

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

META PLATFORMS, INC.,
Petitioner,

v.

EIGHT KHZ, LLC,
Patent Owner.

IPR2023-01003
Patent 9,226,090 B1

PATENT OWNER'S RESPONSE

TABLE OF CONTENTS

	Page
I. INTRODUCTION.....	1
A. The '090 Patent.....	1
B. Asserted Grounds.....	2
C. Claim Construction.....	3
D. PHOSITA	4
II. ARGUMENT	4
A. Dr. Kyriakakis's declaration is conclusory and unsupported, fails to disclose the underlying facts or data on which it is based, and is relied upon by Petitioner to supply a limitation missing from the prior art.....	7
1. Dr. Kyriakakis fails to disclose a key underlying fact relied upon in his declaration analyses.	8
2. Contrary to Dr. Kyriakakis's unsupported opinion, Bedingfield discloses neither binaural recording nor binaural synthesis.....	11
3. Contrary to Dr. Kyriakakis's unsupported opinion, Bedingfield does not disclose ITD.	17
4. Contrary to Dr. Kyriakakis's unsupported opinion, Li does not teach a telephone call an/or electronic call involving live voice exchange.	21
5. Dr. Kyriakakis provided contradictory testimony on other topics.....	30
B. Petitioner and Dr. Kyriakakis misapprehend the scope and meaning of the Challenged Claims.....	35
1. Independent Claim 1.	36
2. Independent Claim 8.	43
3. Independent Claim 15.	48
C. Petitioner's obviousness analysis is flawed, because it is based on an incorrect understanding of the Challenged Claims.	49
1. Ground 1: Bedingfield in view of Andrea does not render obvious Independent Claims 1 and 15.....	49
2. Grounds 2 and 3: Bedingfield in view of Andrea and other references does not render obvious Claims 18-19.....	58

3. Ground 4: Bedingfield in view of Ballas does not render obvious Independent Claim 8.	58
4. Ground 5-8: Li in view of other references do not render obvious any of the Challenged Claims	60
III. CONCLUSION.....	62

TABLE OF AUTHORITIES

	Page(s)
Federal Cases	
<i>Amazon.com, Inc. v. Barnesandnoble.com, Inc.</i> , 239 F.3d 1343 (Fed.Cir.2001)	6
<i>Ashland Oil, Inc. v. Delta Resins & Refractories, Inc.</i> , 776 F.2d 281 (Fed.Cir.1985)	7
<i>Aylus Networks, Inc. v. Apple Inc.</i> , No. 2016-1599, 856 F.3d 1353 (Fed.Cir. (Cal.) May 11, 2017)	3
<i>B/E Aerospace, Inc. v. C&D Zodiac, Inc.</i> , 962 F.3d 1373 (Fed.Cir.2020)	6
<i>Google LLC v. Makor Issues & Rights Ltd.</i> , IPR2017-00815, Paper 31 (PTAB Aug. 10, 2018).....	34
<i>Homeland Housewares, LLC v. Whirlpool Corp.</i> , 865 F.3d 1372 (Fed.Cir.2017)	35
<i>Liberty Mutual Insurance Co. v. Progressive Casualty Insurance Co.</i> , CBM2012- 00003	6
<i>In re Magnum Oil Tools Int'l, Ltd.</i> , 829 F.3d 1364 (Fed.Cir.2016)	5
<i>Meta Platforms, Inc. v. Eight kHz, LLC</i> , IPR2023-1004, Paper 9	7
<i>Philip Morris Prods., S.A. v. RAI Strategic Holdings, Inc.</i> , IPR2020-01602, Paper 32, 34 (PTAB. Mar. 30, 2022).....	5
<i>Phillips v. AWH Corp.</i> , 415 F.3d 1303 (Fed.Cir.2005) (<i>en banc</i>)	3
<i>In re Wilson</i> , 424 F.2d 1382 (CCPA 1970)	61

IPR2023-01003
Patent No. 9,226,090

Xerox Corp. v. Bytemark, Inc.,
IPR2022-00624, Paper 9, 5 (PTAB Aug. 24, 2022).....5, 7

EXHIBIT LIST

Ex.	Description
2001	Defendants’ Preliminary Invalidity Contentions
2002	U.S. District Court Order in <i>Eight kHz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575-ADA-DTG (W.D. Tex. Aug. 30, 2023), Granting Joint Motion to Amend Scheduling Order
2003	U.S. District Court Order in <i>Eight kHz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575-ADA-DTG (W.D. Tex. Sep. 11, 2023), Resetting Markman Hearing Via Zoom
2004	Memorandum from Katherine K. Vidal on Interim Procedure for Discretionary Denials in AIA Post-Grant Proceedings with Parallel District Court Litigation (June 21, 2022) (“Interim Procedures”)
2005	Lex Machina, Patent Cases (All Judges) Time to Trial, U.S. District Court for the Western District of Texas, September 9, 2023
2006	Lex Machina, Patent Cases (Judge Alan D. Albright) Time to Trial, U.S. District Court for the Western District of Texas, September 9, 2023
2007	U.S. District Court Minute Entry in <i>Eight kHz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575-ADA-DTG (W.D. Tex. Oct. 3, 2023), Denying Motion to Transfer Venue
2008	U.S. Patent No. 6,961,439 to Ballas (“Ballas 439”)
2009	Defendant’s Final Invalidity Contentions, October 18, 2023
2010	Defendant’s Final Invalidity Contentions, Ex. C-18, October 18, 2023
2011	Declaration of Durand R. Begault, Ph.D.
2012	U.S. District Court e-mail to parties in <i>Eight kHz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575-ADA-DTG (W.D. Tex. Oct. 2, 2023) (providing preliminary constructions in advance of the <i>Markman</i> ’s hearing)
2013	Transcript of the <i>Markman</i> ’s hearing in <i>Eight kHz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575-ADA-DTG (W.D. Tex. Oct. 3, 2023)
2014	C.V. of Durand R. Begault, Ph.D.
2015	Transcript of Deposition of Kyla Alanna McMullen, Ph.D., <i>Eight Khz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575 (W.D. Tex. April 26, 2023)
2016	Transcript of Deposition of Chris Kyriakakis, Ph.D., 03-28-2024.
2017	<i>Meta Platforms, Inc. v. Eight kHz, LLC</i> , IPR2023-1004, Paper 9 (PTAB Jan. 9, 2024) (decision denying institution)
2018	Declaration of Dr. Kyriakakis in Support of Petition for Inter Partes Review of U.S. Patent No. 9,282,196 in <i>Meta Platforms, Inc. v. Eight kHz, LLC</i> , IPR2023-1004.

2019	D. Hammershøi and H. Møller, <i>Binaural Technique—Basic Methods for Recording, Synthesis, and Reproduction</i> , in COMMUNICATION ACOUSTICS; (Blauert, J., Ed.) Springer, Berlin/Heidelberg, German, 2005; pp. 223–254.
2020	Plaintiff’s Sur-Reply Claim Construction Brief in <i>Eight kHz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575-ADA-DTG (W.D. Tex. June 14, 2023)
2021	Petition for Inter Partes Review of U.S. Patent No. 9,282,196 in <i>Meta Platforms, Inc. v. Eight kHz, LLC</i> , IPR2023-1004.
2022	H. Moller, <i>Fundamentals of binaural technology</i> , Appl. Acoust., vol. 36, pp. 171–218, 1992.
2023	Declaration of Dr. Kyriakakis in Support of Petition for Inter Partes Review of U.S. Patent No. 9,674,628 in <i>Meta Platforms, Inc. v. Eight kHz, LLC</i> , IPR2023-1005.
2024	Defendants’ Reply Claim Construction Brief in <i>Eight kHz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575-ADA-DTG (W.D. Tex. May 31, 2023)
2025	J. Blauert, SPATIAL HEARING: THE PSYCHOPHYSICS OF HUMAN SOUND LOCALIZATION, MIT Press, 1997.
2026	J. Breebaart and C. Faller, SPATIAL AUDIO PROCESSING: MPEG SURROUND AND OTHER APPLICATIONS, John Wiley and Sons Ltd.: Chichester, UK, 2007
2027	<i>Meta Platforms, Inc. v. Eight kHz, LLC</i> , IPR2023-1004, Paper 10 (PTAB Feb. 8, 2024) (Request for Rehearing)
2028	ITU-T Recommendation P.76. Determination of Loudness Ratings: Fundamental Principles, International Telecommunication Union, Geneva 1988

I. INTRODUCTION

Petitioner challenges Claims 1-3, 6, 8-10, 12, 14-19 (the “Challenged Claims”) of U.S. Patent No. 9,226,090 (the “’090 Patent”). Paper 1 (“Petition”; Pet.). But, Petitioner’s obviousness analysis is fundamentally flawed, as explained below. Further, the record provides clear evidence that the Challenged Claims are patentably distinct from the prior art, rendering Petitioner’s arguments ineffective. Patent Owner requests that the Board issue a Final Written Decision affirming the patentability of the Challenged Claims.

A. The ’090 Patent

The ’090 Patent aims to improve the telephone call experience between a listener and a talker located remotely by capturing each and every sound wave at the talker’s system, adjusting the same, and then providing the adjusted sound to the listener. Ex.1001, Claim 1. Unlike a conventional telephone call where “sound [heard by the listener] is not comparable in quality to sound that the listener[] would hear if they were communicating face-to-face” with the talker, (*Id.* 1:9-12), the ’090 Patent discloses methods and systems that “emulate the experience of talking face-to-face” for call participants (*Id.* 7:39-41). Ex.2011, ¶61. The ’090 Patent accomplishes this by externalizing the sound captured for the listener via defining a sound localization point (“SLP”) in empty space that is away from and proximate to the listener. *Id.* ¶62. This SLP *simulates an origin of the sound* that the listener

hears during the telephone call. *Id.* As a result of this externalization process, the sound provided to the listener is fully binauralized regardless of whether it was captured binaurally or not. *Id.* ¶61. It is important to note that the externalization process described in the ‘090 Patent claims is entirely guided by manipulation of the auditory cues—the ‘090 Patent externalization process is not guided by visual cues or the listener’s imagination. *Id.* ¶48. Further, the sound captured at the talker’s system includes not only the talker’s voice but also “various sounds in the environment of the talker.” Ex.1001 38:63-64; Ex.2011, ¶¶63-64, 68-69. In essence, the ‘090 Patent uses the externalization process to place the listener in the sound field of the talker, as if they are in the same room talking to each other. Ex.2011, ¶53 (citing Ex.1001 7:62-8:43 for scenarios where John and Paul are talking like they are in the same room).

B. Asserted Grounds

Petitioner asserts the challenged claims are unpatentable under 35 U.S.C. §103 based on eight grounds. Pet. 5-6. Claims 1, 8, and 15 are independent. This response specifically addresses the grounds challenging the independent claims, and their deficiencies.

Petitioner alleges that the Claims 1 and 15 are obvious over Bedingfield in view of Andrea (Ground 1) as well as over Li in view of Andrea (Ground 5).

Petitioner also alleges that Claim 8 is obvious over Bedingfield in view of Ballas (Ground 4) as well as over Li in view of Ballas (Ground 8).

C. Claim Construction

The claim terms should be given their “ordinary and customary meaning” as understood by a PHOSITA in view of the patent claims, specification and prosecution history. *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312-14 (Fed.Cir.2005) (*en banc*); 37 C.F.R. §42.100(b). There is a “heavy presumption” that the ‘090 Patent’s disputed claim terms “are to be given their ordinary and customary meaning.” See *Aylus Networks, Inc. v. Apple Inc.*, No. 2016-1599, 856 F.3d 1353 (Fed.Cir. (Cal.) May 11, 2017). The ordinary meaning may be readily apparent to a PHOSITA, and claim construction can involve merely accepting the meaning of commonly understood words. *Phillips*, 415 F.3d at 1314.

