks (including different types of masks) in response to COVID-19 in community settings. Other study designs (e.g., observational studies) were excluded.
• Unintended consequences occur when results differ from an expected outcome, in this case masking, and vary across context, country, and age group. Unintended consequences can be positive or negative.

Key points
• We included 46 studies, 30 RCTs and 16 quasi-experimental, conducted across 19 countries.
• Studies examining the physiological effects of masking while exercising (cycling, treadmill, walking) are heterogenous in intervention and outcome measures (32 studies).
• Most studies examining the physiological effects of masking were RCTs and conducted with healthy young adults aged 18-30 years (23/34 studies).
• Most common physiological parameters measured were heart rate (HR) (26 studies), oxygen saturation (SpO₂) (24 studies), and Rating of Perceived Exertion (RPE) (17 studies).
• RPE was recurrent in the data with significant unintended consequences/outcomes related to dyspnea (15 studies) and to overall body discomfort (2 studies).
• Surgical masks were the most frequently studied (35/46).
• Masking while exercising, although uncomfortable, is safe in young healthy people aged 18-30 years (22/26 studies).
• Body temperature and thermal comfort are not significantly altered by masking (3 studies).
• Masking may contribute to dry eye conditions (2 studies).
• Masks significantly impair facial emotion recognition (4 studies), possibly leading to difficulty with expressing and or recognizing emotional states during brief interactions.
• Misperceptions of expressions of emotions such as surprise, fear, anger, disgust, and happiness while masked are common (4 studies).
• Masks attenuate the accuracy of voice recognition in noisy environments (1 study).
• Surgical masks may have a slight edge compared to N95 and cloth in contributing to accuracy of voice recognition (1 study).
• Masking did not negatively impact cognitive performance in healthy young adults (18-30) (3 studies) and school children (grades 5-7) (1 study).
LES 14.1b: Unintended consequences of masking in response to COVID-19

- Mild but statistically significant changes in HR and SpO\textsubscript{2} did not attenuate cognitive performance in university students (1 study).

**Suggested Tweet**
Although heart rate and oxygen levels are largely unaffected by masking, it contributed to feelings of breathlessness and a perception of greater exertion.

**Date of Literature Search:** 22 February 2023


**Please note:** This living evidence synthesis (LES) is part of a suite of LESs of the best-available evidence about the effectiveness of six PHSMs (masks, quarantine and isolation, ventilation, physical distancing and reduction of contacts, hand hygiene and respiratory etiquette, cleaning, and disinfecting), as well as combinations of and adherence to these measures, in preventing transmission of COVID-19 and other respiratory infectious diseases in non-health care community-based settings. The next update to this and other LESs in the series is to be determined, but the most up-to-date versions in the suite are available on the COVID-END website. We provide context for synthesizing evidence about public health and social measures in Box 1 and an overview of our approach in Box 2.

**Box 1: Context for synthesizing evidence about public health and social measures (PHSMs)**

This series of living evidence syntheses was commissioned to understand the effects of PHSMs during a global pandemic to inform current and future use of PHSMs.

**General considerations for identifying, appraising and synthesizing evidence about PHSMs**

- PHSMs are population-level interventions and typically evaluated in observational studies.
  - Many PHSMs are interventions implemented at a population level, rather than at the level of individuals or clusters of individuals such as in clinical interventions.
  - Since it is typically not feasible and/or ethical to randomly allocate entire populations to different interventions, the effects of PHSMs are commonly evaluated using observational study designs that evaluate PHSMs in real-world settings.
  - As a result, a lack of evidence from RCTs does not necessarily mean the available evidence in this series of LESs is weak.
- Instruments for appraising the risk of bias in observational studies have been developed; however, rigorously tested and validated instruments are only available for clinical interventions.
  - Such instruments generally indicate that a study has less risk of bias when it was possible to directly assess outcomes and control for potential confounders for individual study participants.
  - Studies assessing PHSMs at the population level are not able to provide such assessments for all relevant individual-level variables that could affect outcomes, and therefore cannot be classified as low risk of bias.
- Given feasibility considerations related to synthesizing evidence in a timely manner to inform decision-making for PHSMs during a global pandemic, highly focused research questions and inclusion criteria for literature searches were required.
 LES 14.1b: Unintended consequences of masking in response to COVID-19

- As a result, we acknowledge that this series of living evidence syntheses – about the effectiveness of specific PHSMs (i.e., quarantine and isolation; mask use, including unintended consequences; ventilation, reduction of contacts, physical distancing, hand hygiene and cleaning and disinfecting measures), interventions that promote adherence to PHSMs, and the effectiveness of combinations of PHSMs – does not incorporate all existing relevant evidence on PHSMs.
- Ongoing work on this suite of products will allow us to broaden the scope of this review for a more comprehensive understanding of the effectiveness of PHSMs.
- Decision-making with the best available evidence requires synthesizing findings from studies conducted in real-world settings (e.g., with people affected by misinformation, different levels of adherence to an intervention, different definitions and uses of the interventions, and in different stages of the pandemic, such as before and after availability of COVID-19 vaccines).

**Our approach to presenting findings with an appraisal of risk of bias (ROB) of included studies**

To ensure we used robust methods to identify, appraise and synthesize findings and to provide clear messages about the effects of different PHSMs, we:

- acknowledge that a lack of evidence from RCTs does not mean the evidence available is weak
- assessed included studies for ROB using the approach described in the methods box
- typically introduce the ROB assessments only once early in the document if they are consistent across sub-questions, sub-groups and outcomes, and provide insight about the reasons for the ROB assessment findings (e.g., confounding with other complementary PHSMs) and sources of additional insights (e.g., findings from LES 20 in this series that evaluates combinations of PHSMs)
- note where there are lower levels of ROB where appropriate
- note where it is likely that risk of bias (e.g., confounding variables) may reduce the strength of association with a PHSM and an outcome from the included studies
- identify when little evidence was found and when it was likely due to literature search criteria that prioritized RCTs over observational studies.

**Implications for synthesizing evidence about PHSMs**

Despite the ROB for studies conducted at the population level that are identified in studies in this LES and others in the series, they provide the best-available evidence about the effects of interventions in real life. Moreover, ROB (and GRADE, which was not used for this series of LESs) were designed for clinical programs, services and products, and there is an ongoing need to identify whether and how such assessments and the communication of such assessments, need to be adjusted for public-health programs, services and measures and for health-system arrangements.
Findings

- Forty-six studies (30 randomized controlled trials [RCTs], 16 quasi-experimental) are included in this review. Thirty-four studies (n=1331) report on the physiological outcomes of masking, and 12 studies (n=2,148) report on the psychological and/or developmental outcomes of masking. Two of the 34 physiological studies are ocular in focus, with their own unique outcomes.
- JBI quality appraisal/risk of bias (ROB) for included studies rated 35/46 studies at moderate and the remainder at low for ROB. Of the RCTs, 24/30 were assessed as moderate risk of bias, and 6/30 were assessed as low risk of bias.
- A PRISMA 2020 flow diagram of the screening process is shown in Figure 1.

This LES included 46 studies addressing the physiological, psychological and/or developmental outcomes of wearing a face mask of one or more types versus no mask (NM). The studies were conducted in Germany (n=7), USA (n=7), China (n=5), the UK (n=3), Portugal (n=3), Spain (n=3), Canada (n=3), Brazil (n=2), Israel (n=2), and Tunisia (n=2), and the following countries had one study each: Saudi Arabia, Turkey, the Netherlands, Austria, Denmark, Iran, Italy, Japan, and Thailand.

Summary of findings about the primary outcome #1: Unintended physiological consequences/outcomes of masking

Thirty-four studies were included that report on the physiological effects of masking. The characteristics, findings, and assessment of risk of bias for each study are presented in Table 3 and Table 4.

Physiological parameters and masking

Key observations

- Studies examining the physiological effects of masking while exercising are heterogenous (cycling, treadmill, walking).
- Most studies examining the physiological effects of masking while exercising were conducted in Germany (n=7), USA (n=7), China (n=5), the UK (n=3), Portugal (n=3), Spain (n=3), Canada (n=3), Brazil (n=2), Israel (n=2), and Tunisia (n=2), and the following countries had one study each: Saudi Arabia, Turkey, the Netherlands, Austria, Denmark, Iran, Italy, Japan, and Thailand.

Box 2: Our approach

We retrieved candidate studies by searching: 1) MEDLINE; 2) the iCite pre-print server; 3) Embase; 4) CINAHL; and 5) ERIC. Searches were conducted for studies reported in English, conducted with humans, and published since 1 January 2020 (to coincide with the emergence of COVID-19 as a global pandemic). Our detailed search strategy is included in Appendix 1.

Randomized controlled trials (RCTs) and studies that report on empirical data with a comparator were considered for inclusion, with modelling studies, simulation studies, cross-sectional studies, case reports, case series, and press releases excluded. A full list of included studies is provided in Tables 3-5.

Reviews were not included but relevant reviews, including those identified in LES 14.1a, were mined for included studies.


Intervention and control/comparator: Masks and face coverings, including cloth masks and surgical grade masks, in public or private non-healthcare community settings.

Primary outcomes: 1. Unintended physiological consequences/outcomes (e.g., impacts on gas exchange/cardiorespiratory function, discomfort/sensation of inability to breathe, impacts on ability to exercise, impacts on skin, other impacts); 2. Unintended psychological and developmental consequences/outcomes (e.g., anxiety, discrimination, social pressures, reading faces and emotions, facial recognition, inability to lip-read, harder to hear speech, etc.).

Data extraction: Data extraction was conducted by one team member and checked for accuracy and consistency by another using the template provided in Appendix 2.

Critical appraisal: Risk of Bias (ROB) of individual studies was assessed using JBI Critical Appraisal tools. For RCTs we used JBI Critical Appraisal Tool for Assessment of Risk of Bias for Randomized Controlled Trials (Barker et al., 2023), for non-randomized experimental studies we used JBI Checklist for Quasi-Experimental Studies (Non-Randomised Experimental Studies) (Tufanaru et al., 2020), and for case control studies we used JBI Checklist for Case Control Studies (Moola et al., 2020). Our detailed approach to critical appraisal is provided in Appendix 3.

Summaries: We summarized the evidence by presenting key observations by outcome measure. Future versions may include statistical pooling of results if deemed appropriate.

The next update to this document is to be determined.
RCTs and conducted with healthy young adults aged 18-30 (23/32).

- Most common physiological parameters measured were HR, oxygen saturation (SpO₂), Rating of Perceived Exertion (RPE) (See Appendix 4 for more information on RPE).
- RPE was recurrent in the data with significant unintended consequences in relation to dyspnea and to overall body discomfort.
- Surgical masks were the most frequently studied in community settings (26/32).
- Masking while exercising, although uncomfortable, is safe in young healthy people aged 18-30.

Thirty-two studies (see Table 3) involved a variety of physical activity interventions (cycling – 14, treadmill – 7, walking – 4, running – 4, whole body vibration – 1, performing light tasks – 1, warm-up exercises - 1). The physiological parameters most frequently measured wearing masks (surgical, N95 and cloth) versus NM included heart rate (HR), SpO₂, RPE, respiratory/breathing rate (RR), blood pressure (BP), and end-tidal carbon dioxide (etCO₂) (note that for the purposes of this report, we use the term N95 to encompass FFP1, FFP2, FFP3, K95, KN95, etc. mask types). Other physiological parameters less commonly measured included spirometry measures of lung function (forced vital capacity [FVC], forced expiratory volume [FEV1]), and work rate in watts was measured in several studies, as well as serum lactate.

Twenty-five of the 32 studies that involved physical activity had study populations of healthy young adults between the ages of 18-30 years. Six of these studies reported no significant unintended consequences of masking of all types across all physiological parameters measured (Ahmadian et al., 2021; Apolo-Arenas et al., 2022; Guardieiro et al., 2024 low ROB; Martin et al., pre-print; Pasqualetto, 2022 low ROB; Shaw et al., 2020 low ROB). Three of these six studies had a low ROB and the remainder moderate.

