Summary and discussion

Chapter 7
This is a dissertation on the effects of noise, with a specific focus on aircraft noise. The original broad approach, intended to understand more about noise annoyance and the effects of noise, has resulted in two parts: a subjective part and a physiological part about the effects of noise, also in relation to noise sensitivity.

Several factors that have not received much attention in the past years were studied in this dissertation. Research questions were: do factors like identity and attitudes (main focus of chapter 2) and activity (task versus no-task; main focus of chapter 3) influence noise annoyance, and how? Also the order of conditions (as a methodological factor) was taken into account in this chapter. In chapter 4, annoyance by CDAs was compared with that of regular descents, focusing on duration as a predictive variable for noise annoyance. Below, the results of these three chapters will first be recapitulated, followed by reflections. Then, in a similar way, the results and reflections of the physiological part will be addressed. The main focus of chapter 5 was on mismatch negativity (MMN) responses to noise, followed by chapter 6 on heart rate and heart rate variability (HRV) responses to aircraft noise. After the summary and reflections of the two parts, some suggestions for future research directions are given.

Part I – Subjective responses to noise

Summary

In Chapter 2, three experiments are described that were conducted to identify the influence of the noise source identity on noise annoyance. Participants were engaged in a difficult cognitive task (3-back task, Kirchner (1958)), while listening to four noise samples (45 s) at four different noise levels (55, 65, 75 or 85 ASEL). The noise samples included recorded noise from an A320 aircraft, road traffic noise and, for each of these two samples, a transformed/unidentifiable sample with the same spectral energy and build-up as the original sample. Participants rated their annoyance after every noise sample, and also during a few blocks of silence as a reference. It was hypothesized that participants would judge the original, identifiable samples as more annoying than the transformed samples. The rationale behind this hypothesis was that the transformed samples would be heard and rated without attitudes playing a role (solely on acoustic information), while transportation noise would be recognized and regarded as an unwanted byproduct, and thus would lead to more annoyance. Results showed that the annoyance was indeed higher for the original, identifiable samples, than for the transformed samples.

However, the design of experiment 1 did not control for the potential confounding effects of tonal components, which were filtered out in the transformed samples (and spread over the total sample envelope), but were present in the original samples, including
prominent Doppler effects. Tonal components are known to increase annoyance (Landström et al., 1995; Torija et al., 2008; Vos et al., 2010) and could have explained the drop in annoyance for the transformed samples. A third aircraft sample was therefore introduced in experiment 2. The noise in the sample was still identifiable as aircraft noise (as shown in a pilot with a separate group of participants) but the major tonal components were filtered out. Participants in this experiment performed the same cognitive task while listening to this new sample and both the original and transformed A320 samples. The original samples were rated as most annoying, followed by the identifiable aircraft without major tonal components, while the transformed sample was again rated as least annoying. These results indicate that tonal components indeed had influenced the annoyance ratings, but seemed not to explain all variance between the different samples.

To disentangle whether the tonal components were solely responsible for the differences in annoyance, a third experiment was designed. Experiment 3 essentially was a replication of experiment 2, except for the instructions at the beginning for one of the groups. The first group of participants was oblivious of the origin of the samples (replication of experiment 2), while a second group received full disclosure about the making of the samples, but in a casual way as if it were not part of the instruction. The results of the first (oblivious) group replicated those of experiment 2, whereas participants in the second group rated the transformed sample as most annoying. The results of the third experiment confirm the idea that tonal components were not the only explanatory factor, but that the identity of the noise source was also of influence: when aware that all noise had derived from aircraft, people judged the transformed samples very differently from the consistent findings in the other group and in the previous experiments. We therefore conclude that knowing the identity of the noise source, or in other words recognizing the noise source, is also a factor in explaining noise annoyance in addition to tonal components.