On October 3, 2023, the district court issued preliminary constructions, adopting plain and ordinary meaning for all disputed claim terms. Ex.2012. On October 4, 2023, and upon hearing the parties’ argument at the *Markman* hearing, the constructions remained unchanged. Ex.2013, 40 (“I’m going to stay with the preliminary construction”); 55 (“I’ll stay with the preliminary proposal.”); 75 (“I’m going to stay with the preliminary construction”). Accordingly, no express constructions of the claims are necessary. To the extent that any term needs

construction, the plain and ordinary meaning of the term should be adopted.¹ Ex. 2011, ¶8.

D. PHOSITA

For this proceeding, Patent Owner accepts Petitioner’s definition of PHOSITA. Patent Owner’s expert, Dr. Begault is a PHOSITA. *See* Ex.2011, ¶¶3-4; Ex.2014; Ex.2015, 17:11-21 (providing that Dr. Begault’s work is “definitely ... one of the seminal references for this field”); Ex.2016, 110:16-21 (admitting that Dr. Begault’s “is very high quality”); 112:6-11 (“I think very highly of Dr. Begault’s work.”).

II. ARGUMENT

¹ Petitioner contends that the prior art references render obvious the Challenged Claims “under any reasonable construction.” Pet. 15-16. As will be discussed thoroughly in §II.A.1, Petitioner’s assertion is made due to its expert’s significant departure from the standard of plain and ordinary meaning in his declaration. Accordingly, and as discussed in Patent Owner’s Preliminary Response, (Paper 6, 12-13), the Board should reject Petitioner’s “any reasonable construction” contention, along with any analyses, provided by Petitioner’s expert and relied upon by Petitioner, founded on an interpretation other than the plain and ordinary meaning.

Petitioner's obviousness argument is fundamentally flawed for two reasons. *First*, it pivots on the conclusory and unsupported declaration of Dr. Kyriakakis offered "to supply" key limitations "missing from the prior art." *Xerox Corp. v. Bytemark, Inc.*, IPR2022-00624, Paper 9, 5 (PTAB Aug. 24, 2022) (precedential) (citing *Arendi S.A.R.L. v. Apple, Inc.*, 832 F.3d 1355, 1361-62 (Fed.Cir.2016)). Petitioner relies solely on conclusory statements from Dr. Kyriakakis. A "petitioner cannot employ mere conclusory statements. The petitioner must instead articulate specific reasoning, based on evidence of record, to support the legal conclusion of obviousness." *In re Magnum Oil Tools Int'l, Ltd.*, 829 F.3d 1364, 1380 (Fed.Cir.2016). Further, as discussed below, his credibility is compromised by contradictory testimony presented in his testimony. *Philip Morris Prods., S.A. v. RAI Strategic Holdings, Inc.*, IPR2020-01602, Paper 32, 34 (PTAB. Mar. 30, 2022) ("[C]onflicting testimony on cross-examination would be highly relevant to both the Board's analysis of the specific issues on which the expert gave inconsistent testimony and to the Board's overall view of his credibility and any such inconsistencies would likely bear on the overall credibility of the expert." (citing *Ultratec, Inc. v. CaptionCall LLC*, 872 F.3d 1267, 1274 (Fed.Cir.2017))).

Second, and more important, Petitioner's limitation-by-limitation analysis deviates from proper obviousness inquiry as prescribed by law, and fails to prove obviousness of the Challenged Claims by a preponderance of the evidence.

Petitioner's limitation-by-limitation analysis does not follow the analysis prescribed by the courts. Determining the correct meaning and scope of claims is a prerequisite for any obviousness inquiry. *Amazon.com, Inc. v. Barnesandnoble.com, Inc.*, 239 F.3d 1343, 1351 (Fed.Cir.2001) ("Only when a claim is properly understood can a determination be made ... whether the prior art ... renders obvious the claimed invention."). Once scope of the claim is determined, the actual obviousness analysis begins. This requires determining "the scope and content of the prior art; differences between the prior art and the claims at issue ...; the presence or absence of a motivation to combine or modify prior art with a reasonable expectation of success." *B/E Aerospace, Inc. v. C&D Zodiac, Inc.*, 962 F.3d 1373, 1378 (Fed.Cir.2020). The Board has repeatedly stated:

A petitioner who does not state the differences between a challenged claim and the prior art, and relies instead on the Patent Owner and the Board to determine those differences based on the rest of the submission in the petition risks having the corresponding ground of obviousness not included for trial for failing to adequately state a claim for relief.

Liberty Mutual Insurance Co. v. Progressive Casualty Insurance Co., CBM2012-00003, Paper 7 at 2-3 (PTAB Oct. 25, 2012) (precedential).

As discussed herein, Petitioner fails to follow the prescribed approach to obviousness analysis. Petitioner misunderstands the correct scope and meaning of

the Challenged Claims. Further, Petitioner fails to identify the differences between the teachings of prior art references and the Challenged Claims, nor does Petitioner explain properly how a PHOSITA would be motivated to combine prior art references to arrive at the Challenged Claims with a reasonable expectation of success—in spite of the above-mentioned differences. Under the correct limitation-by-limitation analysis, Petitioner’s proposed combinations of various prior art references fail to render obvious the Challenged Claims.

A. Dr. Kyriakakis’s declaration is conclusory and unsupported, fails to disclose the underlying facts or data on which it is based, and is relied upon by Petitioner to supply a limitation missing from the prior art.

Unsupported declarant testimony “is entitled to little or no weight.” 37 C.F.R. §42.65(a); *Ashland Oil, Inc. v. Delta Resins & Refractories, Inc.*, 776 F.2d 281, 294 (Fed.Cir.1985) (“Lack of factual support for expert opinion going to factual determinations, however, may render the testimony of little probative value in a validity determination.”). Such expert testimony “is particularly problematic in cases where, like here, expert testimony is offered not simply to provide a motivation to combine prior-art teachings, but rather to supply a limitation missing from the prior art.” *Xerox Corp.*, IPR2022-00624, Paper 9, 5.

In the decision denying institution in *Meta Platforms, Inc. v. Eight kHz, LLC*, IPR2023-1004, Paper 9 at 22-24; 29-31 (PTAB Jan. 9, 2024) (Ex.2017), the Board

correctly gave little weight to Dr. Kyriakakis's opinion. Specifically, the Board determined that Dr. Kyriakakis's declaration regarding Bedingfield and Li contains "conclusory and unsupported" statements. Ex.2017, 30 ("Dr. Kyriakakis provide[s] no explanation why that would be the case, other than to state that listener 102's headphones are at both ears."); 24 ("[W]e agree with Patent Owner that Dr. Kyriakakis's opinion that he "read[s] Li as disclosing that moving the sound source . . . takes place during a voice exchange between the computer program and a user/person" is conclusory and unsupported."). This proceeding involves almost the same references and substantially similar limitations and concepts. Given the similarity between Dr. Kyriakakis's declaration in this proceeding and that in the - 1004 case, the Board should assign no weight to his declaration.

1. Dr. Kyriakakis fails to disclose a key underlying fact relied upon in his declaration analyses.

In his declaration, Dr. Kyriakakis maintains that he did not apply plain and ordinary meaning only for four terms. Ex.1003, ¶82 ("I have applied the plain meaning of the terms and phrases in the claims in my analysis ... and where applicable, have also applied 8KHZ's positions as I understand them."); *see also* ¶23 ("I understand that in this proceeding, any claim term that is not construed should be given its plain and ordinary meaning."). He concludes by reiterating from the Petition that "the prior art ... disclose the subject matter of the challenged claims

under any reasonable construction, including their plain meaning.” *Id.*; *see also* Pet. 15; *see supra* fn.1.

During his deposition, Dr. Kyriakakis revealed his non-reliance on plain and ordinary meaning for a key term: head-related transfer functions (“HRTFs”). Ex.2016, 74:13-19 (“I was given a definition that Eight KHz uses, and so I used that, which is functions that describe transformation of sound that modify the spectrum and time of arrival at a listener’s ears.”); *see also* 76:8-16; 77:19-78:7 (“I was asked to use the definition provided in the claim construction by -- by Eight KHz, and that’s what I used. So as far as -- as far as my declaration is concerned, that is the definition that I’m using.”). **But nowhere in his declaration does Dr. Kyriakakis mention that he uses Patent Owner’s proposed construction for HRTFs.** Instead, he misleadingly creates an impression that he uses Patent Owner’s proposed construction only for the four terms expressly listed in his declaration. *See* Ex.1003, ¶82. In fact, during the deposition Dr. Kyriakakis referred Patent Owner to read his declaration filed for the IPR2023-1005 case—concerning another patent and involving different prior art references—where he states that he uses Patent Owner’s proposed construction for HRTFs.² Ex.2016, 78:8-81:9; 99:2-20.

² During the deposition, Dr. Kyriakakis once again altered his stance, indicating that his analyses in the declaration were not solely reliant on the Patent Owner’s proposed

Dr. Kyriakakis states in his declaration that he applies the plain and ordinary meaning for claim terms that are not construed “as a POSA would have understood those terms and phrases at the relevant time.” Ex.1003, ¶82. Indeed, he admitted that HRTFs is a “commonly used term.” Ex.2016, 83:6-8. But during deposition he repeatedly refused to opine on the PHOSITA’s understanding of HRTFs in the context of the ‘090 Patent. *See, e.g., id.* 82:17-83:3 (“Generally and in the context of certain applications, the answer could be different, and that’s why I’m not comfortable answering the definition.”); 91:18-92:13 (“I don’t know if a POSITA would have an answer to that without context. Where is it being applied, how -- there is many different ways that this can happen and be implemented, and so that is why context is needed.”); *see generally id.* 81:10-99:20. However, he conceded that the plain and ordinary meaning of HRTFs is different from the definition he used for his

claim construction for HRTFs. He asserted that he also considered other factors. Ex.2016, 90:10-27 (“My opinions are based on the definition in paragraph 83 of the ‘628 plus the ancillary material that I talked about in paragraphs 52 and 53 of my ‘090.”). This inconsistent deposition testimony, coupled with misleading declaration regarding the use of plain and ordinary meaning for most terms, casts doubt on the reliability of his obviousness conclusions.

declaration analyses. *Id.* 90:16-91:6 (stating that the plain and ordinary meaning of HRTFs is “a transfer function related to the head”).

Dr. Kyriakakis also fails to disclose in the 1004 IPR declaration—which involved almost the same prior art and analyses as this case—that he does not use the plain and ordinary meaning of HRTFs. In fact, Dr. Kyriakakis’s declaration in that IPR similarly misleads the Board and Patent Owner that he uses the plain and ordinary meaning for HRTFs. Ex.2018 ¶¶21, 74. As detailed in the following Section, Dr. Kyriakakis’s undisclosed departure from plain and ordinary meaning for HRTFs was orchestrated to show that Bedingfield discloses HRTFs, thereby implying binaural synthesis. The Board should give no weight to Dr. Kyriakakis’s declaration to the extent that it addresses the HRTFs topic due to failure to disclose the definition he uses for his conclusions.

2. Contrary to Dr. Kyriakakis’s unsupported opinion, Bedingfield discloses neither binaural recording nor binaural synthesis.

Here, much like the 1004 IPR, Dr. Kyriakakis states “[b]inaural sound can be created in two ways[:]” binaural recording via “using two microphones” and binaural synthesis via “processing non-binaural content” with HRTFs. Ex.1003, ¶¶54-55; *see also* Ex.2018 ¶¶52-53. The Board in the 1004 IPR, reviewing Bedingfield, “agree[d] with Patent Owner that **Bedingfield does not teach either method** [of creating binaural sound].” Ex.2017, 31. This determination should

foreclose Dr. Kyriakakis's argument to the contrary. Nevertheless, Patent Owner proceeds with explaining why Dr. Kyriakakis's opinion that Bedingfield discloses HRTFs is factually incorrect and unsupported. *See* Ex.1003, ¶173.