Fifteen studies found no statistically significant changes in HR and SpO₂ (Ade et al., 2021; Cabanillas-Barea et al., 2021; Doherty et al., 2021; Driver et al., 2022; Egger et al., 2022; Epstein et al., 2021; Fischer et al., 2022; Ferguson et al., 2023 low ROB; Jesus et al., 2022; Jones et al., 2023; Ng et al., 2022; Poon et al., 2021 low ROB; Dacha et al., 2022; Steinhilber et al., 2022; Yoshihara et al., 2021). These 15 combined with the six above total 21 studies in which HR and SpO₂ were not significantly negatively altered while exercising with a mask in healthy young adults. Five of these 21 studies had a low ROB and the remainder moderate.

Of the remaining studies, one reported statistically significant changes in SpO₂ (Bar-On et al., 2021), and four reported statistically significant changes in HR (Fukushi et al., 2021 low ROB; Lassing, Vogt et al., 2022; Zhang et al., 2021). Five studies did not measure HR and SpO₂ (Dantas et al., 2021; Morris et al., 2021, Slimani et al., 2021, Slimani et al., 2022; Zhou & Dong, 2023).

Twenty-two of the 32 studies that involved an exercise activity measured RPE using the Borg Modified RPE Scale (0-10) and or original Borg (6-20). RPE (0-10) was sometimes referred to by authors of included studies as the Borg dyspnea scale particularly where participants reported breathlessness or a feeling of smothering while exercising. Sixteen of 21 studies examining RPE reported statistically significant unintended consequences/outcomes (Ade et al., 2021 dyspnea; Bar-On et al., 2021 perceived exertion; 2021; Dantas et al., 2021 perceived exertion; Driver et al., 2022 perceived exertion; Steinhilber et al., 2023 dyspnea low ROB; Poon et al., 2021 perceived exertion low ROB; Dacha et al., 2022 dyspnea; Slimani et al., 2021 perceived exertion; Steinhilber et al., 2022 perceived respiratory effort; Vogt et al., 2022 perceived exertion; Wong et al., 2020 perceived exertion; Yoshihara et al., 2021 breathing discomfort; Zhang et al., 2021 dyspnea). Three of these 16 studies had low ROB and the remainder moderate ROB. Further explanation of these instruments is included in Appendix 4.

We purposely decided against a description of each of the 32 studies that involved exercising while masked because most were conducted with healthy young adults and could not be considered representative of the population in community settings. Instead, we focus on the studies with populations that better
approximate community settings. Six studies that involved a physical activity and that examined physiological outcomes with more heterogeneous samples are illustrated in Table 1 and described below (Bar-On et al., 2021; Fischer et al., 2022; Martin et al., pre-print; Morris et al., 2021; Steinhilber et al., 2022; Vogt et al., 2022). ROB for these studies was moderate.

Table 1: Physiological results, physical activity studies with heterogenous samples

<table>
<thead>
<tr>
<th>Citation</th>
<th>Country</th>
<th>n</th>
<th>Mask type</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bar-On et al. (2021)</td>
<td>Israel</td>
<td>21</td>
<td>SM</td>
<td>Sig↑: etCO₂ (during slow walk) and RPE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NS: SpO₂</td>
</tr>
<tr>
<td>Fischer et al. (2022)</td>
<td>Germany</td>
<td>50</td>
<td>SM, N95</td>
<td>Sig↓: peak power cycling (SM &amp; N95 vs NM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sig↑: etCO₂ (N95)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NS: HR, serum lactate</td>
</tr>
<tr>
<td>Martin et al. (pre-print)</td>
<td>Austria</td>
<td>10</td>
<td>SM</td>
<td>NS: across parameters while climbing stairs</td>
</tr>
<tr>
<td>Morris et al. (2021)</td>
<td>Denmark</td>
<td>8</td>
<td>N95</td>
<td>Sig↑: RPE for prolonged N95 use during light tasks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NS: HR, SpO₂, body &amp; face temperature, cognitive performance</td>
</tr>
<tr>
<td>Steinhilber et al., 2022</td>
<td>Germany</td>
<td>39</td>
<td>SM, N95, cloth</td>
<td>RPE trended upward when masked (P &lt; 0.05) while cycling</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NS: physical performance, SpO₂, BP and pCO₂</td>
</tr>
<tr>
<td>Vogt et al. (2022)</td>
<td>USA</td>
<td>19</td>
<td>SM, N95</td>
<td>Sig↑: HR (SM), etCO₂ (N95 vs NM; N95 vs SM), RPE (N95), RPD (N95), workload (N95 &amp; SM) while cycling</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NS: workload across all masks</td>
</tr>
</tbody>
</table>

Sig↓ = Significant decrease; Sig↑ = Significant increase; NS = No significant difference
SM = surgical mask; NM = no mask; HR = heart rate; RPE = rating of perceived exhaustion; RPD = rating of perceived dyspnea; etCO₂ = end-tidal carbon dioxide; pCO₂ = partial pressure of carbon dioxide; SpO₂ = oxygen saturation

Bar-On et al. (2021) (Israel) conducted a controlled trial study to evaluate the effects of surgical facemasks on gas exchange. The sample consisted of 21 healthy adults ranging in age from 29-57 years, mean age 38, 11 females and 10 males. Participants engaged in five-minute slow and brisk walks with mask and without. Parameters measured included SpO₂, etCO₂, and RPE using the Borg scale to rate their feeling of effort/exertion/breathlessness and fatigue during the activity (see Appendix 4 for more information on the Borg scale). There was no significant decrease in SpO₂, etCO₂ increased significantly, P = 0.004 during the slow walk. This increase was speculated to be due to the rebreathing of expired air. During the brisk walk while masked resulted in a statistically significant increase in etCO₂ from baseline without a mask P = 0.009. There was a statistically significant increase in RPE: results while masked with both slow and brisk walking P = -0.002 and P < 0.001 respectively. These authors summarized the results stating, wearing masks is safe, and that the changes in SpO₂ and etCO₂ were small yet reached significance. They speculate whether the gas exchange abnormalities would have been greater if the exercise intensity had been greater.

Fischer et al. (2022) (Germany) conducted a RCT at a University Hospital Medical Centre with a group of participants diagnosed with cardiovascular disease (CVD) and hypertension and scheduled for cycling exercise tests for clinical reasons. Sample included 50 participants, 40 with CVD as the intervention group and 10 healthy untreated participants as controls. Each participant completed the same ET protocol twice or 3 times (SM, N95, NM), respectively and were instructed to maintain a constant speed between 50 and 60 revolutions per minute. The protocol started with a 30 second warm-up period without load, followed by 1 min of constant load at 25 W. Thereafter, the load was increased every minute for 25 W using a ramp protocol. Participants exercised until they were exhausted or showed signs of discomfort such as angina, dyspnea, or dizziness. Parameters measured included peak power, serum lactate, partial pressure of carbon dioxide (pCO₂, capillary gas). In the intervention group power output differed significantly between surgical mask (SM) and NM -5.0 ± 7.0%, P = 0.005; FFP2 vs NM: -4.7 ± 14%, P = 0.03; control group: SM vs NM: -6.8 ± 4.4%, P = 0.008; FFP2 vs NM: -8.9 ± 6.3%; P = 0.01. Earlobe capillary pCO₂ levels increased...
Unintended consequences of masking in response to COVID-19

significantly in CVD group wearing a N95 compared with NM: 36.0 ± 3.2 mm Hg vs 33.3 ± 4.4 mm Hg, \( P = 0.019 \); control group: N95: 32.6 ± 2.8 mm Hg vs NM: 28.1 ±1.7 mm Hg, \( P < 0.001 \). Serum lactate changes were not considered clinically critical and no significant changes in hemodynamic parameters (HR, BP). Authors conclude exercise testing while wearing a face mask is feasible for people with CVD.

Martin et al. (pre-print) (Austria) evaluated (design not reported) the impact of SMs on HR and SpO\(_2\) in a sample of 10 healthy individuals, five <30 years of age and five >45 years of age by having them climb 96 steps. There was no significant difference in parameters measured and authors concluded that healthy adults wearing a SM while exercising did not experience a significant effect on oxygen supply or cardiac load.

Morris et al. (2021) (Denmark) conducted a counter-balance crossover study of effects of mask versus NM on thermal comfort, perceived dyspnea, HR, SpO\(_2\), body temp, skin temp and cognitive function. The study sample \( n=8 \) male participants with a mean age of 35 carried out 45 minutes of light exercise in a climate chamber approximating work in healthcare settings. Cognitive assessments consisted of math calculations, body temp was rectal, skin temp was measured in two places on the face, dyspnea was measured with the Borg breathlessness scale and thermal comfort by a whole-body comfort scale. Perceived dyspnea worsened significantly with prolonged FM use \( P = 0.04 \), neither body temp nor facial temp were affected by masking. Cognitive performance, HR and SpO\(_2\) were not affected by mask wearing.

Steinhilber et al. (2022) (Germany) conducted a randomized cross-over study in a sample \( n=39 \) of healthy adults, 20 men average age 38 and four were smokers. The study aim was to investigate whether face masks (SM, N95, cloth) would impair physical performance and affect physiological and subjective response during submaximal physical activity. Primary outcome was physical working capacity (watts/kg to measure mechanical load of the heart at a rate of 130 beats per minute) with and without mask. Additional outcomes SpO\(_2\), BP, pCO\(_2\), respiratory effort and perceived exertion (Modified Borg 0-10). Participants completed a cycle ergometry protocol to exhaustion. Masking had no statistically significant effect on physical performance, SpO\(_2\), BP and pCO\(_2\). RPE trended upward when masked \(( P < 0.05)\).

Vogt et al. (2022) (USA) in a randomized crossover study examined the effects of different types of masks (NM, SM, and N95) on physiological and perceptual responses during 30-min of self-paced cycle ergometer exercise. The sample \( n=19 \) consisted of community dwelling adults ranging in age from 54-83 years, nine males, 10 females, four with a history of CVD, six taking prescribed medications for hypertension, four with high cholesterol levels, two with sleep apnea, two with multiple sclerosis, one with type I diabetes, one with type II diabetes and two with vertigo. Outcomes included workload measured in watts (W), SpO\(_2\), etCO\(_2\), HR, RR, RPE, RPD. RPE, RPD, and etCO\(_2\) were significantly higher with a N95-mask vs. NM \(( P = 0.012)\), \(( P = 0.002)\), \(( P < 0.001)\) respectively. HR was significantly higher with the SM compared to NM \(( P = 0.027)\) (NM 107.18 ± 9.96) (SM 112.34 ± 10.28), but no significant difference was found when comparing the SM to the N95 condition or when comparing the N95 condition to the NM. Watts increased across time in each mask type \(( P = 0.003)\). In summary no significant differences were found in workload, RR or SpO\(_2\) regardless of mask type. The N95 mask was associated with increased breathing resistance suggesting trapping of expired air as suggested by Bar-On et al. (2021) above. Wearing a N95 mask may be less comfortable for older adults.

Thermal comfort and masking

Key observations
- Body temperature and thermal comfort are not significantly altered by masking.
- Masking does contribute to increased RPE and overall body discomfort.

In this LES of the unintended consequences of masking three studies (Morris et al., 2021; Yoshihara et al., 2021; Zhou & Dong, 2023) meeting the review criteria addressed thermal comfort; however, only one of the three had this as the unique focus (Zhou & Dong, 2023). ROB for these studies was moderate.
In a longitudinal (9 months) experimental study conducted in China, the neutral temperature, outdoor thermal comfort, effects of masks, and outdoor conditions were studied across the seasons with participants while masked (SM), unmasked, sitting quietly and while walking on a treadmill (Zhou & Dong, 2023). The sample consisted of n=42 participants, 20 males, and 22 females who used an outdoor rooftop garden on a regular basis and were in good health. Neutral temperature refers to (point at which the human body can achieve an equilibrium core temperature without shivering or sweating) in terms of skin temperature the neutral zone is in the range of 33 to 35 C, thermal comfort masked and unmasked, and general body discomfort. Results demonstrated significantly lower neutral temperatures in summer meaning the body worked harder to establish core temp equilibrium; discomfort with masks was worse walking than sitting, and facial and chest discomfort increased with mask wearing. Mask wearing was better tolerated at lower outdoor climate temperatures.

Yoshihara et al. (2021) (USA) conducted a RCT focused on thermoregulation to determine if a FM influenced rectal temperature, HR and RPE while walking/jogging for 60 minutes in the heat. The study sample n=12, consisted of 8 males and 4 females, mean age 24 and physically active. Participants had a choice of face mask (SM, N95, gaiter, sport) and performed an exercise protocol for treadmill. Results included no significant difference between FM type, and no significant changes in Temp and HR masked or unmasked. RPE, overall breathing discomfort was significantly higher in masked compared with unmasked. A description of the Morris et al. (2021) study is provided in the previous section.

