



**Uniwersytet  
SWPS  
Sopot**

Sopot, 10.07.2025

dr hab. Konrad Bocian, prof. USWPS  
Uniwersytet SWPS  
Wydział Psychologii w Sopocie  
Polna 16/20  
81-745 Sopot

**Recenzja rozprawy doktorskiej Pani mgr Katarzyny Pypno-Blajdy pod tytułem: “Does listening to music change peoples’ moral beliefs?”**

Przedstawiona mi do recenzji rozprawa doktorska mgr Katarzyny Pypno-Blajdy pod tytułem “Does listening to music change peoples’ moral beliefs?” liczy 236 stron maszynopisu i składa się z czterech części: wstępu, teoretycznej, empirycznej oraz podsumowania. Zawiera streszczenia w językach polskim i angielskim kompletną bibliografię oraz aneksy zawierające materiały pomocnicze. Praca opatrzona jest poprawnym tytułem wskazującym na tematykę związaną z psychologią moralności i wpływem bodźców afektywnych.

Recenzję rozpocznę od przypomnienia trzech kluczowych kryteriów ustawowych określonych w art. 186 ust. 1 pkt 3 ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce. Zgodnie z tym przepisem, rozprawa doktorska: “(...) powinna stanowić oryginalne rozwiązanie problemu naukowego (...) oraz wykazywać ogólną wiedzę teoretyczną kandydata w danej dyscyplinie naukowej (...) oraz umiejętność samodzielnego prowadzenia pracy naukowej (...)”. W dalszej części recenzji odniosę się do tych trzech kryteriów, analizując jakość i wartość merytoryczną rozprawy doktorskiej, ocenę jej struktury, koncepcji teoretycznej, poprawności metodologicznej oraz znaczenia uzyskanych wyników. Szczególną uwagę poświęcę stopniowi

spełnienia powyższych warunków w świetle wyników poszczególnych badań empirycznych przedstawionych w pracy.

## 1. Ogólna wiedza teoretyczna w dyscyplinie

Część teoretyczna pracy podzielona jest na cztery rozdziały. Rozdział pierwszy stanowi próbę osadzenia rozprawy w szerokim kontekście historycznym i teoretycznym, obejmując zarówno klasyczne stanowiska filozoficzne (od starożytności po Nietzschego), jak i nowoczesne ujęcia psychologiczne. Doktorantka wykazuje dobrą orientację w tradycjach myślenia o moralności, sięgając po argumenty z zakresu etyki normatywnej, teorii cnót, deontologii, utilitaryzmu i współczesnej psychologii eksperymentalnej.

Niemniej w mojej opinii fragmenty poświęcone klasycznym filozofom, mimo że są obszerne i erudycyjne, są słabo zintegrowane z problematyką rozprawy. Brakuje jasnego uzasadnienia, w jaki sposób refleksje Platona czy Nietzschego wspierają tezy badawcze lub inspirują model empiryczny, który Doktorantka zamierza wykorzystać w swojej rozprawie. Idąc dalej, Doktorantka przedstawia różne definicje moralności, konkludując, że arbitralnie przyjmują definicję moralności zaproponowaną przez Paruzel-Czachurę (2023). Jednak w samej rozprawie ta definicja ogranicza się do jednego zdania, a Doktorantka nie przedstawia silnego uzasadnienia teoretycznego dla jej przyjęcia, po czym konkluduje, że ta definicja jest nieprecyzyjna, ponieważ moralność może zależeć od czasu, kultury oraz kontekstu.

W dalszych częściach rozprawy Doktorantka powołuje się na przyjętą przez siebie definicję moralności, co samo w sobie nie jest błędem, jednak sama definicja moralności nie jest wystarczająca, by uzasadnić, w jaki sposób muzyka może wpływać na sądy moralne. Pojęcie moralności wskazuje bowiem *czym jest* moralność, natomiast nie wyjaśnia *jakie mechanizmy psychologiczne* prowadzą do formułowania sądów moralnych oraz jak te procesy mogą być modulowane przez czynniki zewnętrzne, takie jak muzyka. Dlatego w rozprawie, która traktuje o wpływie muzyki na sądy moralne, konieczne jest odwołanie się do modelu teoretycznego, który

tłumaczy powiązania między emocjami, nastrojem, procesami poznawczymi i ocenami moralnymi. Na przykład modele takie jak Dual-Process Theory of Moral Judgment (Greene i in., 2001, 2008) czy Social Intuitionist Model (Haidt, 2001) pokazują, że sądy moralne wynikają z dynamicznej interakcji między afektywnymi intuicjami a refleksyjnym rozumowaniem. Teoria Afektywnego Torowania i Przetwarzania Zgodnego z Nastrojem (Clore & Huntsinger, 2007) mówi, że nastrój wpływa na przetwarzanie informacji. Pozytywny afekt potwierdza trafność dostępnych myśli i przekonań, natomiast negatywny afekt je unieważnia, co prowadzi odpowiednio do przetwarzania relacyjnego (powiązań między elementami) lub przetwarzania specyficznego dla elementów. To może sugerować, że osoby w dobrym nastroju mogą być mniej krytyczne lub bardziej konsekwencjalistyczne, a w złym bardziej deontologiczne.