Dr. Kyriakakis's states that Bedingfield discloses HRTFs because "[a] POSA would have understood that Bedingfield's reference to using relative strength and attenuation of different frequencies (for each direction) of a sound for sound localization is the same as an HRTF." *Id.* Since he could not provide a definition for HRTF, as a PHOSITA would understand it, (*See* §II.A.1), his reference to the PHOSITA's understanding of HRTFs should be suspect. Aside from this, his argument suggests that because Bedingfield discloses some of the sound characteristics modified using HRTFs, it necessarily means HRTFs were used to modify them. Wrong, evidenced by his own admission.

Dr. Kyriakakis's declaration concedes that the term HRTFs encompasses many factors—not just attenuation and relative strength. "Such factors include reverberation, reflection, refraction, attenuation, and even the shape of the listener's ears, head, shoulders, and torso, which as the POSA would have readily recognized, all affect the sound that reaches a listener's ears, depending on the environment." Ex.1003, ¶52. Similarly, during the deposition, he was asked to offer an opinion on the following passage:

In the general definition of the HRTF all linear properties of the sound transmission are included. All proposed descriptors of localization cues, such as inter-aural differences in arrival-time, ITD, in phase, IPD, in level/intensity, ILD/IID as well as monaural cues, are contained in the HRTFs. They can thus be derived from the HRTFs, whereas the opposite is not generally the case.

Ex.2019 240. He agreed fully. Ex.2016, 248:6-249:6 (“But if you just had the -- if you -- what they mean is if you just had -- if you were just given ILD and ITD information, you would not be able to reconstruct the HRTFs because you don’t have the spectral modifications.”). As Dr. Begault explains, this passage, with which Dr. Kyriakakis agrees, proves that “one cannot derive an HRTF from a single descriptor such as attenuation.” Ex.2011, ¶37. Therefore, Dr. Kyriakakis’s conclusory statement that HRTFs are the same as attenuation of different frequencies is contradicted by his own testimony.

Furthermore, no PHOSITA would understand that “using relative strength and attenuation of different frequencies” for sound localization is equivalent to HRTFs. Dr. Kyriakakis is wrong because he omits “a key concept: a binaural HRTF must also include an interaural time difference (“ITD”), the difference in time arrival at the two ears of a sound source wavefront.” Ex.2011, ¶¶34, 38.

Additionally, Dr. Kyriakakis explained the basis for his declaration statement that “using relative strength and attenuation of different frequencies” is the same as

HRTFs. Ex.1003, ¶173. He arrives at this conclusion because he does not use plain and ordinary meaning for HRTFs in the declaration—a key fact his declaration failed to disclose. *See* Ex.2016, 231:13-237:6. Dr. Kyriakakis explained that attenuation of different frequencies means to a PHOSITA that “frequencies are changing between front and back,” and that “when frequencies change, that to a POSITA means there are spectral changes.” *Id.* 233-12-16. He explained that he relied on Patent Owner’s proposed construction for HRTFs as “functions that modify the spectrum” to conclude in the declaration that attenuation of different frequencies is the same as HRTFs. *Id.* 232:22-235:8.

Dr. Kyriakakis’s deposition testimony reveals two problems with his declaration. *First*, his declaration unequivocally states “[a] POSA would have understood” that using attenuation of different frequencies for sound localization is the same as HRTFs. Ex.1003, ¶173. Unlike his deposition, Dr. Kyriakakis’s declaration does not say a PHOSITA would understand that “according to the Eight KHz definition [of HRTFs],” (Ex.2016, 235:1-8), using attenuation of different frequencies for sound localization is the same as HRTFs. Here again, Dr. Kyriakakis misleads both the Board and Patent Owner that his conclusions are entirely based on the PHOSITA’s understanding of HRTFs. *See* Ex.2011, ¶34 (stating a PHOSITA’s understanding of HRTFs). Selectively choosing elements from a PHOSITA’s knowledge, blending them with a definition that Dr. Kyriakakis could not establish

as the PHOSITA's, and to subsequently present a conclusion as if it were a PHOSITA's, is sharp practice. The Board should not condone this.

Second, and more important, Dr. Kyriakakis's repeated assertions that he uses Patent Owner's proposed construction for HRTFs are insincere. He does not; nor does he engage with the substance of Patent Owner's proposed construction. Ex.2011, ¶38. In the district court proceeding, Patent Owner maintained that the term HRTFs must be given its plain and ordinary meaning, which is what the court adopted. Ex.2012. Alternatively, Patent Owner proposed HRTFs' construction as "functions that describe transformations of sound that modify the spectrum and time of arrival at a listener's ears." Ex.2020, 9-10. But he also mischaracterizes Patent Owner's alternative construction to imply that HRTFs only "modify the spectrum." Ex.2016, 234:19-21; 235:5-8 ("[T]ransforming the spectrum ... is what I used to -- as the HRTF definition."), whereas it expressly incorporates the concept of "time of arrival at a listener's ears" into the definition of HRTFs. Ex.2011, ¶38. Further, HRTFs only account for head-related spectral transformation, whereas Bedingfield's disclosure of "attenuation of different frequencies" is not caused by the human head

or the outer ears. Ex.2011, ¶¶39-42. Therefore, Bedingfield’s single reference to attenuation has nothing to do with HRTFs.³ *Id.*, ¶43.

The reality is straightforward: Bedingfield does not even remotely suggest or teach HRTFs or HRTF processing, nor mention HRTFs in its disclosure. *Id.*, ¶43. This fact is implicitly acknowledged in the Board’s decision in the 1004 IPR, where it stated that Bedingfield “does not teach” binaural synthesis. Ex.2017, 31. Dr.

³ Even Petitioner’s claim construction expert in the district court admitted under oath that “HRTFs ... describes **the entire, you know, transformation** [i.e., not just spectral transformations] of sound at lots of different locations around a user.” Ex.2015, 87:10-12. Further, she stated the transformation of sound described by HRTFs is that caused by your head: “your ears on your head, and also your torso, your shoulders. Like, basically, everything from here (indicating) up is a pretty big factor in how the HRTF is formulated.” *Id.* 119:17-22. Therefore, Dr. Kyriakakis’s error is two-fold. First, his distorted interpretation of Patent Owner’s proposed construction that HRTFs only modify or transform the spectrum is patently false, because it omits the time of arrival component. Ex.2011, ¶¶34, 38. Second, he errs by stating that attenuation of different frequencies in Bedingfield is the same as HRTFs, because he fails to recognize that non-head-related attenuation has nothing to do with HRTFs. *Id.*, ¶¶39-42.

Kyriakakis acknowledges that processing sound with HRTFs to generate a “virtual auditory space” for transmission to a listener’s ears via headphones constitutes binaural synthesis. Ex.1003, ¶55. Therefore, when the Board determined that Bedingfield does not teach binaural synthesis, despite its reference to “stereo headphones,” (Bedingfield 13:34-36), it logically follows that Bedingfield does not suggest HRTF processing either.

The Board should assign no weight to Dr. Kyriakakis’s declaration vis-à-vis Bedingfield’s disclosure of HRTFs.

3. Contrary to Dr. Kyriakakis’s unsupported opinion, Bedingfield does not disclose ITD.

Petitioner exclusively relies upon a single paragraph in Dr. Kyriakakis’s declaration to argue that Bedingfield discloses ITD. *See, e.g.*, Pet. 29 (citing Ex.1003, ¶147). Examining that paragraph clarifies that Dr. Kyriakakis’s forms his opinion entirely based on a word search for “delay” in Bedingfield—not ITD or interaural time difference. Without any analysis, Dr. Kyriakakis makes yet another conclusory statement that “a POSA would have readily understood that the ‘delay’ in this context is an ITD ... Bedingfield discloses adjusting an ITD of the sound delivered to the listener.” But Dr. Kyriakakis never explains Bedingfield’s context of using the word “delay.”

The word “delay” appears three times in Bedingfield—all in the context of possible sound effects, none concerning ITD. Ex.2011, ¶¶47-48. *The first* appearance relates to the human ear “detect[ing] placement of a speaker.” Bedingfield 3:35-61. Bedingfield begins with “stereo effect,” a phenomenon via which “[a] sound whose volume is stronger in the left ear appears to be coming from a position to the left of the listener,” as one way of sound placement “to ‘position’ instruments left-to-right across the listener’s virtual ‘sound field.’” *Id.* 3:39-44. Just few lines below, Bedingfield describes another technique for “detect[ing] the position of a signal,” which requires attenuating higher frequencies so sounds appear “behind” a listener. *Id.* 3:49-54. In this context, Bedingfield states “changes to delay and echo may heighten this [sound] effect,” i.e., “giv[ing] the listener the impression that a sound source has moved ‘behind’ the listener.” *Id.* 3:54-59. But, per Dr. Begault, ITD is not useful for such impressions, because ITD is the same for front and rear positions.⁴ Ex.2011, ¶49.

⁴ Also, the fact that Bedingfield references attenuation of high frequencies would be undoubtedly understood by a PHOSITA that this passage (including the reference to delay) does not concern ITD, for “ITD is only relevant as a cue for low frequencies.” Ex.2011, ¶49; *see also* ¶48 (“The time differences in the ITDs of HRTFs are miniscule, relative to echoes[.]”).

Even ignoring Dr. Begault's opinion, Dr. Kyriakakis himself contradicted his declaration statement that "a POSA would have readily understood that the 'delay' in this context is an ITD." When he was asked during the deposition whether an echo can create the exact same impression that Bedingfield describes, Dr. Kyriakakis responded "if you change the response with echo[e]s to correspond to the frequency response of a sound source behind the listener, you would be able to mimic that[,]” i.e., “[to] move the sound behind the listener by changing the delay time ... to what would be perceived as a delay for echo.” Ex.2016, 229:14-230:16. Obviously, a PHOSITA would not understand delay, as used by Bedingfield, to teach ITD, when an echo can produce the very sound effect described by Bedingfield and relied upon by Dr. Kyriakakis.

The second appearance of “delay” is among a laundry list of techniques to “produce” sound effects referring to Bedingfield's Fig. 1. Bedingfield 5:35-41. These effects allow the listener “to receive[] audio from the participants 104, 106, 108 through a communication system 116 in a substantially similar (or approximately similar) manner as if the listener was present between all participants.” Bedingfield 4:52-5:3. Thus, “delay” here merely means a technique for simulating a room with participants for a listener, thereby creating a “perception” that the listener is among the participants, whereas ITD is typically used to adjust binaural sound as a function of listener's head orientation. Ex.2011, ¶49; *see also*

¶¶35-36 (defining ITD as “[a] very strong localization cue is the arrival-time difference between left and right ear signals, which positions the auditory event laterally” and stating that it is also frequency-dependent).

The third appearance of the term “delay” is a reference to digital delay control, a musician’s device “to create special dramatic effects or to simulate room reverberation or outdoor environments (e.g., an echo from a canyon), including in multi-channel loudspeaker playback.” As Dr. Begault explains, a digital delay control cannot capture ITDs, for ITDs can only be captured “in a calibrated HRTF measurement in a laboratory.” Ex.2011, ¶49.

Dr. Kyriakakis makes this stretch argument to support his erroneous interpretation of “delay” in Bedingfield: because Bedingfield “orients the listener (listening person) differently within the sound field 122 depending on (in response to) the listener’s head turning from a first to second position,” Bedingfield discloses “adjusting an ITD of the sound that the listening person hears.” Ex.1003, ¶¶146-47. Petitioner almost verbatim pastes these lines in the Petition. Pet. 29. This reasoning is flawed and unsupported. Ex.2011, ¶¶81-83. Bedingfield tracks movement and orientation of the listener’s head “to change the focus of the listener within the sound field 122.” Bedingfield 12:57-58. But Bedingfield emphasizes that any change in the focus of the listener results in the change in the orientation of the sound field by virtue of a change in stereo amplitude volume differences: “As shown in FIG. 15,

after the listener 1402 changes the focus, the sound field may correspondingly change to reflect the new sound field orientation, *thus providing more balanced left and right sound volume.*” Bedingfield 14:35-38; Ex.2011, ¶¶81-83. Thus, Bedingfield does *not* track the listener’s head to alter ITDs, as Dr. Kyriakakis suggests. Ex.2011, ¶¶82-83.