It cannot be ruled out that using a working memory task, such as the 3-back task, altered annoyance ratings relative to what they would have been without the task. Sörvqvist, Stenfelt, and Rönberg (2012) have shown that there is a positive link between early sensory gating and working memory capacity. The higher this capacity, the more the cognitive system is capable of suppressing irrelevant sensory information. Potentially, the fact that working memory was loaded through the task (lowering available working memory capacity), would thus have limited the ability of the participants to filter out the noise. However, even if this were the case, it is not likely to have had much effect on the results, since repeated measures or mixed designs were used in all of our experiments (with equal loading of working memory by the task for all noise samples). Any effect of task would therefore have been orthogonal to those of the manipulations. It is therefore assumed that the differences in annoyance were a result of the manipulations and not of working memory and attentional processes. However, to gain more insight into these processes, a follow-up experiment was conducted to address the potential effects of using a task on the results of an experiment.
Chapter 3 is a methodological follow-up of chapter 2. A task was always performed by all participants in all of the experiments in chapter 2. This in itself could have had an influence on the results. The aim of this study was to find out to what extent activity (task versus no-task) had an effect on noise annoyance and additionally, if the order of noise samples was of influence. Participants listened to both the original A320 and road traffic noise and the transformed twin sounds in randomized order, again at the same four noise levels (55, 65, 75 or 85 ASEL). In this experiment they performed the cognitive task only half of the time (task condition, 20 min), and spent the other half reading a magazine of their choice (no-task condition, 20 min). Annoyance levels were again asked after every noise sample. No effect of the type of activity was found on annoyance. However, participants showed stronger annoyance responses with increasing noise levels in the no-task condition, especially for the transformed samples. It is possible that noise levels became relatively important when the participants were not engaged in a task and had no pre-existing opinion on the noise sample because it was not identifiable. Apart from that, activity did not seem to cast much influence on noise annoyance. Order effects were found however. Noise annoyance was higher in the first condition that people were in, and this effect was strongest for the task condition.

As a result of the findings in chapter 3, additional analyses addressing order were run on the results of the experiments of chapter 2, to ensure that order effects of the sound samples could not have produced a similar order effect. The results of these analyses can be found in the Supplemental Material for chapter 2 (located after the references).

In a study that was more applied than the previous ones, described in Chapter 4, participants rated their annoyance for regular landings (at 2000 ft.) and continuous descent approaches (CDAs, at 3000, 4000 and 5000 ft.) in a virtual environment. The aim of this study was to see whether the CDAs would induce less annoyance than the regular procedures. Against expectation, the CDA at 3000 ft. was rated as more annoying than all other scenarios. So, surprisingly, despite the fact that the \( \text{LA}_{\text{max}} \) level of the 3000 ft. CDA was lower than that of the regular procedure, the annoyance was higher. An explanation for this effect is the duration of the flyover. As can be seen in Figure 4.2, the rise and fall of the noise produced by a regular landing is steeper (due to speed, altitude and angle differences), resulting in a shorter flyover duration. This fact is also visible in the ASEL levels for the regular and the 3000 ft. CDA, which are 79.3 and 79.5 ASEL respectively. The ASEL levels take duration into account, but from these levels equal annoyance ratings would be expected. We therefore think that duration might be a more important factor in explaining annoyance than it has received credit for, as was also supported by earlier findings (Hiramatsu et al., 1978; Zimmer et al., 2008). It seems that duration as a variable needs to be addressed further and is in need for more adequate quantification. For the time being, it seems that ASEL may be a better predictor for annoyance than \( \text{LA}_{\text{max}} \). The regular landing in turn was more annoying than the CDAs at 4000 and 5000 ft., indicating that CDAs do reduce annoyance when still flying.
at higher altitudes. In case of a fixed flight path for CDAs, which is often obligatory for safety reasons, CDAs could result in less noise annoyance in a large area around the airport, but could cause more annoyance in an estimated 2.5 km² directly underneath the flight path, see also Figure 4.4. Considering these findings and the fact that CDAs are less burden for the environment (Clarke et al., 2013; Coppenbarger et al., 2009), we therefore think that the situation should locally be considered before choosing specific procedures, taking local communities underneath the flight path into consideration in addition to safety and additional necessities for flight operations.

Reflections on part I

Theoretical reflections

So how do variables like identifiability, attitude and the activity one is engaged in influence the way we perceive noise? What do these results mean and what can we gain from them? The most likely explanation for the fact that identifiable noise is considered more annoying, lies in the attitude that people already have toward the sound source. When uninformed participants were asked to guess the source of the transformed samples after the experiment, the most frequent answers were ‘sea’ and ‘wind’, i.e. sources that generally have a neutral to happy connotation to them. It is therefore likely that attitudes toward the source are a moderating variable in this process. Furthermore, in the third study of chapter 2, people rated the transformed, unidentifiable sound as least annoying when they were not aware of the source, but as most annoying when they knew it was transformed aircraft noise.

The finding that participants rated identifiable samples as more annoying that transformed ones was replicated four times in chapters 2 and 3. This is therefore consistent, but not in line with earlier findings by Ellermeier et al. (2004), Fidell et al. (2002) and Zeitler et al. (2004). These authors used sound samples that were considerably shorter (2 – 8 s) than was the case in our studies (45 s). Though the difference in length of the samples could have cast their effect, it is well possible that the short everyday noises that were used by Ellermeier et al. (2004) and Zeitler et al. (2004) were considered more neutral or positive by the participants than the transportation noise that was used in our studies.