In these three studies regarding thermal comfort and regulation physiological variables of temperature, HR and SpO2 were not significantly altered while masked compared to NM. Across the dataset for this review the significance of breathing discomfort while masked is recurrent and evident in relation to thermal comfort as well.

Physiological measures of eye surface and masking (no exercise involved)

Key Observations

- Masking may contribute to dry eye conditions.

Two studies focused on the effects of masking on the surface of the eye and although physiologic measures were examined the measures were not related to the cardio-respiratory systems (see Table 4). One study with low ROB and the other moderate.

Marta et al. (2022) low ROB (Portugal) examined the surface of the eye to further understanding of exacerbated dry eye symptoms during the COVID-19 pandemic. Using a retrospective comparative design with a sample of n=274 patients, 548 eyes, mean age 66.15, and range of 18-89, three groups were formed. Group 1 retrospective data pre-pandemic, Group 2 early pandemic before mask mandates, Group 3 mask mandates in place. Parameters measured included Lipid Layer Thickness (LLT), blink rate, tear meniscus height, osmolarity and impact on meibomian (lipid source), and non-invasive tear break-up time (NITBUT). LLT was better in Groups 2 ($P = 0.001$) and 3 ($P < 0.001$); Blink rate and tear meniscus height were similar in group 2 and worse in group 3 ($P < 0.001$ and $P = 0.038$, respectively); Tear osmolarity and loss area of the meibomian glands were worse in group 2 ($P = 0.031$ and $P < 0.001$, respectively) and in group 3 (both with $P < 0.001$). Non-invasive tear break-up time (NITBUT) tests tear film for dry eyes was worse in group 2 ($P = 0.030$) and similar in group 3 ($P = 0.263$). Authors suggest the increase in LLT was to compensate for surface evaporation from increased air flow over the eye from masks, and this plus increased screen time from (e-devices) contributed to decreased tear osmolarity.

Alanazi et al. (2022) (Saudi Arabia) assessed the impact of wearing a mask for one hour on the quantity and quality of tears in participants with normal eye function. In a sample n=104 of participants free of eye
surface disease, an intervention group n=54, mean age 23.8, 14 females and 40 males wore a SM for one hour and had non-invasive tear break-up time (NITBUT), phenol red thread (PRT), osmolarity, tear meniscus height, tear fern (TF, a test of the eye film), and tear evaporation rate (TER) tests before and after. The control group had the same tests one hour apart. Participants also completed questionnaires to rate dry eye presence. Comparisons of participant self-reports of dry eyes with NITBUT results demonstrated strong correlations (r = 0.590; P < 0.001), NITBUT measurements (r = 0.631; P < 0.001), and the tear ferning (drying a small sample of tear fluid onto a microscope slide) grades (r = 0.517; P < 0.001) before and after wearing the mask. Authors conclude wearing SMs for short periods of time alters tear film contributing to dry eye symptoms.

Summary of findings about primary outcome 2: Unintended psychological and/or developmental consequences of masking

Twelve studies were included that report on the unintended psychological and/or developmental consequences of masking. The characteristics, findings, and assessment of risk of bias for each study are presented in Table 5.

Communication and masking

Key observations

- Masks significantly impair facial emotion recognition, possibly leading to difficulty with expressing and/or recognizing emotional states during brief interactions.
- Misperceptions of expressions of emotions such as surprise, fear, anger, disgust and happiness while masked are common.
- Masks attenuate the accuracy of voice recognition in noisy environments.
- SMs may have a slight edge compared to N95 and cloth in contributing to accuracy of voice recognition.

In response to this question the data from included studies addressed emotion (4) (Grenville & Dwyer, 2022; Gulbetekin et al., pre-print 2022; Langbehn et al., pre-print 2020; Rinck et al., 2022; Shepherd & Rippon, 2023 low ROB) as well as speech/acoustic recognition (4) (Lin et al., in press; Joshi et al., in press; Toscano & Toscano, 2021; Polo and Lã, pre-print 2021 low ROB) across nine studies. These studies are summarized in Table 2 and described below. One of the emotion studies had low ROB, and one of the speech/acoustic recognition studies, a pre-print, had low ROB. The remainder were moderate.

Table 2: Psychological and/or developmental results, emotion and speech/acoustic recognition and masking

<table>
<thead>
<tr>
<th>Citation</th>
<th>Country</th>
<th>n</th>
<th>Mask type</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rinck et al. (2022)</td>
<td>the Netherlands</td>
<td>91</td>
<td>superimposed image of N95</td>
<td>Sig↓: emotion recognition for disgust, fear, sadness, surprise, &amp; happiness for masked faces vs unmasked NS: emotion recognition for anger and neutrality Sig↑: emotion confusion in mask faced vs NM (disgust for anger, surprise for fear)</td>
</tr>
<tr>
<td>Shepherd &amp; Rippon (2023)</td>
<td>UK</td>
<td>199</td>
<td>superimposed image of SM</td>
<td>Sig↓: emotion recognition accuracy of happiness, sadness, fear, surprise and disgust when viewing masked faced vs unmasked NS: emotion recognition accuracy of anger Sig↑: emotion recognition accuracy of happiness vs surprise, surprise vs sadness, and disgust vs fear NS: emotion recognition accuracy between anger and disgust</td>
</tr>
<tr>
<td>Langbehn et al. (pre-</td>
<td>USA</td>
<td>222</td>
<td>SM, N95, &amp; cloth</td>
<td>Target emotion was less perceived in masked versus NM faces Happy and disgust expressions less accurate than anger and surprise</td>
</tr>
</tbody>
</table>
Rinck et al. (2022) (the Netherlands) conducted three non-randomized experimental studies to determine the impact of SMs on the recognition of different facial expressions, and which emotions are mistaken with one another. The emotions examined in each experiment were happiness, sadness, anger, surprise, fear, disgust and neutral displayed by male and female actors from the Radboud Faces Database. Study sample n=91, mean age 33.23 years, 23 males, 67 females, 1 non-binary. The sample was drawn from five countries, four European and one from the USA. Twenty-nine participants reported masking was not mandatory. Overall participants correctly identified the emotions of unmasked faces 87.9% compared to 69.2% of masked faces. The untoward effects of masks varied by emotion. The most significant reduction in emotion recognition was for that of disgust (57%) $P < 0.001$, fear 27% $P < 0.001$, sadness 22% $P < 0.001$, surprise 9% $P < 0.001$, happiness 7% $P < 0.001$. For expressions of anger and neutrality significance was not reached. Regarding mistaking one emotion for another, fear was most often mistaken for disgust, 25% in unmasked and 57% in masked. Similarly anger and disgust were mistaken in the unmasked most often with disgust 7% and sadness 6%, while with masked results were 13% $P < 0.001$, and 5% respectively. Authors conclude masks significantly impair facial emotion recognition resulting in misunderstandings, confusion, and mistaken perceptions in communication.
Shepherd and Rippon (2023) (UK) conducted a RCT to examine the impact of FMs on facial emotion recognition in a sample n=199, of masked n=102, and unmasked n=97 participants, with a mean age of 37.44, and an age range of 18-73. The masked group had n=39 males, mean age 43.51, range 20-69, 63 females, mean age 32.10, range 19-68. The unmasked group had n=39 males, mean age 43.72, range 20-66, and n=58 females, mean age 34.93, age range 18-73. Participants briefly viewed facial expressions from the Radboud Faces Database masked (SM) and unmasked conveying anger, disgust, fear, happiness, sadness, and surprise. Results demonstrated that emotion recognition was significantly decreased when viewing masked faces, compared to unmasked (M = 56.22, SD = 7.20), F(1,197) = 245.06, P < 0.001, ηp 2 = .55). There were also significant differences among types of emotion recognition accuracy. In terms of accuracy of recognition of emotion, the most success occurred with happiness, more so than surprise (M = 10.08, SD = 1.94, P < .001); recognition of surprise was more accurate than for sadness (M = 8.36, SD = 2.73, P < .001); and greater success accuracy in recognizing disgust in comparison with fear (M = 4.73, SD = 3.32, P < 0.001).

There was no significant difference between recognition of anger and disgust. The impact of masks was significant in reducing recognition of happiness, sadness, fear, surprise, and disgust at the level of (P < 0.001). Recognition of anger did not reach significance. Authors suggest that wearing FMs may cause difficulty with expressing and or recognizing emotional states during brief interactions.

In a third study (Langbehn et al., pre-print) (USA) focused on emotion recognition, researchers carried out two studies with Amazon employees on (Mechanical Turk, on-demand micro-task platform) the first investigated perceptions of happiness, disgust, anger, and surprise in a sample n=162, mean age 35.25 and age range of 23-68, 60% male, 68.3% White, 28.3% African American, 3.3% Asian. Participants viewed 168 videos of 14 different actors (3 black females, 4 white females, 4 black males, 3 white males) starting with neutral expressions followed by happiness, disgust, anger, and surprise. The face was fully visible for half of the videos and masked (using either SM, N95, cloth, white rectangle, approximately 40 per mask type) for the remaining. Overall, the target emotion was less perceived in masked versus NM faces, no matter the mask, in others words the degree of happy, disgust, angry and surprised were less evident to participants. Perceiving happy and disgust expressions was less accurate than anger and surprise which are reported to be more expressed in the upper part of the face. In study two, participants viewed videos of reward, affiliation and dominance smiles masked (N95) and unmasked and rated them as to the extent each conveyed positive feelings, reassurance, and superiority. Results demonstrated reward smiles signaled significantly more positive feelings (M = 72.5, SE = 1.99) than reassurance and superiority (M = 59.1, SE = 247. 2.43), t(59) = 5.25, P < 0.001, d = .68). Dominance smiles signaled more superiority (M = 64.2, SE = 2.33) than positive feelings and reassurance (M = 58.2, SE = 2.28), t(59) = 3.43, P = .001, d = .44). Affiliation smiles were not significantly more reassuring than positive feelings and superiority. Results were deemed to be unaffected by political attitudes or gender.

Grenville and Dwyer (2022) (UK) conducted a RCT in a sample n=100 (91 female, 9 male) psychology undergraduate students, mean age 34. 8, white, 15 Asian, 1 black, and 4 mixed; all participants reported normal or corrected-to-normal vision. The aim of the study was to determine the accuracy of facial emotional recognition masked and unmasked. Participants viewed images online of masked (SM) and unmasked faces expressing outcomes of anger, disgust, fear, happiness, sadness, and neutrality. Subsequently they were required to identify the emotion expressed by each face. Overall accuracy was higher without a mask than with. Accuracy was highest for recognizing happiness and sadness of masked faces. Without masks accuracy was lower for anger and fear.

In addition to studies of emotion recognition this review also included four studies about masking and voice alteration. Lin et al. (in press) (China) measured Sound Pressure Level (SPL) Parameters (10) in a sample n=83 of healthy participants (25 males, 28 females) mean age 42.62, range 20-85. Participants presented to a clinical exam room where baseline parameters were gathered by a Laryngologist, subsequently participants donned SMs and measures were repeated. Researchers found SPL, particularly loudness increased significantly while masked P = 0.021 and was not impacted by sex and age. This means the various
dimensions of the voice such as frequency, intensity, perturbation (vocal fold vibration) and aerodynamics increased.