Dr. Kyriakakis mischaracterizes teachings of Bedingfield to read into Bedingfield ITD processing and binaural synthesis—the Board should discredit Dr. Kyriakakis’s unsupported testimony.

4. Contrary to Dr. Kyriakakis’s unsupported opinion, Li does not teach a telephone call an/or electronic call involving live voice exchange.

Dr. Kyriakakis admitted during the deposition that the call terms of the Challenged Claims necessarily involves “a live voice exchange” between a speaker and listener. Ex.2016, 176:10-177:6 (stating that “*there is live voice exchange* during a telephone call”); 180:11-21 (admitting that “right now we’re exchanging – we’re having a live voice exchange” on the deposition “Zoom” call); *see also* Ex.2011, ¶¶51-53, 55-56. This admission is particularly important, because Dr. Kyriakakis clarifies that not all telephone calls imply a voice exchange. For example, he testified that telephone calls may not involve live voice exchanges if participants are “placed on hold.” *Id.* 174:6-9. A “live voice exchange,” Dr. Kyriakakis admitted, requires “a talking person and a listening person,” and they must “be using their

voices”—otherwise, there is no live voice exchange when “one person stops talking, [and] the call is still happening.” *See id.* 175:12-177:12. Similarly, a PHOSITA would readily understand that a telephone call involving a person listening to a recorded voice does not imply a live voice exchange. Ex.2011, ¶¶54-55. As Dr. Begault’s declaration explains, when a speaker’s voice is simply recorded and played back for a listener over a telephone call (e.g., robocalls), no live voice exchange occurs. Ex.2011, ¶¶54-55.

In his declaration for this case, much like that in the 1004 IPR, Dr. Kyriakakis almost exclusively relies on Li’s passing reference to “conference call” to conclude that Li discloses the call terms of the Challenged Claims. Ex.1003, ¶243 (citing Li, 22:1-5); *see also id.* ¶¶259, 269, 320, 344. Incorrect. The Board in the 1004 IPR, reviewing Li, agreed with Patent Owner that Li’s conference call does not teach a voice exchange:⁵

⁵ In the 1004 IPR, Petitioner and Dr. Kyriakakis argued a computer program’s voice is no more than the voice of a speaker processed by a computer program. Ex.2018 ¶104; Ex.2021 23-24; 57. Accordingly, Petitioner argued that the voice exchange between a user and a computer program is no more than a voice exchange between a speaker whose voice is processed by a computer program and another user. Ex.2021 23-24; 57. The Board assumed (without deciding) that Petitioner’s

Petitioner does not provide any further explanation as to *how a voice exchange would occur during such a “conference call” when modified to use the processed voice of a speaker*—which the remainder of *Li* describes as recorded in advance in sound tracks. Absent additional explanation, *we are not persuaded that Li’s brief reference to a “conference call” is sufficient.*

Ex.2017, 22. The Board agreed with Patent Owner that Dr. Kyriakakis’s opinion was “conclusory and unsupported,” because “the only support for Dr. Kyriakakis’s opinion is the very brief ‘conference call’ excerpt in *Li*, which we find insufficient.” *Id.* 24. As explained above, Dr. Kyriakakis’s admitted that the call terms of the Challenged Claims necessarily involves a live voice exchange. As such, the Board’s determination in the 1004 IPR should foreclose Dr. Kyriakakis’s conclusory and unsupported opinion that *Li*’s isolated reference to conference call teaches the call terms. *See also* Ex.2011, ¶¶88-91.

argument was correct. Ex.2017 20. Thus, to the extent that Petitioner may argue that voice exchange in the ‘090 Patent’s telephone/electronic call is different from the voice exchange at issue in the 1004 IPR, Petitioner is precluded from making argument, for it represented (and the Board assumed) that the voice exchange between a user and a computer program is the same as the voice exchange between a user and a speaker whose voice is processed by a computer program.

Confronted by the Board’s determination in the -1004 case, during the deposition Dr. Kyriakakis attempted a new argument using hindsight—raised in neither his declaration nor the Petition—to buttress his unsupported and conclusory opinion of Li’s conference call. Dr. Kyriakakis drew an arbitrary distinction between Li’s use of recorded and prerecorded, stating that “recorded ... to me [] means live.” Ex.2016, 282:9-22. Indeed, Dr. Kyriakakis’s statement here echoes the similar “conclusory” statement in the 1004 IPR that “he reads Li as disclosing...,” which the Board rejected as “unsupported.” Ex.2017 24 (citing Ex.2018 ¶104).

Particularly, Dr. Kyriakakis self-contradicted when he conceded several of Li’s statements containing the word “recorded” do not imply “live” sound. When he opined on Li’s statement that “[t]he sound tracks recorded by the microphone array system 902 are stored in a memory or a storage device,” (Li 17:20-22), he admitted that the sound “is not live”—for “it is stored” in the memory. Ex.2016, 283:1-284:1. Upon reading the following excerpt from Li, he again admitted that “[o]f course” Li does not disclose live sound despite using the word “recorded” (*Id.* 284:2-21):

The microphone array system 902 can be configured in a computing device 901 with a universal serial bus (USB) interface for applications in 3D sound recording. The multiple channel sound can be saved in one file in a portable device. Using the 3D sound processing application 1602, users can play the recorded audio as a 3D binaural sound or a 3D surround sound.

Li 34:56-63. When confronted with Li's disclosure that "[w]hen a stereo recording is played back, the recording from one microphone goes into the subject's left ear, while the recording from the other microphone is channeled into the subject's right ear," (*Id.* 1:54-58), Dr. Kyriakakis again admitted that "recording" here does not mean "live two-way communication."⁶ Ex.2016, 286:12-19.

Confronted with these examples, Dr. Kyriakakis's new argument evolved from a categorical statement that "recorded ... to me [] means live," (*Id.* 282:9-22), to a more conditional version suggesting that "[recorded] is not live if ... it was stored on a memory device or something of the type and played back from there." *Id.* 289:15-22. However, Dr. Kyriakakis contradicted himself once again. When presented with Li's disclosure that "[t]he configurable 3D sound system disclosed herein allows individual users to record and reproduce 3D sound for playback on their headsets and home theater speaker systems, thereby allowing users to experience immersive 3D sound," (Li 7:33-37), he could not say if the 3D sound

⁶ In its Request for Rehearing filed in the 1004 IPR, Petitioner treats "recorded" and "recording" as synonyms. *Compare* Ex.2027, 1 ("Li distinguishes 'recorded' ... from 'pre-recorded') *with Id.*, 4 ("The Decision erroneously conflates Li's reference to 'recording' with 'pre-recording.'").

disclosed is live or not—even though the sound is recorded for playback. Ex.2016, 289:3-14.

Dr. Kyriakakis’s repeated self-contradictions emphasize that his attempted distinctions between Li’s use of recorded and pre-recorded is nothing more than thinly slicing the issue just enough to support his conclusory and unsupported opinion about Li’s disclosing a voice exchange. But Dr. Kyriakakis (and Petitioner) cannot make their case. Li consistently uses “recorded” or “recording” to describe recorded sound playing back at a later time—not in real-time. *E.g.*, Li 1:54-55 (“When a stereo recording is played back ...”); 18:16-20 (“The 3D sound processing application records ... such that the recorded sound can be processed and reproduced [i.e., played back] to 3D sound.”)⁷; 7:33-37 (“The configurable 3D

⁷ During the deposition, Dr. Kyriakakis claimed that this exact passage “could be happening in realtime ... [so] yes, it can be live.” Ex.2016, 286:20-287:8. Dr. Kyriakakis explained the basis for his claim that “a microphone is currently recording sound ... means it is picking up live sound and delivering it somewhere,” and “microphone on my computer right now is recording my sound and sending it to you.” *Id.* 288:3-17. Dr. Kyriakakis is incorrect because he completely ignores the word “reproduce” in Li’s passage. He knew or should have known that in context of binaural recording the terms “recorded ... and reproduced” cannot possibly mean

sound system ... allows individual users to record and reproduce 3D sound for playback ...”). Indeed, Li identifies the very problem it tries to solve: provide an option for the user to adjust the source locations of the recorded audio “during reproduction,” i.e., playback:

In existing technology, ... the recorded binaural sound cannot be modified or configured during reproduction. Although the existing technologies are able to achieve a few enhancements on the 3D spatial audio experience for a user, they do not provide an option for the user to adjust the source locations and directions of the recorded audio [during reproduction].

live transmission of sound. In his article, titled *Fundamental and Technological Limitations of Immersive Audio Systems*, which is relied upon by Petitioner, (Ex. 1035), Dr. Kyriakakis referenced a work by Professor Møller, titled *Fundamentals of binaural technology* to describe “[h]ead-related binaural recording.” *Id.* 945 (citing Ex.2022). In this work, Professor Møller clarifies that within context of binaural recording, the term “reproduction” holds an opposite meaning to live transmission: “The binaural technique is used to transmit the event to listeners who are not present at the location, or to store it for later reproduction.” Ex.2022 204. Therefore, the use of “reproduced” in this passage of Li cannot reasonably be interpreted to denote live transmission, contrary to Dr. Kyriakakis’s testimony.

Li 2:36-45.

When Dr. Kyriakakis was asked how he formed his opinion that Li's "recorded sound" can mean "live," he offered two responses. First, he pointed to Li's conference call, stating that "you can't do a conference call that he talks about ... prerecorded, so those are live." Ex.2016, 285:2-4; *see also id.* 290:5-9. But this response plainly demonstrates circular reasoning, because it relies on the circular assumption that Li's conference call involves a live voice exchange.⁸

Second, Dr. Kyriakakis pointed to Li's "example of the music ... [where] [m]usicians are playing together." *Id.* 290:5-9. Specifically, Dr. Kyriakakis referred to Li's FIGs. 17 and 19.⁹ *See id.* ("[Li] has a whole figure on [the music example]."). But these figures do not use live sound as Dr. Kyriakakis suggests. For these figures, Li explains how during "a reconstruction [i.e., play back] of a scene of a concert stage at a music concert," the user "can arrange the position of the four musical instruments ... to experience the 3D sound recording of a regular concert." Li

⁸ As noted above, Dr. Kyriakakis himself admitted that some calls may not involve a live voice exchange like when the listener is "placed on hold." *Id.* 174:6-9. Dr. Begault also provides the example of a robocall that may not involve a live voice exchange between a talking person and a listening person.

⁹ These are the only figures in Li related to the music example.

20:53-21:2. Thus, these embodiments show, contrary to Dr. K's argument, that Li's "example of music" does not encompass live sound or transmission of it in "realtime." See Ex.2016, 287:6-8. These examples relate to a reconstruction of a concert scene, and a user listening to a recorded soundtrack.¹⁰

As Dr. Begault explains, Li distinguished between recorded and pre-recorded based on the recording system. Ex.2011, ¶93. Since Li also describes a method for recording sound, it uses the word "recorded" or "recording" to refer to sounds recorded via Li's method. *Id.*, ¶94. But these sounds are stored in memory for later playback, as explained above. *Id.*, ¶¶95-96. Li's 3D sound processing application is capable of processing both recorded and pre-recorded sounds regardless. *Id.*, ¶¶97. Read in context of Li's specification, a PHOSITA would understand that Li's conference call does not involve a live voice exchange. *Id.*, ¶¶98-101. The listener

¹⁰ Dr. Kyriakakis also testified—without support—that Li discloses "music collaboration." Ex.2016, 280:15. Patent Owner cannot locate any passage in Li on this topic. Dr. Kyriakakis relied on this "music collaboration" to reimagine the remainder of Li: "For example, the music collaboration when those are people that are virtually collaborating with each other and exchanging music. They could be talking to each other if it is a singer and somebody else." *Id.* 280:15-19. There is no support in Li warranting Dr. Kyriakakis's speculation, nor did he offer one.

simply listens to the recorded voices of various speakers that are being played back.

Id.