The findings of chapter 3 may indicate that noise levels became relatively important when the participants were not engaged in a task and had no pre-existing opinion on the noise sample because it was not identifiable. Apart from that, performing a task or not did not seem to cast much influence on noise annoyance. This was not the case for the order of the experimental conditions (methodological factor). Noise annoyance was higher in the first condition that people were in, and this effect was strongest for the task condition. This is could be due to some kind of adaptation process that may run more smoothly when
not having to perform a task, maybe caused by the amount of available attention. Salient was the sharp rise of annoyance when listening to the unidentifiable sample in the no-task condition. Because this noise sample was not identifiable to participants, it is unlikely that the participants had strong associations and attitudes attached to them on the basis of which an annoyance rating can be made. Instead, the exposure level of the sound may have been the dominant determinant of annoyance experienced by the participants in this case. In the challenging task, on the other hand, attention could have been depleted to the extent that an assessment of annoyance caused by the sounds was less precise or accurate. This leads to a whole new discussion: Under what circumstances does annoyance get higher because the noise interferes with an activity? Or can it also be the case that being engaged in an activity can lead the attention away from the annoyance perception, thereby lowering it? My guess would be that it depends on the situation and time, but even more on the person. It would be very interesting to address the parameters involved in this matter. Maybe an attentional and/or physiological approach could shed some light on this, by means of matching annoyance responses to brain activity and/or brain response patterns to specific noise stimuli.

The task influencing noise perception is not the only effect possible in this situation: though it was not the primary interest, a brief look was taken into the effects of noise on cognitive task performance. In the first experiment of chapter 2, it was found that loud noises speeded up the responses on the working memory task. This finding is consistent with earlier findings in the literature on accessory stimuli, which shows that noise can increase arousal leading to faster response times (Hackley & Valle-Inclán, 1999; Han, Liu, Zhang, Jin, & Luo, 2013; Söderlund, Sikström, Loftesnes, & Sonuga-Barke, 2010). However, no effects or correlations of noise, identity and performance were found in any of our later experiments. This suggests that, if there is an effect of noise on task performance, it is not sufficiently strong to be consistently found. This is in line with work by Smith (2012), who writes that working memory tasks are not impaired much by continuous noise. Transportation noise was possibly perceived more as such then we assumed a priori. However, the task does appear to have fulfilled the main purpose it was designed to have: according to informal feedback given afterwards by the participants, the 3-back task indeed stimulated their engagement in the experiments. Although several experiments have shown memory impairment as a result of background noise in both adolescents (Söqvist, 2010) and in adults (Söqvist, Ljungberg, & Ljung, 2010), the fact that no effects of noise on accuracy was found is in line with findings by, for instance, Halin et al. (2014). These authors report that high task difficulty (the 3-back is very difficult) can function as a shield against impairing effects as noise, probably due to the facilitation of selective attention. Noteworthy is that only adults participated in the study.

Surely not only available attention plays a role in these processes. The findings of chapter 1 and 2 suggest that attitudes toward the source may be an important factor when
trying to explain these results. Especially the fact that the annoyance outcomes on the transformed/unidentifiable sample depended this much on the instructions beforehand, suggest that the underlying attitudes that people have, guide them in the way they perceive what they hear. Attitudinal factors, probably highly linked to trust in authorities, fairness and change, have been found to be of influence on annoyance before, as also described in chapter 1 of this dissertation. A lack of trust in the authorities led to higher annoyance levels than would be expected based on the standard exposure-response curves (Schreckenberg, et al., 2017; Schreckenberg et al., 2001), though no such relationship was found in the The Netherlands (Breugelmans et al., 2007). Guski (1999) mentioned that without trust in the authorities, a lot more complaints are to be expected. The problem with trust is that it needs to be earned with years of good behavior1. The factor fairness has received little attention in recent years, though Maris et al. (2007) found that when participants were asked their preference about flight schedules, their noise annoyance was lower, despite the actual schedule that was used. Also a change of situation was found to be related to excess responses of annoyance (higher than could be expected based on exposure-response curves) (Brown & van Kamp, 2009a, 2009b) and even of self-reported physiological symptoms (Hatfield et al., 2001). More studies to confirm or specify the effects of these factors on the attitudes that people develop are needed, especially in case of the perception of fair treatment and honest/better communication.