W wymienionych ujęciach muzyka może wpływać na moralność poprzez zmianę nastroju czy pobudzenia emocjonalnego, co z kolei oddziałuje na równowagę między tymi systemami przetwarzania. Bez takiego teoretycznego zakotwiczenia Doktorantka nie wyjaśnia, dlaczego i w jaki sposób muzyka, będąca przede wszystkim bodźcem afektywnym, miałaby wpływać na oceny moralne. Uzupełnienie teoretycznego fundamentu pracy o wybrany model psychologiczny nie tylko uwiarygodniłoby badania, ale także pozwoliłoby na głębszą interpretację wyników w kontekście współczesnej psychologii moralności. W końcu rozdział nie kończy się syntezą, która pokazywałaby, jak przedstawione w rozdziale podejścia przekładają się na zaplanowane badania empiryczne (np. czy badania dotyczą bardziej moralnego poznania, emocji czy zachowania?). Ten brak utrudnia śledzenie spójności między teorią a empirią.

Podsumowując, pierwszy rozdział pracy doktorskiej ukazuje rozległą wiedzę Doktorantki w zakresie filozofii moralnej i psychologii moralności. Mimo pewnych zastrzeżeń redakcyjnych (brak syntezy, niejasna rama definicyjna), rozdział ten stanowi solidne i wartościowe wprowadzenie do dalszych części pracy. Niemniej ograniczona spójność i funkcjonalność rozdziału nie pozwala na

wyraźne określenie podstaw dla później sformułowanych przez Doktorantkę pytań badawczych i towarzyszącym im hipotez.

Drugi rozdział rozprawy stanowi rozbudowany przegląd badań empirycznych dotyczących wpływu muzyki na emocje, myślenie i zachowanie moralne. Doktorantka porządkuje dotychczasową literaturę w sposób zgodny z zadeklarowaną ramą triady moralności („thinking-feeling-doing”), obejmując badania nad dylematami moralnymi, zachowaniami prospołecznymi, moralną tożsamością oraz emocjami moralnymi. Przegląd obejmuje zarówno efekty pozytywne (np. zwiększone pomaganie po ekspozycji na muzykę), jak i nieistotne statystycznie, co wskazuje na świadomość złożoności badanego zjawiska oraz podejście zgodne z wartościami otwartej nauki.

Rozdział ten dowodzi dobrej znajomości Doktorantki literatury interdyscyplinarnej, łączącej psychologię moralności, psychologię muzyki oraz psychologię emocji. Doktorantka korzysta z wielu źródeł empirycznych, a także wskazuje na ograniczenia istniejących badań, w tym małe próby, heterogeniczność narzędzi, niereplikowalne efekty oraz brak analiz mediacyjnych i moderacyjnych. To świadczy o dojrzałości analitycznej i krytycznym podejściu do materiału źródłowego.

Jednocześnie rozdział nie jest wolny od istotnych braków koncepcyjnych. Przede wszystkim, przegląd ponownie nie prowadzi do spójnej i jasno określonej ramy teoretycznej. Doktorantka tę ramę sygnalizuje odwołując się do triady moralności, jednak szczegółowo jej nie opisuje, co nie pozwala określić jakie są główne założenia teoretyczne prezentowanej rozprawy. Choć Doktorantka omawia wiele badań w kontekście wpływu muzyki na moralność, nie określa, czy rozumie „wpływ muzyki” jako efekt symboliczny, estetyczny, afektywny, czy poznawczy. Nie wyodrębnia także mechanizmu działania, to znaczy czy muzyka działa jako bodziec sam w sobie, czy poprzez nastrój, który generuje. To rozróżnienie jest kluczowe z metodologicznego i interpretacyjnego punktu widzenia, a jego brak utrudnia formułowanie hipotez i interpretację wyników w dalszych częściach pracy.

Konsekwencją niedostatecznego teoretycznego uzasadnienia wyboru przez Doktorantkę triady „thinking-feeling-doing” jest brak jej rozwinięcia w strukturze przeglądu. Część badań została przywołana niejako arbitralnie, bez jasnego wskazania, który komponent triady reprezentują i dlaczego są istotne dla rozprawy. W kilku miejscach pojawiają się także przykłady zachowań (np. wybory konsumenckie), których związek z moralnością nie jest wystarczająco ugruntowany.