For these reasons, the Board should discredit Dr. Kyriakakis's unsupported and conclusory statement that Li's conference call teaches the call terms of the Challenged Claims.

5. Dr. Kyriakakis provided contradictory testimony on other topics.

As discussed above, Dr. Kyriakakis contradicted himself on key topics:

- Dr. Kyriakakis's statement equating attenuation with HRTFs (Ex.1003, ¶173) is contradicted by his declaration statement (*Id.* ¶52), and his deposition testimony, where he agreed with Hammershøi and Møller that HRTFs cannot be derived solely from attenuation or any other linear properties. *See* II.A.2.
- Dr. Kyriakakis's statement that "a POSA would have readily understood that the 'delay' in this context [i.e., Bedingfield] is an ITD," (Ex.1003, ¶147), is contradicted by his deposition testimony that Bedingfield's reference to delay can be delay in the time of echo. *See* II.A.3.
- Dr. Kyriakakis's deposition testimony that "recorded" means "live," (Ex.2016, 282:9-22), is contradicted by his statement that it is not live if stored and played back. *Id.* 289:15-22. He further contradicted himself by testifying

that Li's statement regarding "record and reproduce 3D sound for playback," (Li 7:33-37), may refer to live sound.

Patent Owner does not intend to highlight other instances where Dr. Kyriakakis contradicted himself. But the following also deserves the Board's attention.

- a. **"Binaural recording does not allow a user to arbitrarily manipulate the spatial characteristics of the recorded sound."**

Dr. Kyriakakis disagreed with this header statement made by Dr. Begault. Ex.2016, 119:13-120:5; *see* Ex.2011, ¶¶28, 31. But he contradicted himself while agreeing with the following paragraph from Hammershøi and Møller:

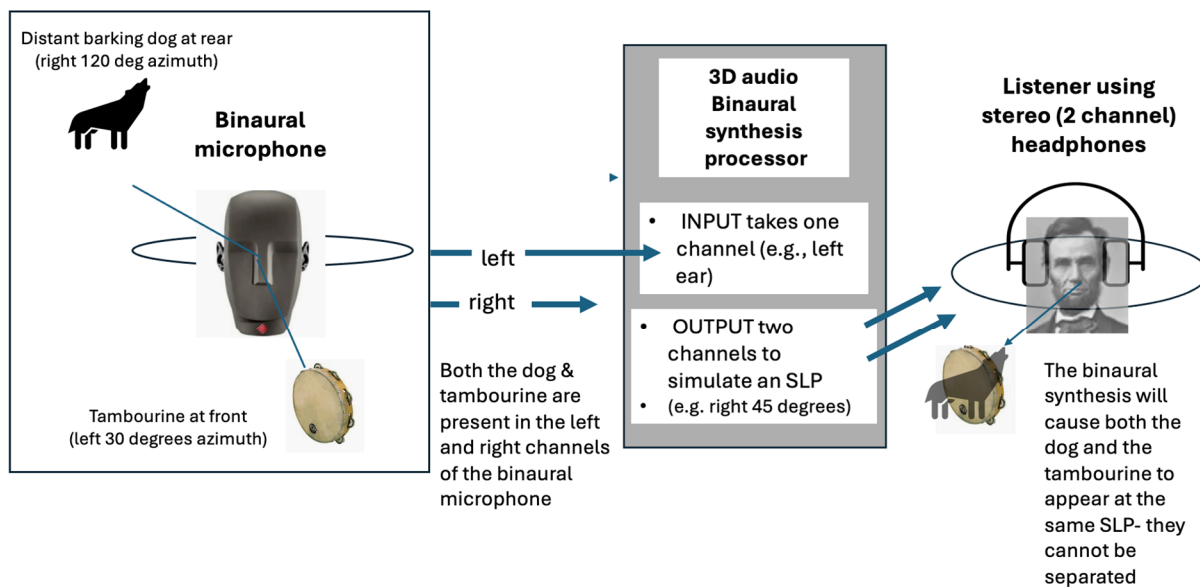
The binaural synthesis can, however, be considerably improved by making use of a single very important difference between recording and synthesis of binaural signals, namely, the possibility in the synthesis to change the HRTFs interactively when the direction of the sound incidence changes due to listeners' head movement. With binaural recording – and static synthesis [3D sound without head tracking], the listener has to sit still to have an authentic listening experience, since a recording cannot respond to later dynamical changes, such as the listener's head movements. This is not an impediment of principle nature for the [binaural] synthesis technique. The position and orientation of the listener can be tracked by so called head trackers, e. g., magnetic devices, and the synthesis can be made responsive to these changes by continuously updating signal processing.

Ex.2019, 245; Ex.2016, 198:3-199:2; Ex.2011, ¶24. In fact, Dr. Kyriakakis testified that this paragraph “strengthens” his opinion. Ex.2016, 300:7-301:15. But, in plain English, this paragraph echoes Dr. Begault’s testimony—with which Dr. Kyriakakis disagrees. This paragraph merely states that unlike binaural synthesis, a binaural recording “cannot respond to later dynamical changes,” because in binaural recording it is impossible to “change the HRTFs interactively.” Ex.2019, 245; Ex.2011, ¶¶24-25, 27-32. As Dr. Begault explains, this is because “the SLPs of sound sources in the resultant recording are fixed.” Ex.2011, ¶28. Dr. Kyriakakis’s double-thinking makes no sense.

Additionally, Dr. Kyriakakis explained that he disagreed with Dr. Begault, because one can divide recorded binaural sound to create two monaural channels, then process “process each monaural channel independently to make binaural sound from each separately.” Ex.2016, 134:20-135-21. But Dr. Kyriakakis himself admitted he is not processing a binaural sound—he is merely processing two monaural channels of sound. *Id.*, 136:13-22. As Dr. Begault explains, this is not the same as modifying a binaural sound in “post-processing.” Ex.2011, ¶¶28-29. Dr. Begault gives an example where two sound sources located at two different spots in a room (a dog and a tambourine) are recorded by two microphones (i.e., binaural recording). *Id.*, ¶¶29, 33. He explains that under Dr. Kyriakakis’s approach, where one monaural channel of the original binaurally-recorded sound is post-processed to

create a resultant binaural sound, the spatial relationship between the two sound sources is destroyed. *Id.* This example, reproduced below, shows that “it would be impossible to separate out the dog from the tambourine and manipulate their locations separately in post-processing of the recorded sound via binaural synthesis.”

Id.



Therefore, not only is Dr. Kyriakakis self-inconsistent in his disagreement with Dr. Begault, but his opinion is scientifically unsound. The Board should assign no weight to Dr. Kyriakakis’s declaration.

b. Ballas does not teach, suggest, or disclose binaural sound.

In his 1005 IPR declaration, Dr. Kyriakakis opines that Ballas discloses binaural sound. Ex.2023, ¶98. Specifically, “Ballas discloses the use of HRTFs to

localize binaural sound at an SLP.” *Id.* However, during the deposition, Dr. Kyriakakis switched again: “Ballas discloses transaural sound.” Ex.2016, 252:9-251:1. He continued that Ballas does not even *discuss* binaural sound. *Id.* Dr. Kyriakakis’s contradictory statements across different IPRs, depending on the need of the case, is not new. As the Board noted in the 1004 IPR, “it is unclear how, if Bedingfield does not capture “binaural sound” from participants 104, 106, and 108 (as Petitioner expressly acknowledges in the 1003 IPR), Bedingfield would provide “binaural sound” to listener 102 (as Petitioner argues in this proceeding). ... Dr. Kyriakakis provide[s] no explanation why that would be the case.” Ex.2017, 30. Like his Bedingfield opinion, Dr. Kyriakakis is inconsistent regarding Ballas—certainly undermining Dr. Kyriakakis’s credibility. *See Google LLC v. Makor Issues & Rights Ltd.*, IPR2017-00815, Paper 31, at 24 (PTAB Aug. 10, 2018) (noting a party’s failure to “explain persuasively” why it has “taken inconsistent positions between IPRs with respect to this limitation.”).

c. Dr. Kyriakakis’s declaration is self-contradictory vis-à-vis teachings of Li.

Dr. Kyriakakis’s declaration takes contradictory positions. In describing Li, Dr. Kyriakakis states “Li generally discloses all the limitations of the independent challenged claims, except for captur[ing] binaural sound or voice, and maintaining the voice’s origin at a fixed SLP while the user moves and changes

head orientation.” Ex.1003, ¶87. But yet again, he changes position, stating that “Li even contemplates binaural recording of a sound.” Specifically, Li describes ‘placing two microphones in inner ear locations,’ which the POSA would have understood to capture binaural sound.” *Id.*, ¶245. The Board should resolve this inconsistency against Petitioner.

In numerous examples, Dr. Kyriakakis submitted inconsistent testimony. Further, his testimony in this case contradicts his testimony in the 1004 and 1005 IPRs. In all instances, his credibility is case dispositive, because Petitioner relies exclusively on his declaration to supply *missing limitations* in Bedingfield (e.g., HRTFs and ITDs) or offer a motivation to combine. *See, e.g.*, Pet. 29 (citing Ex.1003, ¶147 regarding Bedingfield and ITDs); Pet. 38-39 (citing Ex.1003, ¶173 regarding Bedingfield and HRTFs); Pet. 48-49 (citing Ex.1003, ¶¶206-07 regarding motivation to combine Bedingfield and Ballas). Much like the 1004 IPR, the Board should assign no weight to Dr. Kyriakakis’s testimony.

B. Petitioner and Dr. Kyriakakis misapprehend the scope and meaning of the Challenged Claims.

“[W]e must disregard the testimony of an expert that is plainly inconsistent with the record, ... or based on an incorrect understanding of the claim[s].” *Homeland Housewares, LLC v. Whirlpool Corp.*, 865 F.3d 1372, 1378 (Fed.Cir.2017).

1. Independent Claim 1.

Petitioner and Dr. Kyriakakis mischaracterize Claim 1's scope as exclusively directed at binaural recording (but not binaural synthesis), a method of creating binaural sound that requires "using two microphones ... spaced at the same distance as the width of a human head." Ex.1003, ¶54; Ex.2011, ¶23. Petitioner's belief that Claim 1 is exclusively about binaural recording is readily apparent from its use of Andrea as prior art. Dr. Kyriakakis testified "Bedingfield does not disclose *capturing binaural sound*," and "that was the reason for the combination with Andrea." Ex.2016, 202:14-18; *see also* 204:5-11; Ex.1003, ¶¶84, 96, 122; Ex. 2011, ¶¶74-77; Pet. 12 ("Bedingfield generally discloses all independent claim limitations, *except for capturing binaural sound* ..."), 22 ("While *Bedingfield does not expressly teach capturing a binaural sound, this would have been obvious in view of Andrea*"). Dr. Kyriakakis further testified that Andrea does not disclose processing non-binaural content with HRTFs, the only other way to create binaural sound. Ex.2016, 259:2-14 (admitting that Andrea "*has two microphones next to the ears, so it captures binaural everything*," and that he "*did not see [binaural synthesis using HRTF] in Andrea nor did [he] use that*"); *see also* Ex.1003, ¶¶54-55 (listing recording and synthesis as the only two ways to create binaural sound). Therefore, Petitioner's reliance on Andrea—which discloses only binaural recording

per Dr. Kyriakakis’s admission—to teach “capturing binaural sound”¹¹ shows that Petitioner understands Claim 1 as exclusively related to binaural recording.¹² Incorrect. *See generally* Ex. 2011, ¶¶64-73.

¹¹ Dr. Kyriakakis testified that “capturing” binaural sound is equivalent to “recording” it, and that HRTFs are not used when binaural sound is created by recording. Ex.2016, 133:3-16.