Another example of the importance of honest and correct communication follows from the results of and reflections on the Continuous Descent Approach (CDA) study in chapter 4. As far as we know, this study was the first in its kind, and the results have direct and indirect implications. A direct implication is that it seems important to not automatically assume that a CDA flyover or any new kind of flight operation is less annoying for the people on the ground. Though I did not find any official reports stating wrong facts, reports can easily be misinterpreted when not read very carefully. A (Dutch report) for instance correctly stated: “Schiphol airport has the intention to change toward the more broad use of noise reducing take-off and landing procedures (p.1). .... The new procedures could cause higher exposure levels directly underneath the flight path. But because of the smaller

1 Regaining trust is not an easy task. It means years of honest communication, it means that sometimes economic incentives should make place for community interests and it entails excellent communication skills in policy makers, airport representatives and strategists. I once witnessed an important politician burst into laughter when a concerned inhabitant asked him to replace the loss of value of his house due to changes around Schiphol Airport. It does not matter how unrealistic the question is, communities should always be approached with respect and compassion or the result will be more anger and more complaints. A positive change in attitudes is not likely to happen as long as people distrust the authorities and feel that they are treated unfairly.
area that is being exposed and the lower population density, a remarkable decline of the amount of annoyed people and exposed dwellings is to be expected” (Commissie voor de milieueffectrapportage, 2016, p. 3, translation by the author). Articles in newspapers were sometimes phrased less opportune, leaving room for wrong conclusions, for instance an article in newspaper Trouw (Sjouwerman, 2007, translation by the author): “Such gliding flights are also executed in The Netherlands at night, but only to reduce nighttime noise annoyance.” This kind of miscommunication, which is much more likely to be read by the public, can lead to all kinds of positive but false expectations about nighttime noise, in turn increasing annoyance when the actual situation is disappointing. As our results show and the report already indicated, there may be a certain area underneath the flight path where people experience more annoyance. This is especially unfortunate if all CDA procedures are required to follow the same flight path, as is often the case. Good and honest communication is therefore vital when dealing with topics as delicate as aircraft noise and annoyance.

As already mentioned in the chapter summary above, a second implication of the CDA study is of a more technical nature: duration may be a more important predictor for noise annoyance than was previously assumed. Previous research on effects of duration on annoyance (of which there is little) has led to contradictory results. Hiramatsu, Takagi, Yamamoto, and Ikeno (1978) found linear increases of annoyance with white noise samples ranging from 30 ms to 90 s. The article is however unclear about whether the results are statistically significant (but judging from the F-values they are unlikely to be). Varying duration times for noise samples were used in a study containing two experiments (Zimmer et al., 2008). Participants performed a memory task while listening to several kinds of noise samples with a duration of 14 s in the first compared to 10 min in the second experiment. Though the longer exposure duration resulted in more annoyance, the authors stress that the amount of interference by the noise should be taken into account when regarding noise annoyance (Zimmer et al., 2008). Though both these experiments and our study, as described in chapter 4, showed annoyance increases with increasing durations, more research is needed to confirm these findings using several other duration lengths.

Duration is not directly taken into account when calculating exposure-response relationships. This could be one of many reasons why exposure-response curves are less well at predicting annoyance than they used to be (Babisch et al., 2009; Brooker, 2009; Schreckenberg et al., 2010). To predict annoyance more accurately in exposure-response curves, duration should also be taken into account in addition to a number of other factors already mentioned above and in chapter 1, such as change, policy, trust in the authorities and fair treatment.
Methodological remarks

Moving on to the methodological remarks following the results of chapter 3: the order effects that were found suggest that there may be some kind of habituation or adaptation effect causing people to report annoyance primarily at the start of the study and less so later on. Earlier studies, however, did not show any adaptation or habituation effects on sleep quality, mood and performance (Öhrström & Björkman, 1988) nor on tendency to focus on the noise and the ability to cope with the noise (Weinstein, 1982). Other studies on habituation and adaptation usually deal with long-term processes, making them less comparable. For instance, Kuroiwa et al. (2002), found that subjective sleep ratings did show habituation to road traffic noise, while the coinciding physiological parameters did not confirm these subjective results. Some habituation for noise-induced sleep problems was found in another study, but not for annoyance (Bluhm, Nordling, & Berglind, 2004). In their recent WHO review on noise interventions and their effects on health, Brown and van Kamp (2017) found no evidence suggesting that adaptation played a part in any of the processes. Question is however, to what extent these long-term processes are related and comparable to the observations that were made in chapter 3 of this dissertation.

Another explanation could be that lack of familiarity with the annoyance scale has contributed to the order effects. It is possible that people have the inclination to report relatively high annoyance levels until they are familiarized with both the scale, the type of noises and the range of noise levels used in the experiment, after which they switch to answers that they think do more justice to the noise. In future research, a study could be conducted in which one participant group spends time to familiarize themselves with the scale first while a second participant group starts the experiment without. Furthermore, this experiment has shown that even in experiments using multiple conditions, settings or situations using counterbalancing techniques, the results should be checked for order effects to avoid potentially wrong conclusions.