Gulbetekin et al. (2022) (pre-print) (Turkey) studied mask use influence on a) face recognition; b) facial expression recognition; c) interaction with facial expression and race of the other person to influence social distancing in three separate experiments. In experiment 1 the sample was n=102 (80 female, 20 male), with a mean age of 20.4±2.8. Participant accuracy in matching facial stimuli under four conditions was tested: a) both faces unmasked; b) both faces masked; c) sample’s face unmasked and test face masked; d) sample’s face masked and test face unmasked. Accuracy for Caucasian faces (n = 102, M = 0.96±0.08) was higher than for Asian faces (n = 102, M = 0.92±0.09). The mask condition had a significant effect on accuracy F(2, 263.08) = 60.79, P = 0.001 ρ2 = 0.38. Participants demonstrated the highest performance when they were shown an unmasked sample and tested with masked faces. In comparison, they produced the worst performance when shown an unmasked sample and tested with masked faces. In experiments 2 and 3 the sample was n=134 (105 female, 29 male), with a mean age of 21±1.6. In experiment 2, researchers tested the accuracy in correctly matching an emotion to a presented facial expression. The results indicated a significant main effect of mask F(1,133) = 761.96, P = 0.001, ρ2 = 0.85, race F(1,133) = 397.39, P = 0.001 ρ2 = 0.75 and emotion F(1,363.12) = 201.09, P = 0.001 ρ2 = 0.60. Expressions were better recognized on unmasked faces (n = 134, M =0.75±0.08) than masked faces (n = 134, M =0.55±0.09). They were also recognized better on Caucasian faces (n = 134, M =0.71±0.09) than they were on Asian faces (n = 134, M =0.58±0.08). The expressions from the best recognized to the least recognized were happy (n = 134, M =0.84±0.16), neutral (n = 134, M =0.73±0.15), disgust (n = 134, M =0.56±0.10) and fear (n = 134, M =0.45±0.15) respectively. In experiment 3 the researchers looked at preference for amount (meters) of social distance participant would want from presented face/mask/emotion conditions. The results indicated a significant main effect of mask wearing F(1,133) = 67.551, P = 0.001, ρ2 = 0.34 and emotion F(1.48,196.175) = 111.83, P = .001, ρ2 = 0.46. The participants tended to indicate a preference for a wider social distance from unmasked faces (n = 134, M =4.47±1.72) in comparison to masked faces (n = 134, M =3.62±1.49) and preferred greater social distances to faces having an expression of disgust (n = 134, M =4.66±1.59), fear (n = 134, M =4.52±1.56), neutrality (n = 134, M =3.63±1.67) and happiness (n = 134, M =3.37±1.66) respectively.

Joshi et al. (in press) (USA) focused on mask type in relation to six mask conditions (NM, cloth, SM, KN95, SM over KN95, with and without a face shield) and alterations to SPL parameters. In a sample n=19 of 10 females and 9 males the mean age and range were 30.5 years and 18-56 years respectively. Participants voices were assessed and recorded at baseline and considered to have normal voice quality. This was followed by recordings wearing the various masks and at 1-foot and 6-foot distances. Study results found no significant impact of mask type on SPL level parameters of voice intensity, fundamental frequency, Central Peak Prominence (CPP) or formant frequency. There was statistical significance between males and females for intensity measures with males having higher intensity levels.

Similarly, Toscano and Toscano (2021) (USA) investigated the effects of four types of masks (SM, N95, cloth x 2) on speech/voice recognition in low and high levels of background noise. The study sample n= 200 with 73 females and a mean age of 37 working for Amazon’s Mechanical Turk. Recordings were prepared for participants to assess in conditions of low and high noise levels with each of the talkers wearing four different masks. Participants listened to the recordings via headphones and typed what they heard. Masks had little to no effect (5.5% decrease) with low background noise, the SMs having the least effect, however, in areas of high background noise and with N95 and cloth accuracy of speech recognition dropped ranging from 2.8-18.2%.

Polo and Lã (pre-print) (2021) (Spain) investigated self-perceptions of voice-related handicap when masked with Portuguese and Spanish speakers. In a sample of n=558, 297 were Portuguese speakers and 261 Spanish speakers mean age 42.39 with no diagnosed hearing impairment. Male Portuguese participants numbered 97, Spanish 79, 20% from each language group were smokers, most reported wearing SMs
followed by N95. All completed the 10 item Voice Handicap Index (VHI). Results demonstrated both Portuguese and Spanish speakers reported a higher self-perceived voice handicap.

Cognition and masking

Key observations

• Masking did not impact cognitive performance in healthy young adults (18-30) and school children grades 5-7.
• Mild but significant changes in HR and SpO₂ did not attenuate cognitive performance in university students.
• Rating of perceived dyspnea with mask wearing is a possibility.

Four studies examined the effects of masking on cognition, one using a warm-up exercise (Slimani et al., 2021), and three had participants perform cognition related activities (Schlegtendal et al., 2022; Spang and Pieper, 2021 low ROB; Tornero-Aguilera & Clemente-Suarez, 2021 low ROB). Two of the four studies had low ROB and two moderate ROB. One study examined cognitive function via visual attention during maximum aerobic exercise (Slimani et al. 2022).

Slimani et al. (2022) conducted a randomized crossover study to examine the effect of a facemask on cognitive function (visual attention) and RPE during a maximal aerobic fitness test. In a sample of 14 healthy physical education students, participants completed three 20-minute fitness tests, one for orientation, one masked and one unmasked. Results demonstrated that wearing a facemask significantly lowered cognitive performance as participants made more errors in visual attention while masked $P < 0.001$ and changes in RPE did not reach significance.

Slimani et al. (2021) (Tunisia) conducted a RCT to verify the effect of a warm-up protocol on cognitive function with and without FM. In a sample of n=17 healthy physical education students, 9 males, 8 females, mean age 17.6 years, participants were randomized to perform a 15-minute warm-up (run, arm circles, jumping jacks, high knee jogs, back kicking). Following the warmup, participants were given a document to scan and to cross out all letter “d”s. Results of the cognitive exercise demonstrated the warm-up activity improved cognitive performance for both masked and unmasked participants. Therefore, wearing a cloth FM had a positive effect on cognitive function.

Spang and Pieper (2021) (Germany) conducted an RCT to assess cognitive efficacy in relation to masking. In a sample n=45, there were 24 female participants, and the mean age of the sample was 30.3. Most participants were university students and the intervention involved solving mathematical equations while masked and unmasked. Measures included HR, SpO₂, task performance and perceived mental load. None of the parameters showed a significant difference between wearing and not wearing a mask.

Schlegtendal et al. (2022) (Germany) in a RCT studied cognitive performance in school children in grades 5, 6 and 7 with a sample n=133, and 72 were girls. Student participants both intervention (without mask) and control (FM) attended class for two hours in the morning followed by a 15 - minute break, then returned to class where the intervention group removed their masks, controls remained masked, and all students performed a computerized set of cognitive tests. There were no significant differences between groups of masked and unmasked in cognitive performance.

Lastly Tornero-Aguilera and Clemente-Suarez (2021) (Spain) hypothesized that the wearing of a SM in a group of university students would increase the autonomic sympathetic modulation thus decreasing cognitive performance and SpO₂. The sample n=50 had a mean age of 21.2 and 38 males. Student cognitive performance was assessed in a face-to-face class wearing a mask and in a virtual class unmasked. Both lectures were held at 0830 and lasted 150 minutes. Parameters measured before and after lectures included
HR, Heart Rate Variability (HRV), SpO₂, RR, mental fatigue. Results demonstrated significant increase in HR (77.7 ± 18.2 vs. 89.3 ± 11.2 bpm, \( P < 0.001 \) not mask, mask respectively) and SpO₂ significantly lower (98.4 ± 0.5 vs. 96.0 ± 1.8%, mask, not mask respectively) \( P < 0.001 \). Although statistical significance was reached, these results are not clinically significant. Despite these changes they did not significantly effect mental fatigue, reaction time, and RR.

Knowledge gaps and/or methodological gaps in the scientific literature related to masks for COVID-19

- Personal comfort and masking are somewhat at odds. Traditionally masking is associated with healthcare settings and equipment. Mask mandates and equipment for the public were informed by existing healthcare knowledge.
- Studies engaging the public from conception to completion about masking are needed through the lens of people in community settings to further understanding related to perceptions of dyspnea and breathlessness in the absence of physiological alterations.
- Across included studies only biological sex was reported, with one exception (non-binary).
- Optimizing communication while masked requires further study.
- Continued studies on mask fit and comfort are needed.

Breathing is essential to life as is water. Water supply is regulated, air is not nor is masking. Masks go on individual faces and coming up with the material for optimal protection and comfort will not be the obstacle but rather the challenge that one size fits none.
Acknowledgements

To help Canadian decision-makers as they respond to unprecedented challenges related to the COVID-19 pandemic, COVID-END in Canada is preparing evidence syntheses like this one. This living evidence synthesis was commissioned by the Office of the Chief Science Officer, Public Health Agency of Canada. The development and continued updating of this living evidence synthesis has been funded by the Canadian Institutes of Health Research (CIHR) and the Public Health Agency of Canada. The opinions, results, and conclusions are those of the team that prepared the evidence synthesis, and independent of the Government of Canada, CIHR, and the Public Health Agency of Canada. No endorsement by the Government of Canada, Public Health Agency of Canada or CIHR is intended or should be inferred.

References


LES 14.1b: Unintended consequences of masking in response to COVID-19


Langbehn, A., Yermol, D., Zhao, F., Thorstenson, C., & Niedenthal, P. (pre-print). Wearing N95, surgical and cloth face masks compromises the communication of emotion. Research Square. https://doi.org/10.21203/rs.3.rs-133686/v1


LES 14.1b: Unintended consequences of masking in response to COVID-19


LES 14.1b: Unintended consequences of masking in response to COVID-19


Table 3: Summary of studies reporting on the physiological outcomes of masking in response to COVID-19

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
Intervention: SM, flannel mask, and N95 worn during four incremental exercise tests to exhaustion  
Sample: 11 healthy recreationally active participants, 5 men and 6 women, mean age 30 ± 11  
Key outcomes: SpO₂, HR, mean arterial pressure (MAP), stroke volume, cardiac output, RPE and dyspnea, peri-oral end-expired CO₂ and O₂  | • Sig↑ for ratings of dyspnea during submaximal work rates of the ramp test for all mask types (P = 0.007)  
• NS for all other parameters | Moderate       |
Intervention: SM and N95 worn during submaximal and maximal exercise  
Sample: 144 healthy participants, 72 women and 72 men, mean age 29.73  
Key outcomes: HR, BP, SpO₂, Rate Pressure Product (RPP = HR x systolic BP), hematologic (serum) parameters | • NS for all parameters | Moderate       |
## LES 14.1b: Unintended consequences of masking in response to COVID-19

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apolo-Arenas, M. D., Tomas-Carus, P., Galan-Lopez, P., Escribano, J. N., Carvalho, B., Caña-Pino, A., &amp; Parraca, J. A. (2022). <em>The influence of surgical mask on heart rate, muscle saturation of oxygen, and hemoglobin during whole-Body vibration exercise</em>. BioMed Research International, 2022, Article 3958554. <a href="https://doi.org/10.1155/2022/3958554">https://doi.org/10.1155/2022/3958554</a></td>
<td>22 November 2022</td>
<td>Portugal</td>
<td><strong>Design</strong>: Randomized controlled trial&lt;br&gt;<strong>Intervention</strong>: SM worn during whole-body vibration exercise&lt;br&gt;<strong>Sample</strong>: 47 university students, 16 women, 31 men, mean age 19.9 ± 2.0 years&lt;br&gt;<strong>Key outcomes</strong>: HR, muscle oxygen saturation (SatO₂), oxyhemoglobin, deoxyhemoglobin</td>
<td>• NS for all parameters</td>
<td>Moderate</td>
</tr>
<tr>
<td>Bar-On, O., Gendler, Y., Stafler, P., Levine, H., Steuer, G., Shmueli, E., Prais, D., &amp; Mei-Zahav, M. (2021). <em>Effects of wearing facemasks during brisk walks: A COVID-19 dilemma</em>. Journal of the American Board of Family Medicine, 34(4), 798-801. <a href="https://doi.org/10.3122/jabfm.2021.04.200559">https://doi.org/10.3122/jabfm.2021.04.200559</a></td>
<td>26 July 2021</td>
<td>Israel</td>
<td><strong>Design</strong>: Non-randomized experimental study with comparator&lt;br&gt;<strong>Intervention</strong>: SM worn (a) at rest, (b) during a slow walk, and (c) during a brisk walk.&lt;br&gt;<strong>Sample</strong>: 21 participants without underlying cardiopulmonary morbidity, 11 female, 10 male, median age 38, range 29-57 years&lt;br&gt;<strong>Key outcomes</strong>: SpO₂, CO₂, RPE, etCO₂, subjective perception questionnaire</td>
<td>• Small but sig↑ in etCO₂ during the slow walk while wearing a SM vs NM (P = 0.004)&lt;br&gt;• Small but sig↓ in O₂ range while walking with a SM vs (P = 0.08).&lt;br&gt;• In the subjective perception questionnaire, participants described that walking briskly with SM vs NM caused difficulty breathing (86%), shortness of breath (33%), choking feeling (57%), and dizziness (19%).</td>
<td>Moderate</td>
</tr>
<tr>
<td>Cabanillas-Barea, S., Rodríguez-Sanz, J., Carrasco-Uribarren, A., Lópbez-De-celis, C., González-Rueda, V., Zegarra-Chávez, D., Cedeño-Bermúdez, S., &amp; Pérez-Bellmunt, A. (2021).</td>
<td>25 November 2021</td>
<td>Spain</td>
<td><strong>Design</strong>: Randomized controlled trial</td>
<td>• Sig. ↑ perception of dyspnea with the N95 mask and the SM compared to NM</td>
<td>Moderate</td>
</tr>
<tr>
<td>Reference</td>
<td>Date released</td>
<td>Setting</td>
<td>Study characteristics</td>
<td>Summary of key findings in relation to the outcome</td>
<td>JBI critical appraisal/ROB</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------</td>
<td>---------</td>
<td>-----------------------</td>
<td>--------------------------------------------------</td>
<td>--------------------------</td>
</tr>
</tbody>
</table>
**Sample**: 50 healthy volunteers, 26 men, 24 women, mean age 20.96 (SD=3.6)  
**Key outcomes**: distance walked, oxygenometry, HR, sensation of dyspnea, tone of the inspiratory accessory muscles | | |
**Intervention**: SM, N95, and cloth masks worn during a 6-minute walk test  
**Sample**: 29 healthy participants, 19 women, 3 men, mean age of 22  
**Key outcomes**: respiratory symptoms (dyspnea and breathing effort), SpO₂, and functional capacity | • Sig. ↑ in breathing effort for wearing a cloth mask and a N95 mask compared to NM, between SM and N95 mask, and between a SM and cloth mask.  
• Sig. ↑ in dyspnea between wearing a cloth mask and NM | Moderate |
**Intervention**: Wearing a cloth mask during sprint training – 5 | • Wearing a mask may impact subjective feelings of training (e.g., RPE) without necessarily harming sprint performance | Moderate |
### LES 14.1b: Unintended consequences of masking in response to COVID-19