¹² Dr. Kyriakakis could not confidently say whether the “electronic earphone” in Claim 1 includes two microphones. *See* Ex.2016, 172:2-16; 173:3-11; 177:16-178:7. Specifically, he testified that he could not unconditionally state that “capturing with an electronic earphone located at the head of a talking person binaural sound that will be provided to a listening person during a call is binaural recording,” because “[t]hose words are not sufficient for that, because it doesn’t say how many microphones the earphone has and what the listener is wearing during the telephone call.” *Id.* 172:2-16. He stated that Claim 1 captures binaural sound “[i]f there are two microphones.” *Id.* 173:3-11; *see also* 177:16-178:8. However, at no point in his declaration does Dr. Kyriakakis mention that the “capturing binaural sound” in Claim 1 hinges on the number of microphones in the electronic earphone. Instead, he and Petitioner assume—without explaining—that the electronic earphone necessarily possesses two microphones, conclude that Claim 1 pertains to binaural

A PHOSITA would comprehend that Claim 1 does not solely encompass the concept of capturing sound binaurally through two microphones. Limitation 1[a] describes “capturing, with an electronic earphone located at a head of a talking person, binaural sound that will be provided to a listening person.” The qualifier “that will be provided to a listening person” restricts the preceding term “binaural sound.” In context, this limitation indicates that the binaural sound captured by the microphones “located at the head of a talking person” will also be presented to the listener as binaural sound, as recited in forthcoming action in limitation 1[d]. Therefore, by employing the restrictive clause “that will be provided to a listening person,” limitation 1[a] forwards to limitation 1[d]. Ex. 2011, ¶67.

Indeed, a PHOSITA would understand the fundamental distinction between *capturing binaural sound* (limitation 1[a]) and *providing binaural sound* (limitation 1[d]), when the microphones capturing the sound are located at the head

recording, and utilize Andrea’s two-microphone system to teach capturing binaural sound. Indeed, Dr. Kyriakakis’s hesitation during the deposition to definitively affirm that the electronic earphone in Claim 1 encompasses two microphones contradicts Petitioner’s assumption that Claim 1 is exclusively aimed at binaural recording. The Board should resolve this inconsistency against Petitioner. Claim 1 also encompasses binaural synthesis, as explained herein.

of the talker in a telephone call. The sound being captured by the electronic earphone has two components: (1) the sound of environment (e.g., sound sources present in the talker's room) and (2) the talker's voice. Ex. 2011, ¶¶63-64. Petitioner and Dr. Kyriakakis miss a key point that "sound" encompasses a broader concept than "voice," effectively equating the two. For example, in discussing Claim 1, that recites "capturing binaural sound," Petitioner repeatedly uses the term "binaural voice" that is not recited in Claim 1. *See* Pet. 17 ("A POSA would have recognized that capturing binaural sound (including binaural voice) and providing it to the listener ..."); 18 ("A POSA would have had a reasonable expectation of success in applying Andrea's capturing binaural voice to Bedingfield."); *Id.* ("Modifying Bedingfield would have been as simple as adding a second microphone to Bedingfield's WED for capturing binaural voice."). Petitioner's confusion is readily ascertainable from other places of the Petition. *See* Pet. 33 (citing Ex.1003, ¶160 to claim that Bedingfield-Andrea teaches 15[b] because the combination discloses binaural sound or voice); ¶¶84, 87 (using binaural sound or voice interchangeably).

The terms are different. "Voice" is the vocal component of "sound," but "sound" also includes non-vocal components. For example, for a talker talking in the presence of a chirping bird, the sound captured by the electronic earphone includes both vocal (i.e., the talker's voice) and non-vocal (i.e., chirping sound) components. *See* Ex.2011, ¶¶60, 66. The '090 Patent specifically uses different

terms in the Claims: Claim 1 is directed at capturing and providing binaural sound during a call, whereas Claims 8 and 15 have to do with “voice” of the talker. Ex.2011, ¶¶57-58. The specification distinguishes between sound and voice. *See* Ex.1001 38:57-39:3; Ex.2011, ¶69.

This distinction is important, because an electronic earphone worn by a talker captures the spatial perspective of the talker. Ex.2011, ¶¶60, 85. This spatial perspective is transmitted to the listener during a call. Since the talker perceives the sound of the environment (e.g., chirping sound of a bird) as external to them, the listener also perceives the sound captured by the electronic earphone as external, when the same is provided to the listener. *Id.* However, while the electronic earphone worn by the talker can capture the talker’s voice binaurally, the listener would not perceive it binaurally, or, as Claim 1 recites, as being originated “at the sound localization point in empty space that is away from and proximate to the listening person.” *Id.* This is because the listener would hear the talker’s voice from the spatial perspective of the talker, which is inside of their head or in front of their mouth. *Id.* As Dr. Begault explains, the means of providing the talker’s voice to the listener binaurally—in cases where the microphones capturing the talker’s voice are worn by the talker—is binaural synthesis. *Id.*, ¶64.

Even Petitioner’s expert in district court admitted that “it’s not standard to capture the binaural voice of a person who’s speaking with two microphones in their

own ears.” Ex.2015, 108:11-15. Dr. Begault explains this is because the talker’s voice captured by binaural microphones is not perceived binaurally by a listener listening to the captured voice. Ex.2011, ¶¶72-73 (citing Ex.2028, Figure 3). To provide the talker’s voice binaurally to a listener, one must process (i.e., adjust) the captured voice by HRTFs—which is the processes described in limitations 1[b] and 1[c]. Ex.2011, ¶¶68-71. On the other hand, for all other sound sources in the room of the talker (i.e., the sound of the environment), the listener could perceive them as externalized as the talker perceives them as externalized.

Therefore, Claim 1 discloses both binaural recording (for sounds other than the talker’s voice) and binaural synthesis (for the talker’s voice). The specification also supports this:

Consider an example in which a listener and a talking person wear or use an electronic device that captures, transmits, and/or provides binaural sound during a telephone call. In the first part of the call, a VMP at the location of the speaker is designated to be at the point of the actual physical microphones being worn by the speaker (e.g., being at the ears of the speaker). As such, *various sounds in the environment of the speaker are perceived by the listener to emanate from locations relative to the ears of the listener, and the SLP of the speaker’s voice is perceived by the listener as approximately at the point of the mouth of the listener. Later in the call it is determined that the vocal part of the sound being sent by the speaker should be moved to a SLP off to*

the right side of the face of the listener. By way of example, the vocal part of the sound captured by the device worn by the speaker can be identified by the intersection of the set of frequencies associated with voice and with amplitudes and sound pressures matching those that likely emanated from the oral cavity a few inches away from the microphones

Ex.1001, 38:57-39:9; Ex.2011, ¶68. As this embodiment shows, the vocal component of the sound being captured binaurally needs further processing for a listener to hear it binaurally at a SLP away from and proximate to the listener. Ex.2011, ¶69. Dr. Begault explains Claim 1 discloses a mixed recording-synthesis method wherein the captured binaural sound is overlayed with binaural synthesis to fully externalize the sound for the listener. *Id.*, ¶69 (citing ¶26).

Furthermore, it is important to note that the SLP in Claim 1, located in “empty space that is away from and proximate to” the listener, “simulates *an origin* of the binaural sound” such that the sound “*originates* ... from the [SLP].” This describes an externalization process, exclusively caused by manipulation of auditory cues, which can only be accomplished by binaural synthesis. *See generally* Ex.2011, ¶¶9-22. Otherwise, cognitive and/or visual cues may help a listener imagine a sound not externalized in nature (e.g., stereo) as external to the listener—but even in that case the origin of the sound is inside the listener’s head not “away from and proximate to” them. *Id.*, ¶17. For example, a loudspeaker placed at the corner of a room is a

SLP by and of itself—but this SLP does not represent the origin of the stereo sound that the loudspeaker emits; as to the listener, the sound originates inside their head, not from the SLP. *Id.*, ¶45. Thus, Claim 1 recites more than designating an SLP—it requires the SLP “simulates an origin” of the sound.

2. Independent Claim 8.

Petitioner and Dr. Kyriakakis misunderstand the meaning and extent of this claim, failing to recognize that externalizing the voice of a talker at a location away from and proximate to a listener necessitates binaural synthesis. *Id.*, ¶22 (“[Binaural sound] is the only true way for producing externalization.”); *see also id.*, ¶¶78-80. As Dr. Begault explains, processing the sound with “accurate replicas of the human pinnae and head” (i.e., HRTFs) is a *must* for reliable externalization. *Id.*, ¶12. This is particularly the case for earphone externalization. *Id.*, ¶11. However, reliable sound externalization, regardless of the listener’s familiarity with the sound source, cannot be accomplished without tracking the change in head location and orientation of the listener (i.e., 6 degrees of freedom head-tracking system). *Id.*, ¶¶11, 14. Accordingly, without head-tracking, “localization errors may occur,” and thus the externalization cannot be realistic even with HRTF-based binaural processing. *Id.*, ¶14 (citing Ex.2025, 382-83).

Both experts agree that Claim 8 provides the talker’s voice to the listener as externalized sound. *Id.*, ¶¶78-80. Specifically, Dr. Kyriakakis admitted in his

deposition—despite his hesitation—that the sound appears to be external to the listener. Ex.2016, 193:5-19 (stating that based on “the definition I used in my -- in my report” the sound is external to the listener). Even Petitioner’s expert in the district court confirmed that “[i]n the context of spatial audio, away from ... mean away from, like, outside of the head of the person.” Ex.2015, 193:1-9.

Further, Claim 8 necessarily discloses a head-tracking mechanism used to determine the change in head orientation and location with respect to an SLP, in response to which the ITD of the talker’s voice is altered. Ex.2011, ¶¶78-80. Claim 8 inherently requiring a head tracking mechanism is supported by both Claim 8 language and the specification: “The location and the head orientation of the listener are tracked with the wearable electronic device 1720 so the sound can be adjusted with movements of the listener and continue to appear to originate from the sound localization point as the listener moves.” Ex.1001, 30:14-18. Even Petitioner’s district court expert confirmed that “determine[ing] head location orientation of the listener with respect to the sound localization point, this does mention a -- some sort of head-tracking system.” Ex.2015, 135:24-136:2.

Therefore, a PHOSITA, reading Claim 8, would understand that the talker’s voice is provided to the listener (1) via earphones, and (2) as externalized at a SLP away from and proximate to the listener, wherein (3) 6 DOF tracking of the listener’s head with respect to the SLP is accounted for. Ex.2011, ¶¶78-80. To a PHOSITA,

this means binaural sound, because there is no other way for accomplishing all these three steps. Again, Petitioner's district court expert agreed that inclusion of these three elements necessary imply HRTF processing (i.e., binaural synthesis). *See* Ex.2015, 134:22-140:11 (discussing Fig.2 of the '090 Patent).

Dr. Kyriakakis refused to admit this basic technical fact. While he stated a voice can be truly "externalized but not binauralized," (Ex.2016, 193:20-194:1), he could not explain how, simply providing a conclusory answer:

No, it could be binaural or it could be other things as well. ... I don't know if they have a name. I know here are processes that you do -- you can process signals, so that when they arrive at your ears, externalize the sound without being binaural. ... I don't know of a name. It involves multiple things like adding reverberation, adding other things, filtering. I don't -- I am not aware of a name for it. I'm just aware of the effect because I have experienced it.

Id., 196:19-197:10; Ex.2011, ¶¶18, 22. This response, lacking specificity and scientific rigor, highlights Dr. Kyriakakis's tendency toward unsubstantiated positions, such as the assertion in his 1004 IPR declaration that "he reads Li as disclosing..." which the Board dismissed as "unsupported." Ex.2017, 24 (citing Ex.2018, ¶104). However, he maintained that externalized sound cannot be stereo

either, because “stereophonic requires loudspeakers.”¹³ Ex.2016, 195:20-196:17. It raises questions about how Dr. Kyriakakis asserts the existence of externalized sound that is not binauralized if stereo sound is not viable for externalization, undermining the consistency of his argument.¹⁴

Claim 8 inherently using binaural synthesis and HRTF processing to externalize the voice of a talker for a listener is ascertainable by comparing Claim 10 with Claim 8. In Claim 8, an SLP is designated for the listener at a location “away from and proximate to” that “simulates” for the listener “the origin” of the voice of

¹³ This is yet another unscientific statement from Dr. Kyriakakis. As Dr. Begault explains, the reason that stereo-based method cannot externalize sound is not because they use loudspeakers; even “stereo headphone listening does not provide externalized imagery.” Ex.2011, ¶19. This is because in stereo “the width of the auditory spatial image is limited to being inside of the head of the listener.” *Id.* (citing Ex.2026, 9-10). Otherwise, stereo sound can be reproduced using headphones, but the reproduced sound cannot be perceived as externalized by the listener. *See id.*, ¶10.