For this dissertation, it was specifically chosen not to follow the standard annoyance categories of highly annoyed (HA), annoyed (A) and low annoyed (LA) that are often used in the field. Reason for this choice is that it seems a shame to throw away valuable information by transforming a linear scale into a categorical scale in a research context. Secondly, even though standardized questions to assess noise annoyance were formulated by Fields et al. (2001; ISO/TS 55666:2003), other questions and response options are still used on a regular basis for various reasons. Categorizing from several different scales enables researchers to compare their studies to previous work in the field. A methodological question is however to what extent these comparisons can indeed be made, taking into account that different noise sources and noise levels and quality of noise may have been used. Furthermore, also cultural differences and subsets of participants can influence the outcomes, resulting in categories that may look the same, but are actually not (entirely) comparable. Recently, also Guski (2017), in his review for the new WHO Environmental Noise Guidelines, has expressed his
concerns about these categories, stating that also changes in annoyance over time will be more difficult to detect without the raw scores.

**Part II – Physiological responses to noise**

**Summary**

Acute physiological responses to noise were addressed in part II of this dissertation. Noise sensitivity was given a very important role in this second part, because it is one of the most predictive personal non-acoustic factors explaining noise annoyance. When addressing physiological responses to noise, it is important to be aware that potential biological markers of noise sensitivity could be able to account for a substantial amount of variance. If so, it would be a mistake to analyze the participant pool as a unity, instead of making a division based on noise sensitivity. We thus decided to specifically address the physiological base of noise sensitivity in addition to acute physiological responses to noise.

Firstly, for *chapter 5*, a Mismatch Negativity (MMN) evoking experiment was conducted to learn more about noise sensitivity and its potential biological markers. Though all brain response differences between the sensitivity groups would be interesting, we specifically chose to address the MMN and P3 deflections as outcome measures. The MMN response is a brain deflection that is only present when hearing (or seeing in a visual context) stimuli that are deviant from preceding stimuli. In case the MMN response would be present more often or show larger amplitudes in more sensitive individuals, this could indicate that a more sensitive brain is constantly overwhelmed by new and deviant stimuli. The P3 response, when present, shows that the person is paying attention to a stimulus. We addressed the P3 response because we assumed that high noise sensitive people would have more difficulty to not pay attention to stimuli even though they were asked to in the instructions. EEG measurements were made while participants passively listened to an optimized version of the oddball task (Näätänen et al., 2004) while watching a silent film by Buster Keaton (1922) to distract them from the noises. Participants were selected for a high or low noise sensitivity group based on their scores on the NSS and NoiSeQ questionnaires. We expected to see higher MMN and P3 amplitudes in high noise sensitive people than in low noise sensitive people. However, no differences between the noise sensitivity groups were found on either of these measures.

Acute responses of heart rate (HR) and heart rate variability (HRV) responses to aircraft noise were addressed in *chapter 6*. The role of noise sensitivity in this process was also taken into account in this study. ECG recordings were made in a sound insulated room in three 8-min conditions: baseline (always first), task performance in silence and task performance
with aircraft noise (counterbalanced order). Higher mean heart rates (around 8 bpm (beats per minute) extra), were found in the condition with noise compared to the one without. This is a larger increase than found in previous studies. When looking at the HRV, we found that parasympathetic activity was lower in the noise condition compared to the silence condition. The parasympathetic nervous system manages bodily recovery, repair and build-up. Noise diminished the activity of the parasympathetic nervous system, which indicates that noise exposure inhibits these important recovery processes. We also expected to see increased activity of the sympathetic nervous system, which is in charge of fight, flight and fright responses, as well as general activation of the individual. This is not what we found. There was no increased power of the low frequency band, LF, nor were the sympathovagal balance (LF/HF) or the pre-ejection period (PEP) affected by the noise. It thus seems that stress levels were not increased by the aircraft noise.

Again the participants were divided in two groups based on their noise sensitivity scores, but this time using a median split (not selecting extremes). Even though the resulting two groups indeed did not show extreme ratings on the scale, some interesting physiological differences between the two groups were found nonetheless. High noise sensitive participants exhibited higher heart rates, lower parasympathetic activity and marginally higher levels (trend) on the sympathovagal balance compared to low noise sensitive participants. In addition, low noise sensitives demonstrated lower parasympathetic activity (indicative of body restoration processes) during noise than in the baseline condition, while this was not the case in the sensitive subjects.