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>in Sports Medicine.</strong> <a href="https://doi.org/10.1080/15438627.2021.2010202">https://doi.org/10.1080/15438627.2021.2010202</a></td>
<td></td>
<td></td>
<td>30 metre sprints on an outdoor track</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sample:</strong> 10 competitive track and field athletes, mean age of 23 ± 4 years, 3 women, 7 men</td>
<td></td>
<td></td>
<td><strong>Sample:</strong> 10 competitive track and field athletes, mean age of 23 ± 4 years, 3 women, 7 men</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Key outcomes:</strong> exercise performance (sprint times, acceleration, jump height), subjective affect (e.g., RPE, pleasure, or discontent)</td>
<td></td>
<td></td>
<td><strong>Key outcomes:</strong> exercise performance (sprint times, acceleration, jump height), subjective affect (e.g., RPE, pleasure, or discontent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Summary of key findings in relation to the outcome</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Summary of key findings in relation to the outcome</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Doherty, C. J., Mann, L. M., Angus, S. A., Chan, J. S., Molgat-Seon, Y., &amp; Dominelli, P. B. (2021). Impact of wearing a surgical and cloth mask during cycle exercise. Applied Physiology, Nutrition and Metabolism, 46(7), 753-762. <a href="https://doi.org/10.1139/apnm-2021-0190">https://doi.org/10.1139/apnm-2021-0190</a></strong></td>
<td>7 May 2021</td>
<td>Canada</td>
<td><strong>Design:</strong> Randomized controlled trial</td>
<td>• Sig. ↑ in dyspnea with wearing a cloth mask vs. NM</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Intervention:</strong> SM and cloth mask worn while doing 3 5-8 minute submaximal cycle tests</td>
<td>• Wearing a mask during short-term moderate-intensity exercise may increase dyspnea but has minimal impact on the cardiopulmonary response</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Sample:</strong> 12 healthy adults, 5 women, 7 men, age 26 (SD=3).</td>
<td>• Differences in respiratory gas pressures were inconsequential during exercise</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Key outcomes:</strong> mask resistance/mouth pressure, HR, SpO₂, breathing frequency, relative respiration, dyspnea, face temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

23
### Reference

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Date released</strong></td>
</tr>
<tr>
<td>13 April 2021</td>
</tr>
<tr>
<td><strong>Setting</strong></td>
</tr>
<tr>
<td>USA</td>
</tr>
<tr>
<td><strong>Study characteristics</strong></td>
</tr>
<tr>
<td><strong>Design:</strong> Randomized controlled trial</td>
</tr>
<tr>
<td><strong>Intervention:</strong> cardiopulmonary exercise tests on a treadmill wearing a cloth mask versus NM, exhaustive incremental exercise (exercise to volitional fatigue)</td>
</tr>
<tr>
<td><strong>Sample:</strong> 31 participants, 14 women, 17 men, median age 23 (range=18-29)</td>
</tr>
<tr>
<td><strong>Key outcomes:</strong> BP, HR, SpO₂, RPE, perception of wearing facemasks</td>
</tr>
<tr>
<td><strong>Summary of key findings in relation to the outcome</strong></td>
</tr>
<tr>
<td>• Wearing a cloth mask led to a sig. reduction in exercise time, maximal oxygen consumption, minute ventilation, maximal HR, beats per minute, and increased dyspnea</td>
</tr>
<tr>
<td>• 30/31 participants agreed or strongly agreed that that it was harder to give maximal effort while they were wearing a cloth mask</td>
</tr>
<tr>
<td><strong>JBI critical appraisal/ROB</strong></td>
</tr>
<tr>
<td>Moderate</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Date released</strong></td>
</tr>
<tr>
<td>6 June 2021</td>
</tr>
<tr>
<td><strong>Setting</strong></td>
</tr>
<tr>
<td>Germany</td>
</tr>
<tr>
<td><strong>Study characteristics</strong></td>
</tr>
<tr>
<td><strong>Design:</strong> Randomized cross-over</td>
</tr>
<tr>
<td><strong>Intervention:</strong> Cycling 50 RPM till exhaustion while wearing SM or FFP2</td>
</tr>
<tr>
<td><strong>Sample:</strong> n=16, healthy male athletes, mean age 27</td>
</tr>
<tr>
<td><strong>Key outcomes:</strong> Max performance, HR, SpO₂, BP, VO₂, VCO₂, RPE</td>
</tr>
<tr>
<td><strong>Summary of key findings in relation to the outcome</strong></td>
</tr>
<tr>
<td>• Statistically significant decrease in maximum performance measured in watts with SM</td>
</tr>
<tr>
<td>• Significance not attained with other parameters</td>
</tr>
<tr>
<td><strong>JBI critical appraisal/ROB</strong></td>
</tr>
<tr>
<td>Moderate</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Setting</strong></td>
</tr>
<tr>
<td>Israel</td>
</tr>
<tr>
<td><strong>Study characteristics</strong></td>
</tr>
<tr>
<td><strong>Design:</strong> Multiple cross-over trial</td>
</tr>
<tr>
<td><strong>Summary of key findings in relation to the outcome</strong></td>
</tr>
<tr>
<td>• Significant increase in etCO₂ while wearing N95 (P = 0.001)</td>
</tr>
<tr>
<td><strong>JBI critical appraisal/ROB</strong></td>
</tr>
<tr>
<td>Moderate</td>
</tr>
</tbody>
</table>
### Reference


**Date released:** March 2023  
**Setting:** Canada  
**Study characteristics:**  
**Intervention:** Cycle ergometry ramp protocol till exhaustion while wearing SM and N95  
**Sample:** n=16 healthy adult males, mean age 34  
**Key outcomes:** HR, SpO₂, etCO₂, BP, RR, RPE  
**Summary of key findings in relation to the outcome:**  
- Significance not reached with other parameters

https://doi.org/10.1249/MSS.0000000000003074

**Date released:** 10 April 2022  
**Setting:** Germany  
**Study characteristics:**  
**Intervention:** Incremental cycling tests over 3 visits while wearing SM  
**Sample:** n=16 healthy young adults, 9 women, mean age 25  
**Key outcomes:** Spirometry (FVC, FEV₁), HR, RR (breaths per minute), dyspnea (Borg 0-10), work rate in watts (W)  
**Summary of key findings in relation to the outcome:**  
- Mean exercise duration (min) significantly reduced while masked (16.4 ± 4.0 vs 15.9 ± 4.0 min, P = 0.02). No significant cardio-respiratory impact  
- Dyspnea unpleasantness significantly greater with FM at both submaximal exercise level and at peak (5.9 ± 1.7 vs 3.9 ± 2.9 Borg 0–10 units, P = 0.007) and at peak exercise (7.8 ± 2.1 vs 5.9 ± 3.4 Borg 0–10 units, P = 0.01)


**Date released:** 10 April 2022  
**Setting:** Germany  
**Study characteristics:**  
**Intervention:** Cycle ergometer 30 seconds warm-up period without load, followed by 1  
**Summary of key findings in relation to the outcome:**  
- Significant reduction in peak power output (W) across groups while masked compared to NM. SM vs NM: (5.0 ± 7.0%, P
LES 14.1b: Unintended consequences of masking in response to COVID-19

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Cardiology, 173, 1-7.</em> <a href="https://doi.org/10.1016/j.amjcard.2022.02.056">https://doi.org/10.1016/j.amjcard.2022.02.056</a></td>
<td>30 September 2021</td>
<td>Japan</td>
<td>min of constant load at 25 W increased every minute for 25 W using a ramp protocol while wearing SM or N95</td>
<td>= 0.005; FFP2 vs NM: 4.7 ± 14%, P = 0.03; control group: SM vs NM: 6.8 ± 4.4%, P = 0.008; FFP2 vs NM: 8.9 ± 6.3%, P = 0.01) • Wearing N95 compared to NM significantly higher pCO₂ CVD group: FFP2: 36.0 ± 3.2 mm Hg vs NM: 33.3 ± 4.4 mm Hg, P = 0.019; control group: FFP2: 32.6 ± 2.8 mm Hg vs NM: 28.1 ± 1.7 mm Hg, P &lt;0.001)</td>
<td>Low</td>
</tr>
<tr>
<td>Fukushi, I., Nakamura, M., &amp; Kuwana, S. (2021). <em>Effects of wearing facemasks on the sensation of exertional dyspnea and exercise capacity in healthy subjects.</em> PLoS One, 16(9), Article e0258104. <a href="https://doi.org/10.1371/journal.pone.0258104">https://doi.org/10.1371/journal.pone.0258104</a></td>
<td>30 September 2021</td>
<td>Japan</td>
<td>Design: Randomized controlled trial</td>
<td>SpO₂ unaffected • Wearing cloth FM significantly worsens dyspnea more so than SM • RPE dyspnea: Control 5 (3.75–6) &amp; SM 6 (5–7.25) (P &lt; 0.05).</td>
<td>Low</td>
</tr>
</tbody>
</table>
### Reference

**during moderate to heavy exercise.** *Journal of Physical Activity & Health, 20*(1), 35-44. [https://doi.org/10.1123/jpah.2022-0145](https://doi.org/10.1123/jpah.2022-0145)


### Table

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
| **during moderate to heavy exercise.** *Journal of Physical Activity & Health, 20*(1), 35-44. [https://doi.org/10.1123/jpah.2022-0145](https://doi.org/10.1123/jpah.2022-0145) | 9 November 2021 | Portugal | treadmill, starting at 4-5km/hour up to a max of 13-14km/hour, until exhaustion | **•** Sig↓ in VO$_2$ and minute ventilation during moderate and severe exercise while wearing a SM ($P < 0.0001$)  
**•** Sig↓ in time to exhaustion while wearing a SM ($P = 0.014$)  
**•** NS in HR and respiratory exchange ratio | Moderate |
Intervention: SM work while using a cycle ergometer cycle for 3-min at 60 W, then the work rate was increased by 15 W min until the participant was unable to continue  
Sample: 32 healthy and active volunteers, 16 female, 16 male, mean age 24 (SD 3.3)  
Key outcomes: SpO$_2$, HR, BP, VO$_2$, minute ventilation, time to exhaustion, respiratory exchange ratio | **•** NS in SpO$_2$ or HR with any of the three masks  
• Wearing a face mask caused additional symptoms such as breathlessness (n=13, 18%) and dizziness (n=7, 10%) | Moderate |
Intervention: Mask (surgical, FFP3, or cloth) worn during four 15-min bouts of moderate-to-high intensity (mostly running, 1 rowing) | **•** NS in SpO$_2$ or HR with any of the three masks  
• Wearing a face mask caused additional symptoms such as breathlessness (n=13, 18%) and dizziness (n=7, 10%) | Moderate |
# LES 14.1b: Unintended consequences of masking in response to COVID-19