¹⁴ To the extent, by “other things,” (Ex.2016, 196:14-17), Dr. Kyriakakis might have been referring to the “decorrelation” effect, Dr. Kyriakakis’s argument fails. *See id.*, ¶¶20-21.

the [talker]” such that it “appears to the [listener] to originate from” the SLP when the electronic call is executed. Claim 8 further discloses that its method “maintain[s] the origin of the voice of the second individual at the [SLP]” by “alter[ing] [the ITD] of the voice that the listener hears in response to” the listener’s head movement with respect to the SLP.

Claim 10 adds two limitations to Claim 8, effectively allowing the listener to “localiz[e] the voice of the [talker] during the electronic call.” Claim 10 requires that “movements of the head orientation of the [listener]” be “monitor[ed]” in response to the act of listener localizing the talker’s voice. Claim 10 further requires that “[HRTFs] of the [listener]” be adjusted “to more accurately simulate real HRTFs of the [listener] in response to the [monitored] movements.”

Both Claims deal with “simulat[ing] a location in empty space from where an origin of a voice of the [talker] occurs for the [listener].” The real difference between the two claims is the way each performs the simulation as a function of the change in head orientation. Claim 8 alters ITDs, whereas Claim 10 adjusts HRTFs, thereby improving simulation accuracy. The nature of this simulation is HRTF-based, because Claim 10 improves the accuracy of this specific simulation via adjusting HRTFs of the listener, rather than altering the ITD of the voice heard by the listener. Differently put, a PHOSITA, reading both claims, would understand that the word “simulate” in Claim 8 refers to a HRTF-based simulation whose accuracy is

improved by adjusting HRTFs of the listener. This is consistent with a PHOSITA's understanding of earphone externalization, as discussed by Dr. Begault. Ex.2011, ¶¶70-80.

Further, like 1, Claim 8 does not merely disclose a SLP—it requires the SLP “simulates ... an origin of a voice” such that the voice “originates ... from the [SLP].” This language points to an externalization process that is exclusively caused by manipulation of auditory cues, which can only be accomplished by binaural synthesis, as opposed to imagination-based externalization. *Id.*, ¶¶17, 46.

3. Independent Claim 15.

Petitioner and Dr. Kyriakakis advance the same argument for Claim 15 as Claim 1. *See* Pet. 32-36. In their view, Claim 15 is also directed at binaural recording, evidenced by using Bedingfield-Andrea in Ground 1. *Id.* However, they fail to recognize that Claim 15 is more specific than Claim 1, focusing only on the vocal component of the sound in the talker's room: “captures a binaural voice of the second person during an electronic phone call with the first wearable electronic device that the first person wears.” Ex.2011, ¶84. As explained in greater detail above, while a system equipped with a microphone array worn by a talker is capable of capturing a talker's voice binaurally, it cannot *provide* the captured voice to the listener biannually. *Id.*, ¶85. To provide the binaural voice of the talker to the listener such that the listener perceives the “origin of the binaural voice” at a SLP “away from and

proximate to” to the listener, binaural synthesis methods *must* be employed. *Id.*, ¶¶86-87. Accordingly, Claim 15 is exclusively directed at binaural synthesis.

Further, as noted with regard to Claims 1 and 8, Claim 15 does not merely disclose an SLP—it requires the SLP “represents an origin of the binaural voice.” This language points to an externalization process that is exclusively caused by manipulation of auditory cues, which can only be accomplished by binaural synthesis, as opposed to imagination-based externalization. *Id.*, ¶¶17, 46.

C. Petitioner’s obviousness analysis is flawed, because it is based on an incorrect understanding of the Challenged Claims.

1. Ground 1: Bedingfield in view of Andrea does not render obvious Independent Claims 1 and 15.

As discussed in §§II.B.1 and II.B.3, Petitioner misunderstands the correct scope and meaning of Claims 1 and 15. First, Petitioner confuses the distinction between “sound” and “voice,” treating them as equivalent, thereby advancing the same argument, based on binaural recording, for Claims 1 and 15. Second, Petitioner fails to recognize both Claims require not only that sound (or voice in Claim 15) be captured binaurally, but also to *provide* sound to the listener binaurally. Third, Petitioner does not explain how the vocal component of sound captured binaurally by an electronic earphone worn by the talker (or the wearable electronic device in Claim 15) can be provided to the listener binaurally without binaural synthesis. As explained in §§II.B.1 and II.B.3, binaural synthesis is the only way to provide to the

listener the talker's voice captured binaurally by microphones worn by the talker. Under this correct understanding of the Claims, Bedingfield in view of Andrea does not render obvious Claims 1 and 15.

As explained in §II.A.2, and determined by the Board in the 1004 IPR, Bedingfield discloses neither binaural synthesis nor HRTF processing of captured sound. Further, both experts agree that Andrea does not disclose binaural synthesis and HRTF processing. Ex.2016, 259:2-14; Ex.2011, ¶¶74-77. Therefore, by combining Bedingfield with Andrea a PHOSITA cannot possibly arrive at Claims 1 and 15, when both involve binaural synthesis as discussed in §§II.B.1 and II.B.3—notably, Claim 15 is exclusively directed at binaural synthesis.

Indeed, Petitioner's understanding that Claim 1 is directed at binaural recording, as opposed to a mix of binaural recording (for sound of environment) and binaural synthesis (for the talker's voice) is technically and scientifically incorrect. “[A] POSITA would know that the SLP of a sound created by binaural recording is fixed and cannot be arbitrarily transformed to new locations.” Ex.2011, ¶71. Thus, if Claim 1 was exclusively directed at binaural recording, as Petitioner understands it, “limitation 1[b] and 1[c],” reciting designating a new SLP for and adjusting a sound created by binaural recording, “would make no sense to a POSITA.” *Id.*

However, even under Petitioner's incorrect read of Claims 1 and 15, Petitioner's obviousness argument is insufficient. *First*, Andrea uses a beamforming

system with two microphones. This emphasizes the talker's voice by directing the microphone beam, while de-emphasizing surrounding sound as a form of noise cancellation. *See* Andrea, 11:6-24 (discussing Fig. 10 where an external loudspeaker to the side of the talker generates a noise signal); Ex.2011, ¶76. Andrea's two microphones work together to cancel this noise while providing direction for the talker's voice. Andrea, 11:6-24; Ex.2011, ¶76. Therefore, integrating Andrea's noise cancellation microphone system into Bedingfield's headphones would cancel or minimize the environment sounds in the talker's room. This contradicts Claim 1's goal, which is to capture the talker's voice and all other sounds present in the talker's room environment (e.g., chirping bird). As explained in §II.B.1, Claim 1 uses sound to encompass every possible sound wave—not just the talker's voice.

Second, Petitioner relies on Dr. Kyriakakis's statement that “[i]t would have been simple for a POSA to modifying the system of Bedingfield to include a wearable device that implements Andrea's second microphone for noise cancellation.” Ex.1003, ¶110. According to Petitioner, “[t]he modification would have been a natural extension of Bedingfield's WED.” Pet. 17. In essence, Petitioner states that if a talker replaces Bedingfield's communication devices that capture multi-channel audio with a wearable device with two microphones (as in Andrea), the binaural sound captured would be provided to the listener. Pet. 16-19; Paper 9 (“Decision”) 22. This argument is technically incorrect.

Even assuming that Claims 1 and 15 are directed at providing binaural voices via binaural recording, adding a second microphone to Bedingfield merely enables listening to the captured binaural voice with Bedingfield's stereo headphones; whereas both Claims require adjusting the captured binaural voice of the talker **before providing** it to the listener. Claim 1 discloses a method where binaural sound is captured, adjusted and then provided to the listener. Similarly, Claim 15 discloses a system where binaural voice is captured and then adjusted "in order to position" it around the listener. The specification expressly explains that the listener is provided with the captured sound *after* adjustment: "Sound waves can thus be captured, processed, and altered with HRTFs to change sound properties before being provided to a listener. Upon hearing the artificially altered sound wave, the listener will believe that the sound originates from a location different than the real location." Ex.1001, 29:4-8; Ex.2011, ¶65.

Third, both Claims describe adjustments (externalized position in 3D space) that can *only* be accomplished through binaural synthesis. Claim 1 requires that the "binaural sound captured" being "adjust[ed]" so "the binaural sound originates during the telephone call from the sound localization point in empty space that is away from and proximate to the listening person." Similarly Claim 15 requires "adjust[ing] the binaural voice of the second person during the electronic phone call in order to position a sound localization point of the binaural voice to a location that

is away from and proximate to the first person.” Indeed, the specification states “[t]hese adjustments include, but are not limited to, calculations for and changes to ITDs, ILDs, HRTFs, attenuation, reverberations, and other aspects of sound.” Ex.1001, 35:52-55. Bedingfield does not teach, suggest, or disclose binaural sound; thus, it cannot possibly teach, suggest, or disclose adjusting binaural sound as prescribed by the ‘090 Patent. Ex.2011, ¶43. If binaural sound captured by Andrea’s two-microphone system is added to Bedingfield, all Bedingfield can do with it is play it for the listener—as captured and without manipulation—through the disclosed stereo headphones.

Furthermore, as Dr. Begault explains, with respect to the talker’s voice captured by Andrea’s two-microphone system worn by the talker, Bedingfield’s listener will not hear the captured voice *as binaural*. Ex.2011, ¶¶60, 85. Specifically, the listener might hear the captured voice via Bedingfield’s stereo headphones inside their head, but they would not perceive the talker at an externalized SLP away from and proximate to the listener. *Id.* Even with Andrea’s two microphones without noise cancellation processing applied to Bedingfield’s headphones, the resulting spatial perspective for the listener will be fixed to the spatial perspective of the talker, i.e., from inside their head, even if the listener moves their head or their position. Ex.2011, ¶¶70, 85. Therefore, a PHOSITA would have no expectation of success to arrive at Claims 1 and 15 by combining Bedingfield and Andrea—such combination

would not supply the “adjusting” limitation, even under Petitioner’s read of the Claims.

Petitioner does not, nor can it explain how Bedingfield’s stereo-based system discloses methods for adjusting binaural sound. Such adjustment necessarily requires binaural synthesis—a concept unknown to Bedingfield. Petitioner attempts to address Bedingfield’s deficiency by taking Patent Owner’s district court position completely out of context, alleging that “auditory adjustments for spatial audio location like ITD and ILD were ‘well-known.’” Pet. 26. Petitioner mischaracterizes Patent Owner’s position.

In district court, Petitioner claimed that the specification lacks sufficient structure (i.e., algorithm) for any of these adjustments. Ex.2024, 15-16 (“[H]ere, the claims recite special purpose functions requiring special programming, and they do not recite sufficient structure for performing those functions and all that Patent Owner said was that specification provides sufficient structure.”). Showing portions of the specification describing HRTFs and ITD, Patent Owner responded that “[t]he disclosure of the well-known conventional method of convolution serves as adequate disclosure for performing part of the disclosed algorithm.” Ex.2020, 20. Specifically, Patent Owner argued that a PHOSITA may arrive at the structure recited in the claims by relying on all “structural clues,” including “additional structural definitions in the description, well-known methods for performing

claimed functions, and using the function itself.” *Id.*, 17. Patent Owner never said that the specific auditory adjustments recited in the claims are well-known. Patent Owner merely said a PHOSITA is familiar with the concept of auditory adjustments such as ITD adjustment, and thus “structures well-known to PHOSITAs may be relied upon in the algorithm.” *Id.*, 20 (citing *Enfish, LLC v. Microsoft Corp.*, 822 F.3d 1327, 1339-40 (Fed.Cir.2016)).