Reflections on part II

Theoretical reflections

We were surprised not to find any effects in the MMN study, because in a study that was carried out around the same time (Kliuchko et al., 2016), MMN deflections were found to have bigger amplitudes in low compared to medium and high noise sensitive individuals. Question is why these results were not replicated in the current study. When comparing our stimuli with those of Kliuchko et al. (2016), it is clear that a location, an intensity and a pitch (frequency) deviant were used in both studies, but three different types of deviants were used in their study: noise, pitch slide and rhythm, compared to gap and duration deviant in ours. Both the standard and the deviants were longer in the Kliuchko et al. (2016) study, around 200 ms, and were synthesized piano tones in different pitches. Both studies used a separate time window for each deviant to locate the matching MMN response. In case of Kliuchko et al. (2016), these windows started between 70 – 150 ms after stimulus onset, which means that the MMN response must have already started before the stimulus had ended. These stimulus differences may have led to the difference in results between the studies.
Though the study by Kliuchko et al. (2016) had substantially more statistical power, the results in our study did not suggest that more power would have made a difference. A reason for these findings could be that we encountered a problem in setting the baseline for the event related potentials (ERPs). The 200 ms interstimulus interval that was used to set the baseline for every subsequent trial (following the instructions by Näätänen et al., 2004, who created this paradigm) should in hindsight have been substantially longer. Luck (2005) recommends an interstimulus trial of 2 s or more to prevent late potentials of one trial to influence the next trial. Indeed, a lot of drift was visible in our data as the residue potential of the P3 response was within the timeframe of setting the baseline for subsequent trials. Longer interstimulus intervals could have prevented this problem, but it is also possible that more trials were needed in addition. We chose to use the paradigm that was described by Näätänen et al. (2004) as a successful attempt for an optimal paradigm for the original oddball task. By using a complex stimulus as a standard, each deviant alters just one of the characteristics of the standard, and thereby confirms the standard features that are altered by the other deviants. Because each deviant also functions as a standard in this system, it is possible to strongly reduce the amount of trials that is typically needed for an oddball MMN experiment, rendering the duration to approximately 18 minutes, which sounded very enticing. Although Näätänen et al. (2004) describe successful results using these settings, it seems probable that the data in the current experiment could have been better and less noisy in case we had used significantly more trials and longer inter-stimulus intervals.

The results by Kliuchko et al. (2016) are interesting and suggest a biological substrate for noise sensitivity, but more research is needed to confirm these findings, as our results do not replicate them. While the present MMN study did not confirm a biological base for noise sensitivity, some interesting findings in that direction were found between the sensitivity groups in the HR study in chapter 6. Though higher HR in high noise sensitive people was found before by Stansfeld (1992), marginally higher power (trend) of the sympathovagal balance and lower parasympathetic power compared to low noise sensitives are interesting new findings. Most salient was the finding that this low parasympathetic power in the high noise sensitive group was consistently low, while the low noise sensitive group showed a drop in parasympathetic activity after the baseline condition. This effect suggests that high noise sensitive people are under a constant strain, resulting in attenuating restorative processes of the body. The question is whether noise induces such a strain on these people that the effect lingers on after the noise has stopped, or that these results should be interpreted along the lines of a more generic sensitivity (see for instance van Kamp & Davies, 2013). It seems at least that the experiment was more stressful in general for the high than for the low noise sensitive group. Noteworthy is that in a similar study using road traffic noise, no results were found between the groups on HR and sympathovagal balance (White et al., 2010). What is similar between that study and the one in chapter 6, however, is the unresponsiveness of the HRV in high noise sensitive people which was observed in both studies. It appears that the
HRV in the high noise sensitive group is unresponsive across conditions, while this is not the case in the low noise sensitive group. This is counterintuitive, but could be a marker for a constantly overloaded system. This is something interesting to address in future studies.

It is also possible that the unresponsiveness of the nervous system is a mathematical bias due to the high HR in the high sensitive group. However, in White et al. (2010), the heart rates were generally a bit lower in the high noise sensitive group than observed in this study, while similar results were found. Potential explanations for this difference in HR across the studies are that the participant sample in this study was fairly young, so their hearts may have been more responsive than would be expected from slightly older people. Though aircraft and road traffic noise differ in many ways, the noise levels used in the two studies were fairly similar. It is also possible that the test situation was partly responsible for the unexpected results in the baseline condition, which showed unexpectedly high heart rates, LF and HF power. Speaking to participants after the experiment, it became clear that sitting in an unfamiliar well-insulated room with shut eyes may have been more stressful than we had realized up front. In hindsight, it would have been better to start with a longer period of adjusting to the situation (for instance by introducing a pretest situation), followed by the task conditions and to have ended with the baseline condition. This however, still does not explain the differences between this study and the one by White et al. (2010), as a similar procedure was used there. All in all, it is interesting that these results between the sensitivity groups were found in the present study even though we did not select extremely high or low noise sensitive people. More research is needed to confirm these effects.