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
**Intervention**: SM worn during maximal lactate steady state  
**Sample**: 14 participants (all male), mean age 25.7  
**Key outcomes**: HR, RR, airway resistance | • HR while wearing SM significantly increased 160.1±11.2, $P < 0.01$, compared to NM 154.5±11.4  
• Respiratory rate significantly increased with NM 34.03±7.29, $P = 0.02$ compared to SM 32.09±5.4  
• Surgical face masks increase airway resistance: NM 0.32 ± 0.08 vs SM 0.58 ± 0.16 $P < 0.01$ | Moderate |
| Martin, E., Stefan, O., & Reinhold, K. (pre-print). Effects of wearing face masks under moderate physical effort. *medRixv*. https://doi.org/10.1101/2021.03.18.21253539 | 20 March 2021 (pre-print only) | Austria | **Design**: Randomized controlled trial  
**Intervention**: SM worn while climbing up four floors (96 steps)  
**Sample**: 10 healthy volunteers, 6 female, 4 male, 5 < 30 years old, 5 > 30 years old  
**Key outcomes**: HR and SpO₂ | • NS for all parameters | Moderate |
<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
**Intervention**: KN95 worn during light 45 min of light exercise simulating work in healthcare and related settings  
**Sample**: Eight participants, all male, mean age 35  
**Key outcomes**: perceived dyspnea, thermal comfort, moto-cognitive performance, and rectal, skin, and facial temperature | • Sig↑ in perceived dyspnea for prolonged KN95 use ($P = 0.04$)  
• NS for all other parameters | Moderate |
**Intervention**: Taped filter mask worn during graded exercise test compared to SM and NM  
**Sample**: 8 participants (4 male, 4 female), mean age 24.5  
**Key outcomes**: Perception of breathlessness, HR, blood lactate concentration, $\text{SpO}_2$, maximal workload | • Taped face mask (247±56) significantly reduces maximal workload compared to no face mask (278±56) and SM (269±56), $P < 0.001$  
• Taped face mask (25.6±6.2) significantly decreased time (min) to exhaustion compared to no face mask (29.2±6.6) and SM (27.5±6.3), $P < 0.001$ | Moderate |
| Pasqualetto, M. C., Tuttolomondo, D., Gaibazzi, N., Baratella, M. C., Casolino, P., Stefani, M., Reato, S., Tattan, E., Sorbo, M. D., Bigon, I., Giada, F., Nizzetto, M., Ferrara, C., Galiotto, A., Scevola, M., & Rigo, F. (2022). *Safety of surgical* | June 2022 | Italy | **Design**: Randomized crossover | • Wearing of SM while exercising had no significant effects on parameters measured | Low |
### Reference

**masks during physical activity evaluated with graded cycle ergometry test. The Journal of Sports Medicine and Physical Fitness, 62(6), 846-850.**
https://doi.org/10.23736/S0022-4707.21.12814-2

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
**Sample:** n=32 17 men, mean age 30  
**Key outcomes:** exercise time, max power (W), BP, SpO₂, HR | • Significant increase in RPE during vigorous exercise wearing SM mask: 15.5 ± 1.5 vs. NM: 14.2 ± 2.1, *P* < 0.05  
• Significance not reached across remaining parameters | Low |
**Intervention:** Wearing a face mask during progressive cycle ergometer test | • No detrimental effect on exercise performance (HR, SpO₂ or RPE) while wearing either a non-disposable cloth mask or disposable SM | Low |
### Reference

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
Intervention: Warm-up exercises of 15 minutes with and without cloth facemasks  
Sample: 17 healthy physical education students, 8 females 9 males, mean age of 17.6  
Key outcomes: attention assessment (concentration performance, errors), RPE | • Sig. ↑ rates of concentration performance and rate of perceived exertion for those wearing cloth masks.  
• Wearing a cloth facemask during warm-up may stimulate the cognitive function. | Moderate |
Intervention: Maximal running aerobic fitness test for 20 minutes using cloth masks and NM  
Sample: 14 physical education students, 9 males, 5 females, mean age of 17.5.  
Key outcomes: physical performance, attention | • Lower maximal aerobic speed (MAS), VO_{2max}, and distance covered  
• Significantly greater decreases in concentration performance and more errors | Moderate |
<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
**Intervention:** Wearing a face mask during sub-maximal physical activity  
**Sample:** 39 participants (20 men, 19 women), mean age 38.2±14.2  
**Key outcomes:** physical working capacity, SpO₂, BP, pCO₂, respiratory effort and RPE | • Perceived respiratory effort one point higher (zero-to-ten Likert scale) while using face masks compared to control, \( P < 0.05 \)  
• Differences in physical performance, SpO₂, BP and pCO₂ did not reach significance | Moderate |
**Intervention:** 30 minutes of self-paced exercise on an exercise cycle ergometer, SM and 95 masks  
**Sample:** 19 community-dwelling participants, 9 males, 10 females, age 54-83 who could tolerate exercising for 30 minutes without adverse effects  
**Key outcomes:** pulmonary function; HR, SpO₂, etCO₂, RPE and dyspnea, subjective questions, power output | • HR was significantly higher with the surgical compared to the NM condition  
• N95 was significantly associated with increased RPE, rating of perceived dyspnea compared to NM.  
• petCO₂ was significantly higher in the N95 and SMs compared to NM.  
• petCO₂ was significantly higher in the N95 compared to the SM | Moderate |
**Reference**

<table>
<thead>
<tr>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
<tbody>
<tr>
<td>28 July 2020</td>
<td>China</td>
<td>Design: Randomized controlled trial</td>
<td>• The results suggest CO₂ trapping inside the mask. Wearing an N95 mask may be less comfortable for older adults during prolonged exercise.</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intervention: SM during walking on a treadmill at 4 km/hour for 6 minutes</td>
<td>• Sig. ↑ of HR and RPE while wearing a SM</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample: 23 healthy adults, 10 males (mean age 35), 13 females (mean age 32.7)</td>
<td>• Those wearing SMs reported subjectively higher physiological demands</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Key outcomes: HR, RPE</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


| 1 July 2021  | USA     | Design: Randomized controlled trial | • Sig. ↑ in RPE (breathing discomfort) for all mask types vs. NM | Moderate |
|             |         | Intervention: 60 minutes of walking and jogging between 35% and 60% of relative VO₂max. Four face mask trials (SM; N95; cloth mask and gaiter) and 1 control trial (NM) in the heat (32.3°C ± 0.04°C; 54.4% ± 0.7% relative humidity) | | |
|             |         | • Gaiter and sport masks retained more water vapor and sweat after the exercise compared with surgical and N95 masks | | |
|             |         | • Face mask use is feasible during | | |

### Reference

<table>
<thead>
<tr>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
| 10 September 2021 | China | **Design:** Randomized controlled trial  
**Intervention:** SM influence on changes in cardiopulmonary function during cycle ergometer exercise test  
**Sample:** 12 participants, 8 males, 4 females, mean age 24 | - Mean HR significantly increased with surgical face mask 26.2±14.2 compared to NM 22±12.9, *P* = 0.006  
- Peak exercise VO2 with NM 1628.6±447.2, and with mask 1454.8±418.9, *P* < 0.001  
- RR significantly different with SM 33.8±7.98 compared to NM 37.9±6.72, *P* < 0.001  
- NS difference in SpO2 | Moderate |
| 1 February 2023 | China | **Design:** Non-randomized experimental study with comparator  
**Intervention:** SM worn while walking, performed on a treadmill at a speed of 1.2 m/s | - SMs significantly increased facial discomfort especially when walking (TSV Thermal sensation sig at *P* < 0.05)  
- Wearing a mask in a warm environment | Moderate |
**LES 14.1b: Unintended consequences of masking in response to COVID-19**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Sample:</strong> 42 healthy young adults, 22 female, 20 male, mean age males 25.5, females 24.5</td>
<td>decreases subjective thermal comfort</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Key outcomes:</strong> thermal comfort questionnaire (subjective perception survey)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Sig↓ = Significant decrease; Sig↑ = Significant increase; NS = No significant change*

*BP = blood pressure; CVD = cardiovascular disease; etCO2 = end-tidal carbon dioxide; HR = heart rate; pCO2 = partial pressure of carbon dioxide; SpO2 = oxygen saturation; VO2 = maximal oxygen consumption; VCO2 = carbon dioxide production; RPE = rating of perceived exertion; RR = respiratory rate; RF = respiratory frequency; SM = surgical mask; NM = no mask*
Table 4: Summary of studies reporting on physiological measures of eye surface and masking (no exercise involved)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
**Intervention:** Wearing a face mask versus not wearing a face mask  
**Sample:** 104 participants, 52 in intervention group (14 female, 40 male, mean age 23.8±4.4), 50 in control group (15 female, 35 male, mean age 22.9±4.1)  
**Key outcomes:** Severity of dry eye symptoms using the SPEED questionnaire. NITBUT was recorded as the number of seconds between blinks, appearance of dry spot in the tear film.  
- Control group: no significant differences between variables  
- Significant differences were found between the SPEED scores ($P=0.002$) and the NITBUT ($P<0.001$), before and after wearing a face mask. | | Moderate |
**Intervention:** Multimodal ocular surface evaluation between August 2019 and April 2021  
**Sample:** 274 participants (43.4% male, 56.6% female, mean age 66.15±13.4). Group 1 (before lockdown Aug 2019 – Mar 2020), Group 2 (after lockdown without mask mandate Apr 2020 – Oct 2020), Group 3 (after lockdown with |  
- Mean lipid layer thickness significantly better in group 2 ($P=0.001$) and 3 ($P<0.001$) compared to group 1  
- Schirmer test better in group 3 ($P=0.002$) compared to group 1  
- Tear osmolarity and loss of area to meibomian glands significantly worse in group 2 ($P=0.031$ and $P<0.001$, respectively) and in group 3 ($P<0.001$) compared to group 1 | Low |
<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
|           |               |         | mask mandate Nov 2020 – Apr 2021 | • Blink rate and tear meniscus height significantly worse in group 3 compared to group 1 ($P < 0.001$ and $P = 0.038$, respectively)  
• Non-invasive break up test significantly worse in group 2 ($P = 0.03$) compared to group 1 |               |

**Key outcomes:** Differences in tear film properties (lipid thickness, Schirmer test, tear osmolarity, blink rate, tear meniscus height, non-invasive break-up test)
Table 5: Summary of studies reporting on the psychological and/or developmental outcomes of masking in response to COVID-19

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grenville, E., &amp; Dwyer, D. M. (2022). Face masks have emotion-dependent dissociable effects on accuracy and confidence in identifying facial expressions of emotion. Cognitive Research: Principles and Implications, 7, Article 15. <a href="https://doi.org/10.1186/s41235-022-00366-w">https://doi.org/10.1186/s41235-022-00366-w</a></td>
<td>15 July 2022</td>
<td>United Kingdom</td>
<td>Design: Randomized controlled trial&lt;br&gt;Intervention: NM, a posed mask, or imposed mask&lt;br&gt;Sample: 100 (91 female, 9 male) psychology undergraduate students, mean age 19.5±2.34. 80 white, 15 Asian, 1 black, and 4 mixed; all participants reported normal or corrected-to-normal vision.&lt;br&gt;Key outcomes: Participants’ accuracy and confidence in identifying the emotions (i.e., anger, disgust, fear, happiness, neutral, sadness) portrayed in photographs presented to them of people wearing NM, a posed mask, or imposed mask.</td>
<td>• The accuracy of emotion recognition from faces can become impaired when the lower part of the face is obscured by masks, but the effect is not consistent across all emotions investigated in this study. It was absent or reversed for the emotions of anger, fear, and neutral emotions.&lt;br&gt;• Accuracy was highest for happiness and sadness, there was a clear effect from the masks.&lt;br&gt;• Accuracy was generally the same for the posed mask and imposed mask conditions.&lt;br&gt;• Participants’ confidence in their emotion judgements was higher for stimuli not obscured by masks and was similar between the posed and imposed conditions.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Gulbetekin, E., Fidanci, A., Altun, E., Er, M. N., &amp; Gürcan, E. (pre-print). Effects of mask use and race on face perception, emotion recognition, and social distancing during the COVID-19 pandemic. Research Square. <a href="https://doi.org/10.21203/rs.3.rs-692591/v2">https://doi.org/10.21203/rs.3.rs-692591/v2</a></td>
<td>18 July 2021</td>
<td>Turkey</td>
<td>Design: Randomized controlled trial&lt;br&gt;Intervention: Mask use influence on a) face recognition; b) facial expression recognition; c) interaction with facial expression and race of the</td>
<td>• Experiment 1: Accuracy for Caucasian faces (n = 102, M = 0.96±0.08) was higher than for Asian faces (n = 102, M = 0.92±0.09). The mask condition had a significant effect on accuracy F(2.6, 263.08) = 60.79, P = 0.001 ηp² =</td>
<td>Moderate</td>
</tr>
</tbody>
</table>
Unintended consequences of masking in response to COVID-19

**Sample:** Experiment 1: 102 participants (80 female, 20 male) mean age 20.4±2.8. Experiments 2 & 3: 134 participants (105 female, 29 male) mean age 21±1.6. Undergraduate students.