But Petitioner ignores the obvious: the recited adjustments in Claims 1 and 15 are performed in context of an electronic call between at least two individuals involving live voice exchange. As Dr. Begault explains, adjusting the SLP or other characteristics of captured voice in context of “two-way communication for participants who move” requires “intensive signal processing” to account for “control and compensation for system latency, signal quality, and algorithms to implement multiple non-linear time varying filters.” Ex.2011, ¶56. A PHOSITA needs to overcome these challenges to arrive at Claims 1 and 15, otherwise they cannot “emulate the experience of talking face-to-face directly with a human” participant during a call, which is the ultimate goal of the invention. Ex.1001, 7:39-41. Therefore, Petitioner’s argument that the “adjusting” limitation in Claims 1 and 15 is well-known is incorrect.

Fourth, as discussed in §II.B, both Claims identify SLPs away from and proximate to the listener that simulates (or represents for Claim 15) the origin of the

binaural sound (or voice for Claim 15). Claim 1 requires an SLP be designated “in empty space that is away from and proximate to the [listener] such that the [SLP] simulates an origin of the binaural sound at the empty space that the [listener] hears during the telephone call.” Similarly, Claim 15 requires a SLP that “represents an origin of the binaural voice of the [talker] to the [listener],” which is located at “a location that is away from and proximate to the [listener].” Petitioner relies on “change a participant’s location” feature in Bedingfield to conclude that Bedingfield discloses designating an SLP away from and proximate to the listener. But Petitioner misses the key distinction between stereo audio-based systems like Bedingfield and a system truly based on binaural synthesis: the former is not capable of externalizing sound “such that the sound localization point simulates an origin of the binaural sound at the empty space that the listening person hears during the telephone call.” Petitioner mentions “change a participant’s location” but fails to acknowledge that the SLP designated by Bedingfield cannot truly externalize the sound, because it is a stereo-based system. Ex.2011, ¶46. Thus, while Bedingfield’s listener may be able to designate an SLP, that SLP does not “simulate[] an origin of the binaural sound at the empty space,” and the listener does not perceive the sound originating from that SLP as “away from and proximate to” them. *Id.*

Further, Petitioner refers to Bedingfield’s Fig. 19a embodiment where the call participants are positioned around a circular table. Pet. 24. Petitioner also cites to

Ex.1037 and Ex.1038 to contend that allocating an empty space to a virtual participant is known in the field. *Id.* 24-25. But Petitioner fails to mention that in all these examples (including Bedingfield’s Figs. 19a-c), the imagination of the listener—as opposed to manipulation of auditory cues—allows the listener to perceive the stereo sound as externalized. Ex.2011, ¶¶45-46. This perception is guided by either memory and familiarity or visual cues.¹⁵ For example, Dr. Kyriakakis explains in the call described in Ex.1038, “voices of remote participants ... appear to originate from the projected image,” because “the phone participant’s image is // projected in the location of the selected seat.” Similarly, as Dr. Begault explains, Bedingfield’s Figs. 19a-c are also image-driven externalization. The listener must first imagine a circular table or a classroom setting to perceive externalized participant voices. However, Bedingfield’s auditory cues do not guide the perception of externalization, because stereo sound is naturally heard inside the head. Ex.2011, ¶¶45-46, 15-20. This is different from externalization process claimed in Claim 1, where the auditory cues alone—without images or reliance on

¹⁵ “[F]or example, a baseball game heard on an old-fashioned monaural AM radio with a single earpiece can be imagined to be outside of the head, independent of a conflict with the normal spatial auditory cues a person who was actually present at the baseball game would have.” Ex.2011, ¶17.

memory or familiarity—simulate an SLP at an empty space away from and proximate to the listener. *Id.*, ¶46.

Bedingfield in view of Andrea does not render obvious Claims 1 and 15. For the same reasons, the proposed combination does not render obvious other Claims discussed under Ground 1: Claims 2, 3 and 6 (depending from Claim 1) and Claims 16 and 17 (depending from Claim 15).

2. Grounds 2 and 3: Bedingfield in view of Andrea and other references does not render obvious Claims 18-19.

Because Challenged Claims 18 and 19 of Grounds 2 and 3 depend from independent Claim 15, and rely on Bedingfield in view of Andrea for at least the teachings of the limitations of Claim 15 above, these Claims are not rendered obvious by the proposed combinations.

3. Ground 4: Bedingfield in view of Ballas does not render obvious Independent Claim 8.

As discussed in §II.B.2, Petitioner fails to recognize Claim 8 is directed at binaural synthesis. Notably, Petitioner does not mention that the voice being provided to the listener in Claim 8 is necessarily binaural; otherwise, the SLP in Claim 8 cannot possibly “simulate[] a location in empty space from where an origin of a voice of the [talker] occurs for the [listener] through the electronic earphone such that the [SLP] in empty space is located away from and proximate to the [listener].” This misdirected Petitioner to use two prior art references that have

nothing to do with binaural sound or binaural synthesis. *See* §II.A.2. (discussing Bedingfield); §II.A.5.b. (discussing Ballas); Ex.2011, ¶¶102-103. Petitioner does not explain how a PHOSITA is motivated to combine two non-binaural references to arrive at claims based on binaural audio, especially with Ballas teaching away from using earphones. Ballas, Abstract (“In this system, a set of speakers is mounted in a location near the temple of a listener’s head ..., rather than in earphones.”); *see also* ¶17 (“In the present invention, a pair of speakers is mounted in a location near the temple of a listener’s head, such for example, on an eyeglass frame or inside a helmet, rather than in headphones. A head tracking system also mounted on the frame where speakers are mounted determines the location and orientation of the listener’s head and provides the measurements to a computer system for audio signal processing in conjunction with a head related transfer function (HRTF) filter to produce spatialized audio.”).

More importantly, combinations of Ballas and Bedingfield still do not teach certain limitations of Claim 8. Petitioner contends that Bedingfield alone teaches designating the SLP limitation of Claim 8, relying on the same argument it raised with respect to limitation 1[b]. For the reasons discussed above, the Board should reject Petitioner’s argument. *See* §II.C.1 (discussing the fourth reason).

Petitioner also contends that Bedingfield alone teaches the following limitation: “alter an interaural time difference of the voice that the first individual

hears in response to the first individual moving the head orientation with respect to the sound localization point.” Pet. 52. As explained in §II.A.3, Bedingfield does not teach ITD, and thus Petitioner’s obviousness analysis necessarily fails. To cure this deficiency, Petitioner makes a conclusory statement that “[a] POSA would have appreciated it was commonplace to alter ITDs to maintain the origin of a sound while the user moves head orientation.” Pet. 52. Assuming this is true, Petitioner never explains how a PHOSITA reading Bedingfield could alter ITDs, when such a feature is absent in Bedingfield. The absurdity of Petitioner’s position can be illustrated via this example: A deaf person, though maybe recognizing that a concert stage is a place for listening to music played by musicians, cannot actually “listen” to the music because they lack the necessary feature for listening: hearing.

Bedingfield in view of Ballas does not render obvious Claim 8 and any of the dependent claims depending from it.

4. Ground 5-8: Li in view of other references do not render obvious any of the Challenged Claims

Petitioner’s Li-based Grounds are based on the same incorrect understanding of the Challenged Claims discussed in §II.B. For Claims 1 and 15, Petitioner relies on Andrea for capturing binaural sound—but fails to recognize that the vocal component of the sound cannot be provided to the listener other than via binaural synthesis and use of HRTF processing. Similarly, in Claim 8 Petitioner does not

address how combination of Li and Ballas “simulates a location in empty space from where an origin of a voice of the [talker] occurs for the [listener] through the electronic earphone such that the [SLP] in empty space is located away from and proximate to the [listener].” As discussed above, such simulations necessarily require binaural synthesis and use of HRTFs. *See* §II.B.2.

Aside from this flaw in Petitioner’s argument, Petitioner does not explain how the proposed combinations teach several of the limitations in its limitation-by-limitation analysis. “All words in a claim must be considered in judging the patentability of that claim against the prior art.” *In re Wilson*, 424 F.2d 1382, 1385 (CCPA 1970). As the Board correctly observed, while Petitioner addresses limitation 1[b], it fails to address portions of limitations 1[c] and 1[d] that are worded differently than limitation 1[b]. Decision 31-32. “At this stage, we question whether Petitioner’s approach of using its analysis of limitation 1[b] as a shorthand for ten other allegedly ‘similar’ limitations that use different language and pertain to different steps, without any further explanation, is sufficient to meet its burden.” *Id.* 32.

Finally, even if this is overlooked, all Li-based Grounds necessarily fail, because Li does not teach an electronic/telephone call involving a live voice exchange between a talker and a listener as explained in §II.A.4. Petitioner exclusively relies on Li’s conference call embodiment in its obviousness analysis to

teach the call terms in the Challenged Claims. For the reasons discussed above, and similarly for the reasons explained by the Board in the 1004 , Li's conference call does not involve a live voice exchange. *See* §II.A.4. Similarly, Petitioner's *ad hoc* reasonings based on a fabricated dichotomy of recording/pre-recording fail as discussed above. *Id.*

Accordingly, Petitioner's proposed combinations based on Li fail, and Li in view of other references does not render obvious any of the Challenged Claims.

III. CONCLUSION

For the reasons above, Petitioner's proposed grounds of unpatentability fail. Thus, Patent Owner asks the Board to find all of the Challenged Claims patentable.

IPR2023-01003
Patent No. 9,226,090

Date: April 23, 2024

Respectfully submitted,

/Kevin S. Tuttle/

Kevin S. Tuttle (Reg. No. 52,307)
Kyle L. Elliott (Reg. No. 39,485)
Homayoon Rafatijo (Reg. No. 80,870)
SPENCER FANE LLP
1000 Walnut Street, Suite 1400
Kansas City, MO 64106
ktuttle@spencerfane.com
kelliott@spencerfane.com
hrafatijo@spencerfane.com
Tel: (816) 474-8100
Fax: (816) 474-3216

Erick S. Robinson (Reg. No. 51,354)
Patrick M. Dunn (Reg. No. 70,474)
Brian G. Strand (Reg. No. 62,442)
Jayme Partridge (Reg. No. 39,011)
Spencer Fane LLP
3040 Post Oak Boulevard, Suite 1300
Houston, TX 77056
erobinson@spencerfane.com
pdunn@spencerfane.com
Tel: (713) 212-2638
Fax: (713) 963-0859

Attorney for Eight KHZ, LLC

CERTIFICATE OF WORD COUNT

Pursuant to 37 C.F.R. §42.24(d), the undersigned hereby certifies that the foregoing Patent Owner's Response contains 13,997 words using the word count feature of Microsoft Word.

Dated: April 23, 2024

/Kevin S. Tuttle/
Kevin S. Tuttle (Reg. No. 52,307)
Attorney for Eight KHZ, LLC

CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. §42.6(e), the undersigned hereby certifies that the above-captioned Eight KHZ LLC's Patent Owner Response, and its cited exhibits, were served via electronic mail, as authorized by Petitioner's Mandatory Notices, on April 23, 2024 at the following email correspondence addresses of record:

Lisa K. Nguyen
Eric E. Lancaster
Paul Hastings LLP
1117 S. California Avenue
Palo Alto, CA 94303

David M. Tennant
Alan M. Billharz
Paul Hastings LLP
2050 M St., N.W.
Washington, DC, 20036

Grace I. Wang
Paul Hastings LLP
200 Park Avenue
New York, NY 10166

Date: April 23, 2024

Respectfully Submitted,
SPENCER FANE LLP

By: /Kevin S. Tuttle/
Kevin S. Tuttle (Reg. No. 52,307)
Spencer Fane LLP
1000 Walnut Street, Suite 1400
Kansas City, Missouri 64106-2140
Customer No. 21129
Tel: (816) 474-8100
Fax: (816) 474-3216
Attorneys for Eight KHZ LLC