Other studies that looked into physiological differences between high and low noise sensitive individuals have addressed several other outcome variables. High noise sensitives showed attenuated filtering processes of incoming stimuli (sensory gating) with a potential overload of information as a result (Shepherd et al., 2016), more active early attentional processes (Kliuchko et al., 2016), higher brain arousal (White et al., 2010) and a combination of higher sympathetic and decreased vagal arousal during noise (Shepherd et al., 2016) compared to low noise sensitive people. Moreover, genetic markers for noise sensitivity were found in a twin study (Heinonen-Guzejev et al., 2005). Though these are several interesting findings suggesting biological markers of noise sensitivity, hardly any of these findings have been replicated - even though noise sensitivity as a topic has received a lot of interest again in the past decade. Question is of course if any studies have been carried out addressing noise sensitivity and biological markers which were not published; it is possible that there is a publication bias concerning this topic, leaving null results (and replication fails) on this topic unpublished, such as the results of chapter 5 currently. It has been said many times before, but if journals do not start accepting manuscripts with well-designed and well-executed studies returning zero results, the research community will continue to waste time and means to address the same questions over and over again. At this point, several more studies are needed to shed light on these processes and to replicate or undermine the state-of-the-art on this topic.
Methodological remarks

As is the case with every kind of research, laboratory experiments have their pros and their biases. Generalizability can be pointed out as the main issue, considering that both the surroundings of the experiment and the participants were not representative for the field and the general population. Measuring mostly students was not merely a choice of convenience. It was deliberately chosen not to target inhabitants of the municipalities close to Amsterdam Airport Schiphol. For the studies on identifiability I specifically wanted people to be honest in the experiments, without a potential urge to ‘make a point’ hoping that the study results would lead to an improvement of their daily situation at home. Furthermore, for the EEG study targeting noise sensitivity, it was necessary to broaden the search criteria for participants. However, the advertisements did not generate as many participants we had hoped for and after participants had filled out the questionnaires, quite a few low noise sensitive people refused to participate in the EEG experiment. Question is to what extent this process may have influenced the results.

Another laboratory bias could have derived from the use of a very quiet, sound-insulated room for all of the experiments described in this dissertation. It is possible that some participants did not (immediately) feel at ease in these quiet surroundings. In chapter 6 (HRV study), this may have led to unusually high heart rates in the baseline condition. Furthermore, in all experiments using aircraft and road traffic noise, it is unclear to what extent participants felt immersed in their environment. It is to be expected that immersion was better in the VR experiment than in the others, but it remains a black box. The reason to still feel confident about the results, is that a within-participant design was used in all the experiments. It is possible that the exact annoyance ratings in the field would have been different, but it can be expected that the differences between the conditions would remain the same.

Due to the increasing popularity of the NoiSeQ (Schütte et al., 2007) in recent years, we have been inconsistent in the use of noise sensitivity scales. For the analyses of the HRV experiment (chapter 6, carried out in the first half of 2014) the Noise Sensitivity Scale (NSS) was used (Weinstein, 1978), while a combination of the NSS and the NoiSeQ was used to select high and low noise sensitive individuals for the MMN experiment (chapter 5, carried out in 2015/2016). The main reason for this inconsistency were discussions during Internoise2014 (which took place in November), where several researchers expressed the idea that the NoiSeQ may be a more reliable and valid instrument than the NSS. Around this time a shift toward the use of the NoiSeQ is also visible in the literature. While both questionnaires were collected in all experiments, we deliberately decided not to analyze the results of chapter 6 a second time with the NoiSeQ because data phishing then would have been too easy. We still plan to write a methodological paper with a re-analysis of all experiments to compare the questionnaires. This article falls outside the scope of this dissertation.
To assess annoyance, we stayed as close as possible to the Dutch version of the standardized question that was proposed in Fields et al. (2001; ISO/TS 55666:2003). Unfortunately, because of a software issue, it was not possible to use the proposed 11-point Likert scale. Instead, we used a 10-point scale in chapters 2 and 3. The proposed 11-point scale is used in chapter 4. As a result, the annoyance scores cannot directly be compared between the experiments and with results in the field using the same question. Most important for the results in this dissertation however, is that the results of chapter 2 and 3 are comparable as they are strongly connected. Furthermore, because the scale is so large, we expect that scores in our experiments will be very close to those had an 11-point scale been used. It is highly likely that other differences between experiments had larger effects on annoyance than the effect of missing a point on the scale. Another potential bias of the use of the standardized question (intended to assess annoyance at home) is the fact that participants were asked to assess how annoying they expected the noise to be in their home situation, while they were sitting in the laboratory room. It is unclear how this mental translation of location has affected the annoyance scores and if this translation differed much between participants. Though the effects between conditions probably were not affected much because of the within subject designs, it is advised not to introduce this type of mental translation if one can avoid it. In other words, a good standardized question is in need for laboratory situations.