Podsumowując, drugi rozdział rozprawy prezentuje szeroką wiedzę Doktorantki w zakresie badań nad relacją muzyki i moralności, a także jej krytyczne podejście do dotychczasowej literatury. Niemniej jednak brakuje w nim spójnej i wyraźnie zdefiniowanej ramy teoretycznej, która pozwalałaby zintegrować omówione wyniki i przełożyć je na strukturę programu badawczego. Kluczowym niedopatrzeniem jest brak rozróżnienia pomiędzy wpływem muzyki jako zjawiska kulturowo-symbolicznego a wpływem wywołanego przez nią nastroju, co ma istotne konsekwencje dla dalszej części rozprawy.

Trzeci rozdział pełni w rozprawie funkcję pomostu między przeglądem literatury a częścią empiryczną. Doktorantka dokonuje tu krytycznej analizy dotychczasowych badań nad wpływem muzyki na moralność, a następnie prezentuje stan wiedzy dotyczący wpływu muzyki na trzy główne zmienne psychologiczne: utylitaryzm, moralną tożsamość i fundamenty moralne. Celem tej części jest zidentyfikowanie luk badawczych i sformułowanie uzasadnienia dla własnego programu badawczego.

Po pierwsze, Doktorantka trafnie identyfikuje problemy metodologiczne wielu dotychczasowych badań, m.in. małe i niezrównoważone próby, brak analiz mocy statystycznej, różnorodność narzędzi pomiarowych oraz ograniczoną możliwość uogólniania wyników ze względu na wykorzystywanie próbek WEIRD. Podnosi również problem braku spójnych i porównywalnych wskaźników, co stanowi realną barierę dla integracji wiedzy w tym obszarze. Dalej, Doktorantka wskazuje na potrzebę odejścia od dominacji prób amerykańskich i włączenia kontekstu chińskiego (USA vs. Chiny) jako przeciwwagi dla uprzedzeń kulturowych. Jest to ważny

i uzasadniony postulat badawczy. Z perspektywy metodologii psychologii społecznej, ten kierunek zwiększa trafność zewnętrzną i znaczenie aplikacyjne badań. W końcu Doktorantka odwołuje się do koncepcji transparentności, prerejestracji i replikowalności, co świadczy o dobrej orientacji we współczesnych trendach metodologicznych w psychologii.

Biorąc pod uwagę, że celem rozdziału jest formułowanie uzasadnienia dla własnego programu badawczego, należy rozważyć, czy Doktoranta zrobiła to w sposób nie budzący wątpliwości i czy proponowany program proponuje oryginalnie rozwiązuje problemu naukowego. Doktorantka deklaruje istnienie luki badawczej, ale nie zawsze jasno wyjaśnia, dlaczego dane pytanie jest istotne teoretycznie. Na przykład w sekcji 3.2 stwierdza, że *“badania nad wpływem muzyki na moralność są nieliczne”*, ale nie rozwija tego w kontekście wartości poznawczej, praktycznej lub aplikacyjnej tych badań. Samo stwierdzenie istnienia „luki” nie wystarcza jako uzasadnienie naukowej oryginalności, a jest to ustawowe kryterium oceny rozprawy doktorskiej.

Podobnie jak w Rozdziale 1 i 2, w Rozdziale 3 Doktorantka nie przedstawia zintegrowanego modelu teoretycznego, który łączy wpływ muzyki, afektu i ocen moralnych. Choć Doktorantka omawia trzy konstrukty (użyteczność, moralna tożsamość, fundamenty moralne), w Rozdziale 3 nie odnosi się wystarczająco do ich operacjonalizacji przedstawionej wcześniej. Nie pokazuje też, jak konkretnie jej projekt badawczy buduje się na istniejących teoriach. Może więc zadać pytanie, w świetle informacji przedstawionych w sekcji 3.1, czy Doktorantka prezentuje oryginalne rozwiązanie problemu naukowego, czy zamierza replikować badania, które jak sama słusznie zauważa, obciążone są wieloma problemami metodologicznymi.

Co w mojej opinii jest ważne, to fakt, że Doktorantka nie podejmuje wystarczająco refleksji nad tym, czy obserwowane w badaniach efekty muzyki nie wynikają przede wszystkim z wpływu na nastrój (a nie z semantyki, symboliki czy struktury muzycznej). Brakuje uwzględnienia alternatywnego modelu przyczynowego: muzyka - afekt - sądy moralne, który byłby zgodny z klasycznymi modelami psychologii afektu. Pominięcie takich modeli jest poważnym

ograniczeniem, zarówno koncepcyjnym, jak i projektowym. Uwzględnienie (a następnie testowanie lub odrzucenie) alternatywnych ścieżek przyczynowych jest nieodzownym elementem rzetelnej analizy naukowej i świadczy o umiejętności syntetyzowania różnych modeli teoretycznych.