**Key outcomes:** Experiment 1: Accuracy in matching facial stimuli under four conditions: a) both faces unmasked; b) both faces masked; c) sample’s face unmasked and test face masked; d) sample’s face masked and test face unmasked. Experiment 2: Accuracy in correctly matching an emotion to a presented facial expression. Experiment 3: Preference for amount (meters) of social distance participant would want from presented face/mask/emotion conditions.

0.38. Participants demonstrated highest performance when they were shown an unmasked sample and tested with unmasked faces. In comparison, they produced the worst performance when shown an unmasked sample and tested with masked faces.

- **Experiment 2:** The results indicated a significant main effect of mask F(1,133) = 761.96, P = 0.001, ηp² = 0.85, race F(1,133) = 397.39, P = 0.001 ηp² = 0.75 and emotion F(1,363.12) = 201.09, P = 0.001 ηp² = 0.60. Expressions were better recognized on unmasked faces (n = 134, M =0.75±0.08) than masked faces (n = 134, M =0.55±0.09). They were also recognized better on Caucasian faces (n = 134, M =0.71±0.09) than they were on Asian faces (n = 134, M =0.58±0.08). The expressions from the best recognized to the least recognized were happy (n = 134, M =0.84±0.16), neutral (n = 134, M =0.73±0.15),
**LES 14.1b: Unintended consequences of masking in response to COVID-19**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joshi, A., Procter, T., &amp; Kulesz, P. A. (in press). COVID-19: Acoustic measures of voice in individuals wearing different facemasks. <em>Journal of Voice</em>. <a href="https://doi.org/10.1016/j.jvoice.2021.06.015">https://doi.org/10.1016/j.jvoice.2021.06.015</a></td>
<td>19 June 2021</td>
<td>United States</td>
<td><strong>Design:</strong> Non-randomized experimental study&lt;br&gt;&lt;br&gt;<strong>Intervention:</strong> Face mask influence on measures of voice</td>
<td>- Males (1ft mean = 83.46 dB SPL, SE = 1.34; 6ft mean = 64.13 dB SPL, SE = 1.11) had higher intensity levels than females (1ft mean = 76.23 dB SPL, SE = )</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

• Experiment 3: The results indicated a significant main effect of mask wearing $F(1,133) = 67.551, P = 0.001, \eta^2_p = 0.34$ and emotion $F(1.48,196.175) = 111.83, P = .001, \eta^2_p = 0.46$. The participants tended to indicate a preference for a wider social distance from unmasked faces ($n = 134, M =4.47\pm1.72$) in comparison to masked faces ($n = 134, M =3.62\pm1.49$) and preferred greater social distances to faces having an expression of disgust ($n = 134, M =4.66\pm1.59$), fear ($n = 134, M =4.52\pm1.56$), neutrality ($n = 134, M =3.65\pm1.67$) and happiness ($n = 134, M =3.37\pm1.66$) respectively.
## LES 14.1b: Unintended consequences of masking in response to COVID-19

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
| Langbehn, A., Yermol, D., Zhao, F., Thorstenson, C., & Niedenthal, P. (pre-print). *Wearing N95, surgical and cloth face masks compromises the communication of emotion.* | 23 December 2020 | United States | **Sample:** 19 participants (9 male, 10 female), mean age males 39.4 (21-67), mean age females 30.5 (18-56)  
**Key outcomes:** 1) Intensity level for /a/ at 1ft and 6ft distances from the speaker, 2) F0, F1, F2 for /a/ and /i/, 3) Cepstral peak prominence for /a/ | 1.21; 6ft mean = 58.84 dB SPL, SE = 1.01) at both 1ft and 6ft  
- Intensity while wearing a KN95 + face shield (mean = 81.97, SE = 1.03) was significantly higher than KN95 alone (mean = 78.54, SE = 1.6) for /a/ at 1ft  
- Main effects of sex F(5,17) = 5.18, \( P = 0.005 \) were significant for cepstral peak prominence value for /a/ with males (mean = 14.61, SE = 0.45) having higher cepstral peak prominence values than females (mean = 10.99, SE = 0.43), and KN95 + face shield (mean = 12.17, SE = 0.31) resulting in a lower cepstral peak prominence than KN95 alone (mean = 13.02, SE = 0.39)  
- For F0 and F1 for /a/, only a main effect of sex was significant F(1,17) = 82.93, \( P < 0.001 \); F(1,17) = 34.34, \( P < 0.001 \), respectively | Moderate |

**Design:** Non-randomized experimental study  
**Intervention:** Wearing a variety of masks (N95,
## Reference

**Research Square.**
https://doi.org/10.21203/rs.3.rs-133686/v1

<table>
<thead>
<tr>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 May 2021</td>
<td>China</td>
<td>Surgical, cloth while communicating emotion</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Sample

- **Experiment 1:** 162 participants (55% male, 45% female), mean age 36.46 ± 10.85
- **Experiment 2:** 60 participants (60% male, 40% female), mean age 32.25 ± 9.01

### Key outcomes

- **Experiment 1:** Rating the extent to which faces presented were expressing each emotion during each mask condition
- **Experiment 2:** Rating smiles in terms of the extent to which each conveyed a) positive feelings b) reassurance c) superiority, while faces were masked with an N95

- Significant interaction effect between face presentation (masked versus unmasked) and facial action predominance (upper versus lower face), F(1,158) = 7.17, P = 0.008, np² = 0.043

- **Experiment 2:** Reward smiles were seen as signaling significantly more positive feelings (M = 72.5, SE = 1.99) than reassurance and superiority (M = 59.1, SE = 2.33), t(59) = 5.25, P < 0.001, d = 0.68. Dominance smiles also signaled more superiority (M = 64.2, SE = 2.33) than positive feelings and reassurance (M = 58.2, SE = 2.28), t(59) 248 = 3.43, P = 0.001, d = 0.44. The interaction between face presentation (visible, masked), smile type (reward, affiliation, dominance), and rating scale (target, non-target), F(2,118) = 4.79, P = 0.01, np² = 0.075, was significant

### Design

- Non-randomized experimental study

- Significant increases while wearing masks

### Reference

<table>
<thead>
<tr>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="https://doi.org/10.1016/j.jvoice.2021.04.028">https://doi.org/10.1016/j.jvoice.2021.04.028</a></td>
<td></td>
<td></td>
<td>included sound pressure level (P = 0.021), jitter (P = 0.005), and shimmer (P = 0.002)</td>
</tr>
<tr>
<td>16 August 2021</td>
<td>Spain</td>
<td><strong>Intervention:</strong> Medical masks worn while producing /a/ sound</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• A significant decrease was noted in F3 (P = 0.004) while wearing a face mask</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Significant decrease in maximum phonation time among participants &gt;45yr while wearing masks, and a significant increase in participants &lt;45yr (P = 0.032)</td>
<td></td>
</tr>
</tbody>
</table>

#### Summary of key findings in relation to the outcome

- **Intervention:** Medical masks worn while producing /a/ sound
- **Sample:** 53 participants, 25 male, 28 female, mean age 42.62±14.43
- **Key outcomes:** 1) Acoustic parameters including a) fundamental frequency b) sound pressure level c) percentage of jitter d) percentage of shimmer e) noise to harmonic ratio f) cepstral peak prominence; 2) Formant parameters including a) F1 b) F2 c) F3 d) maximal phonation time

#### JBI critical appraisal/ROB

- Low

---


- **Design:** Non-randomized experimental study
- **Intervention:** Voice handicap index as a function of facemask use in the general working population
- **Sample:** 558 participants (31.5% male, 68.3% female, mean age 42.39±13.9), 261 Portuguese participants (37.2% male, mean age 44.8±15.9) and 297 Spanish participants (26.6% male, mean age 40.3±11.5)
- **Key outcomes:** Voice handicap index scores in

- Both Portuguese and Spanish speakers perceived higher voice-related handicap when using a facemask
- Without facemasks, Spanish speakers perceived significantly higher overall voice-related handicap compared to Portuguese speakers \(P = 0.007\)
- All dimensions of visual handicap index (functional, physical, emotional) were significantly different between the masked and unmasked conditions for
## Reference

**Date released**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
**Intervention:** Face masks worn during emotional expression recognition  
**Sample:** Experiment 1: 91 participants (23 male, 67 female, 1 non-binary), mean age 33.23±10.42. Experiment 2: 89 participants (29 male, 60 female), mean age 38.66±12.09. Experiment 3: 153 participants (36 male, 117 female), mean age 21.61±6.37  
**Key outcomes:** Experiment 1, 2 & 3: a) Recognition of different facial expressions while faces are masked b) Emotion confusion | In all experiments, facial emotion recognition was 20% worse for masked faces than unmasked faces (68% correct versus 88% correct)  
Impairment was largest for disgust, followed by fear, surprise, sadness, and happiness  
Participants frequently confused emotions that share activation of the visible muscles in the upper half of the face  
Participants frequently misinterpreted disgust as anger, fear as surprise, and sadness as neutral | Moderate |
**Intervention:** Face masks worn during cognitive performance tests  
**Sample:** 133 participants total. Group 1 had 65 (Grade 5 = 56.9%, Grade 6 = 23.1%, Grade 7 = 20%, Sport-focused class = 52.3%) and Group 2 had 68 (Grade 5 = | Overall, there were no significant differences in cognitive performance between both groups, masks vs. NMs. Wearing face masks has no significant influence on attention and executive functions of pupils and can still be recommended during school lessons | Moderate |
**LES 14.1b: Unintended consequences of masking in response to COVID-19**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
Intervention: Facial stimuli presented wearing a mask  
Sample: 199 participants (78 male, 121 female), mean age 37.44±6.36. Intervention group (masked faces) had 102 participants. Control group (unmasked faces) had 97 participants  
Key outcomes: Emotion recognition and empathic response | • Significant main effect of “facial covering” indicated that emotion recognition was significantly lower when viewing masked faces (M = 40.33, SD = 6.73) in comparison to unmasked faces (M = 56.22, SD = 7.2), F(1,197) = 245.06, P < 0.001, np² = 0.55  
• Significant main effect of “type of facially expressed emotion” on facial emotion recognition accuracy F(5,985) = 260.59, P < 0.001, np² = 0.57  
• Significant interaction effect between “facial covering” and “type of facially expressed emotion” F(5,985) = 35.9, P < 0.001, np² = 0.15  
• There was a simple main effect of “facial covering” in which viewing masked faces significantly reduced | Low |
LES 14.1b: Unintended consequences of masking in response to COVID-19

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
Intervention: Face mask (FFR) worn while solving basic arithmetic equations (addition, subtraction, multiplication, or division)  
Sample: 45 participants (24 female, 21 male), mean age 30, healthy university students  
Key outcomes: SpO2, HR variability, task performance (ratio correct/all tasks, ratio correct/responses given, mean response time, mean response time of correct responses), perceived mental load | • None of the dimensions showed a significant difference between wearing an FFR and not wearing an FFR.  
• Participants who viewed masked faces reported higher levels of empathic concern (M = 23.35, SD = 3.44) than unmasked faces (M = 22.42, SD = 3.22), t(197) = 1.97, P = 0.05; d = 0.28)  
• Recognition of happiness, t(197) = 7.60, P < 0.00; d = 1.09, sadness, t(197) = 13.8, P < 0.001; d = 1.96, fear, t(197) = 8.2, P < 0.001; d = 1.16, surprised, t(197) = 7.83, P < 0.001; d = 1.11, and disgust, t(197) = 14.18, P < 0.001; d = 2.01 | Low |
## LES 14.1b: Unintended consequences of masking in response to COVID-19