Another point, concerning generalizability is air pollution. Noise is rarely the only component affecting health when studying people’s home situations. It is known that air pollution also affects the heart (Pope III et al., 1999; Sinharay et al., 2018). Although some ambient pollution level of NO₂ and particulate matter will surely have been present in the lab rooms that were used, it is likely that levels were lower than outside on the street. In that sense, the findings in the lab are likely to be purer indications of noise effects than findings of field studies that did not control for air pollution. Though air pollution is taken into account in noise field studies more and more, it is a factor to be aware about. Furthermore, it may be time that these factors will be more integrated in research, i.e. studies could focus on annoyance, cognitive and/or health effects in exposed areas, taking all kinds of exposure into account depending on the area.

**Future directions**

With this dissertation, contributions have been made to the field of noise annoyance by aircraft noise in a laboratory setting, with key research questions concerning subjective (part I) and physiological responses to noise (part II). In part I, subjective responses to factors such as the role of source identity and the type of activity that one is engaged in during...
noise exposure and their influences on noise annoyance were studied. Furthermore, the effects of CDAs on noise annoyance was addressed for the first time and condition order was looked at as a methodological factor to take into account when setting up an experiment. In part II, the physiological effects of noise on health outcomes such as brain responses and responses of the heart and nervous system were addressed, taking also noise sensitivity into account. As is usually the case, these research findings have in turn led to new questions. Below, some suggestions for future research are formulated.

Several follow-up studies on the ones presented in this chapter would be worthwhile, for instance addressing the role of attitudes and habituation on noise annoyance. Regarding attitudes, it would be interesting to repeat the last experiment of chapter 2 (with 2 groups, one of which is aware of the production method of the samples) in a few different settings. The experiment could be repeated with different kinds of noise with different connotations to replicate the findings of chapter 2 in a broader perspective. For instance, noises with happy connotations could be used. Furthermore, an intervention study on attitudes could be added to unravel if the current results are indeed mediated by attitudes. This could be done by actively trying to change people’s attitudes about a sound source, using for instance stories and gadgets to affect people. It could be worthwhile if also communication experts would address annoyance topics more often. Bad communication potentially accounts for a lot of unnecessary annoyance. This is in line with one of the conclusions by Brown and van Kamp (2017), in which is stated that policy makers should be informed about change effects that can coincide interventions that are made on the infrastructure.

Other topics that deserve renewed attention are for instance: fear, the difference between continuous and intermittent noise, coping, perceived control, effects of policy and trust in the authorities, feelings of unfair treatment, and identity of the noise source. This list is far from complete, but it seems to me that the topics I have mentioned here may be more important predictors of noise annoyance than they are credited for at this point.

It may not be a new direction, but I think that there is promise within the field of soundscaping. A soundscape can be seen as the total of all heard events in an environment (Schafer, 1977), taking into account all meanings, expectations and emotions that are prompted by the location (Botteldooren et al., 2011). Within the field of soundscaping, researchers and urban designers work together to shape the environment. When attenuation of exposure levels is not a realistic option, then organizing public areas in a smart way can modify the perception of these areas for the better. For instance by adding a water feature such as a fountain, part of the background noise may be masked by the fountain sound (Axelsson, Nilsson, Hellström, & Lundén, 2014), which is considered as pleasant by many people. According to Brown (2012) it is not even about masking, but about dominance of preferred sounds over unwanted sounds. Also visual attributes with a positive connotation can add to the perception of an acoustically more pleasant environment (Lugten, Karacaoglu,
& White, 2017). While soundscaping is not a topic of this dissertation (and without wanting to introduce it in large detail on the last page of it), it is worthwhile to mention that, to the best of my knowledge, the study by Lugten et al. (2017) is the first soundscaping study using aircraft noise. More studies are needed to see if soundscaping interventions can be of use in public spaces with aircraft noise.

Last but not least, vulnerable groups should be taken very seriously. In case of noise, especially children and noise sensitive people deserve to be protected from a human and health point of view.