<table>
<thead>
<tr>
<th>Reference</th>
<th>Date released</th>
<th>Setting</th>
<th>Study characteristics</th>
<th>Summary of key findings in relation to the outcome</th>
<th>JBI critical appraisal/ROB</th>
</tr>
</thead>
</table>
Intervention: SM worn during 150-minute lecture  
Sample: 50 participants (38 male, 12 female), mean age 20.2±2.9, university students in biomedical class  
**Key outcomes:** SpO₂, HR, HR variability, mental fatigue, reaction time | • HR was significantly higher while being masked (89.3± 11.1) versus unmasked (77.7±18.2)  
• SpO₂ was significantly lower while being masked (96±1.8) versus unmasked (98.4±0.5) | Low |
Intervention: Face masks (surgical, N95, two cloth) worn during sentence recognition  
Sample: 200 participants (127 male, 73 female), mean age 37  
**Key outcomes:** Speech recognition in variety of mask conditions | • Significant main effect of signal-to-noise-ratio (b = 5.48, SE = 0.21, z = 25.81, P < 0.001), confirming that listeners performed better at the higher signal-to-noise ratio  
• Significant main effect of talker (b = 1.01, SE = 0.2, z = 5.13, P < 0.001), confirming that listeners were more accurate at recognizing speech produced by talker 1 (female) than talker 2 (male)  
• Significant talker x signal-to-noise ratio interaction (b = 0.51, SE = 0.25, z = 2, P = 0.045) | Moderate |

**NM** = no mask; **SM** = surgical mask; **SpO₂** = oxygen saturation; **HR** = heart rate
LES 14.1b: Unintended consequences of masking in response to COVID-19

Figure 1: PRISMA 2020 Flow Diagram

Identification of studies via databases and registers

Records identified from:
- Databases (n = 4963)
  - MEDLINE (n = 1526)
  - Embase (n = 1740)
  - CINAHL (n = 358)
  - iCite (n = 1298)
  - ERIC (n = 41)

Records removed before screening:
- Duplicate records removed (n = 1012)

Identification of studies via other methods

Records identified from:
- Citation searching (n = 24)

Records screened (n = 3951)

Records excluded (n = 3895)

Reports excluded:
- Wrong study design (n = 9)
- Wrong setting (n = 8)
- Wrong outcomes (n = 5)
- Conference abstract only (n = 2)
- Wrong intervention (n = 2)
- Wrong methods (n = 1)

Studies included in review (n = 46)

Appendices

Appendix 1: Detailed search strategy

**Database:** Ovid MEDLINE(R) ALL 1946 to February 22, 2023

**Date:** February 22nd, 2023

**Filters:**


Therapy - MEDLINE. Maximizes sensitivity. In Search Filters for MEDLINE in Ovid Syntax and the PubMed translation. Health Information Research Unit, McMaster University. [https://hiru.mcmaster.ca/hiru/HIRU_Hedges_MEDLINE_Strategies.aspx#Therapy](https://hiru.mcmaster.ca/hiru/HIRU_Hedges_MEDLINE_Strategies.aspx#Therapy)

Line 3 adapted from:


<table>
<thead>
<tr>
<th>#</th>
<th>Searches</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>COVID-19/ or SARS-CoV-2/ or coronavirus infections/ or exp coronavirus/ or exp betacoronavirus/ or (covid 19* or covid19* or covid-19* or covid or severe acute respiratory syndrome coronavirus 2 or severe acute respiratory syndrome corona virus 2 or wuhan coronavirus* or wuhan corona virus* or betacoronavirus* or corona virus* or coronavirus* or ncov* or 2019 ncov* or 19 ncv* or 19ncov or cov or cov2 or hcv-19 or sars2 or sarscov or sars cov 2 or sars-cov-2 or sarscov-2 or sars-cov2 or sars-cov2 or sars-cov2 or sars coronavirus 2 or hcv).ti,ab,kf. (Masks/ or Respiratory Protective Devices/ or (mask or masks or masking or face-mask or facemask or face-masks or face-masks or face masks or face covering or facial covering or mouth covering or face piece* or facepiece* or face protect* or face shield* or faceshield* or respirator or respirators or respiratory protection or respiratory device or respiratory devices or n95 or n 95 or kn95 or kn 95 or kf94 or kf 94 or ffp or ffp1 or ffp 1 or ffp2 or ffp2 or ffp3 or ffp 3 or ffp 3 or n97 or n 97 or n99 or n 99 or p2).ti,ab,kf.) not mechanical.ti,ab,kf.</td>
</tr>
<tr>
<td>3</td>
<td>(ae or co).fs. or (safe or safety or unsafe or complication? or consequence? or react* or harm* or side-effect? or side effect?).ti,ab,kf. or ((adverse* or negative* or undesirable) adj2 (effect? or event? or outcome? or incident? or impac?)).ti,ab,kf. ((physiolog* or physical or biolog* or health* or medical* or function* or psych* or mental* or emotion* or cogniti* or development* or economic* or finance* or environment* or eco*) adj5 (adverse* or negative* or undesirable or harm* or effect? or reaction? or event? or outcome? or complication? or incident? or consequence? or impact? or stress* or strain or anxiet* or anxious or aggravat* or exacerbate*)).ti,ab,kf.</td>
</tr>
<tr>
<td>5</td>
<td>3 or 4 [unintended consequences]</td>
</tr>
<tr>
<td>6</td>
<td>Meta-analysis/ or meta analysis.mp,pt. or review.pt. or search*.tw.</td>
</tr>
<tr>
<td>7</td>
<td>clinical trial.mp. or clinical trial.pt. or random*.mp. or tu.xs.</td>
</tr>
</tbody>
</table>
| 8 | Non-Randomized Controlled Trials as Topic/ or (comparative study or controlled clinical trial).pt. or (quasiexperiment or quasi experiment or quasiexperimental or quasi experimental or quasi-randomized
LES 14.1b: Unintended consequences of masking in response to COVID-19

<table>
<thead>
<tr>
<th>9</th>
<th>cohort studies/ or longitudinal studies/ or follow-up studies/ or prospective studies/ or retrospective studies/ or (cohort or longitudinal or prospective or retrospective).ti,ab,kf.</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Case-Control Studies/ or retrospective studies/ or Control Groups/ or ((case and control) or (cases and controls) or (cases and controlled) or (case and comparison*) or (cases and comparison*) or control group or control groups).ti,ab,kw.</td>
</tr>
<tr>
<td>11</td>
<td>1 and 2 and 5 and 6 [will retrieve Reviews]</td>
</tr>
<tr>
<td>12</td>
<td>1 and 2 and 5 and 7 [will retrieve RCTs]</td>
</tr>
<tr>
<td>13</td>
<td>1 and 2 and 5 and 8 [will retrieve Quasi-experimental studies]</td>
</tr>
<tr>
<td>14</td>
<td>1 and 2 and 5 and 9 [will retrieved Cohort studies]</td>
</tr>
<tr>
<td>15</td>
<td>1 and 2 and 5 and 10</td>
</tr>
<tr>
<td>16</td>
<td>(or/11-15) and English.lg.</td>
</tr>
<tr>
<td>17</td>
<td>16 not (exp Animals/ not (exp Animals/ and exp Humans/))</td>
</tr>
<tr>
<td>18</td>
<td>limit 17 to yr=&quot;2020 -Current&quot;</td>
</tr>
</tbody>
</table>
Appendix 2: Data extraction form

General Information:
- Covidence ID
- Lead author
- Title of article
- Country in which the study was conducted

Methods:
- Aim of study
- Study design
- Methods
- Mask type

Participants:
- Sample description
- Participant age
- Participant gender
- Inclusion criteria
- Exclusion criteria
- Exercise type (exercise studies only)
- Exercise duration (frequency and time) (exercise studies only)

Outcomes:
- Table of outcome measures (exercise studies only):

<table>
<thead>
<tr>
<th></th>
<th>No mask</th>
<th>N95</th>
<th>Surgical mask</th>
<th>Cloth or community mask</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpO₂</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heart rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood pressure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPE/SoB/dyspnea</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respiratory rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO₂</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Unique measures (exercise studies only)
- Key findings
- Significance
- Authors conclusions
Appendix 3: Approach to critical appraisal

JBI Critical Appraisal Tool for Assessment of Risk of Bias for Randomised Controlled Trials (Barker et al., 2023) was used to access RCTs. JBI Checklist for Quasi-Experimental Studies (Non-Randomized Experimental Studies) (Tufanaru et al., 2020) was used to assess non-randomized experimental studies. JBI Checklist for Case Control Studies (Moola et al., 2020) was used to assess case control studies. All studies included for data extraction were evaluated in full.

Quality rankings were assigned according to scores. For RCTs, high risk of bias (ROB) is 0-5/13 points, moderate is 6-9/13 points, and low is 10-13/13 points. For non-randomized experimental studies high ROB is 0-3/9 points, moderate is 4-6/9 points, and low is 7-9/9 points. For case-control studies high ROB is 0-4/10 points, moderate is 5-7/10 points, and low is 8-10/9 points.
Appendix 4: Rating of Perceived Exertion (RPE) and Borg Scale

The consistent use of the Rating of Perceived Exertion (RPE) and/or the Borg scale is noteworthy in the exercising while masked and unmasked across the review studies. Overall, authors did not clearly distinguish the two but rather seemed to use the terms interchangeably. There is some difference between the two. The original Borg scale has a range of 6-20, no exertion to maximum effort and considered to correlate with the person’s heart rate (HR), and self-perception of how hard they are working. Effort was measured by selecting a number between 6-20, adding a 0 which should reflect the HR, therefore a score of 11 would equal a HR of 110. Borg also developed the modified RPE with scores ranging from 0-10, 0 meaning no exertion and 10 meaning maximum. According to Borg "1" on this scale equals lying on a couch and 10 equals pushing a car up a steep hill. This scale corresponds more with breathlessness. Studies in this review frequently used the term Borg dyspnea scale rather than RPE, however, some used both. If comparing the two scales the Borg 12-14 (moderate intensity) corresponds with the RPE (4-5) and Borg 15-17 (vigorous activity) corresponds with the RPE (6-8). The chart below is intended to assist with further understanding and interpreting the use of these scales.

In general studies reporting significant unintended consequences related to breathlessness while exercising and masked (9) used the Borg Modified RPE (0-10) scale, while those reporting significant unintended consequences related to perceived exertion used the Borg (6-20) scale. However, there was some inconsistency in the use of these instruments.

This chart gives you an idea of how these scales and activities compare.

<table>
<thead>
<tr>
<th>Exertion</th>
<th>RPE scale</th>
<th>Borg scale</th>
<th>Activity examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>0</td>
<td>6</td>
<td>laying on the couch</td>
</tr>
<tr>
<td>just noticeable</td>
<td>0.5</td>
<td>7 to 8</td>
<td>bending over to put on your shoes</td>
</tr>
<tr>
<td>very light</td>
<td>1</td>
<td>9 to 10</td>
<td>easy chores, such as doing laundry</td>
</tr>
<tr>
<td>light</td>
<td>2 to 3</td>
<td>11 to 12</td>
<td>leisurely walking that does not increase your heart rate</td>
</tr>
<tr>
<td>moderate/ somewhat hard</td>
<td>4 to 5</td>
<td>13 to 14</td>
<td>brisk walking or moderate activity that speeds up your heart rate without making you out of breath</td>
</tr>
<tr>
<td>hard</td>
<td>6 to 7</td>
<td>15 to 16</td>
<td>vigorous activity, such as jogging, biking, or swimming (increases your heart rate and makes you breathe harder and faster)</td>
</tr>
<tr>
<td>very hard</td>
<td>8 to 9</td>
<td>17 to 18</td>
<td>the highest level of activity that you can continuing doing without stopping, such as running</td>
</tr>
<tr>
<td>maximum effort</td>
<td>10</td>
<td>19 to 20</td>
<td>a short burst of activity, such as a sprint, that you cannot keep doing for long</td>
</tr>
</tbody>
</table>