Na końcu chciałbym zwrócić uwagę, że w rozdziale nie pojawiają się wyraźne pytania badawcze ani hipotezy. Mimo iż sekcja nazywa się „(...) *Aims of the Present Research*”, nie zawiera listy celów zidentyfikowanych na podstawie krytyki wcześniejszych badań. Stanowi to istotną lukę w rozprawie, ponieważ plan badawczy powinien wynikać bezpośrednio z analizy literatury. Ta luka jest do pewnego stopnia uzupełniona w drugiej, empirycznej części rozprawy, jednak nadal bez konkretnego odniesienia do wcześniej przywoływanej literatury. Na początku części empirycznej Doktorantka pisze: *“Building on the issues presented in previous sections, I conducted two experimental studies examining how music can influence one component of the moral triad – moral thinking.”* Uważam, że jest to uzasadnienie bardzo ogólne i nie wystarczające do określenia jaki konkretnie problem Doktoranta zamierza rozwiązać. Tym samym budzi to wątpliwości, czy rozwiązanie proponowanego problemu naukowego jest oryginalne, ponieważ ten problem nie zostało jasno zdefiniowany, tak samo jak nie została podkreślona jego oryginalność.

Rozdział 4 stanowi przejście od części teoretycznej rozprawy do realizacji programu empirycznego. Doktorantka prezentuje podstawowe założenia metodologiczne swojej pracy: opisuje typy planów eksperymentalnych (between-, within-, mixed-subject designs), uzasadnia wybór zmiennych współzmiennych oraz wyjaśnia powody uwzględnienia próbek międzykulturowych (USA i Chiny). Wskazuje również na stosowanie standardów otwartej nauki, co świadczy o dobrej orientacji w aktualnych tendencjach metodologicznych.

Na tle części teoretycznej (rozd. 1–3), rozdział ten pozytywnie wyróżnia się świadomością ograniczeń typowych dla psychologii eksperymentalnej (np. dominacji próbek WEIRD), a także deklaracją przestrzegania zasad transparentności i prerejestracji. Stanowi to kontynuację refleksji metodologicznej z Rozdziału 3, w którym Doktorantka krytykowała niespójność i niską jakość

wcześniejszych badań w omawianym obszarze. Jednakże, w relacji do poprzednich rozdziałów, widoczne są również poważne ograniczenia.

Przede wszystkim brakuje bezpośredniego powiązania metodologii z wcześniej zarysowanymi pojęciami i modelami teoretycznymi. W Rozdziale 3 Doktorantka wskazała szereg luk badawczych i zasygnalizowała potrzebę porównania efektów muzyki na różne komponenty moralności (np. deontologia, tożsamość moralna, fundamenty moralne), jednak w Rozdziale 4 nie doprecyzował, jak zaproponowane metody badawcze mają na celu ich przetestowanie. Struktura rozdziału metodologicznego nie jest oparta na założeniach teoretycznych, lecz pozostaje z nimi równoległa, co osłabia spójność całości rozprawy.

W części teoretycznej (rozd. 1–2) Doktorantka omawia znaczenie zmiennych indywidualnych (np. empatii, przywiązania narodowego, preferencji muzycznych), ale w Rozdziale 4 nie wskazuje ich roli analitycznej: czy miały pełnić funkcję moderatorów, mediatorów czy też zmiennych kontrolnych. To utrudnia ocenę adekwatności konstrukcji badawczej względem zakładanych mechanizmów psychologicznych. W Rozdziale 3 Doktorantka wskazywała na braki w metodach analitycznych wcześniejszych badań (np. brak analizy interakcji, niską moc testów), natomiast w Rozdziale 4 nie zaprezentowała własnej strategii analitycznej. Nie określiła, czy i w jaki sposób testowane będą interakcje, mediacje, czy zastosowane zostaną modele liniowe mieszane (LMM) lub inne narzędzia adresujące złożoności danych eksperymentalnych i międzykulturowych.

Podsumowując, choć Rozdział 4 spełnia podstawowe wymogi poprawności metodologicznej, jednak nie realizuje potencjału analitycznego zapowiadanego w poprzednich rozdziałach teoretycznych. Brakuje mu funkcji łącznika między teorią a badaniem. Doktorantka nie wyjaśnia, w jaki sposób przyjęte decyzje metodologiczne (np. wybór planu eksperymentalnego, zmiennych współzmiennych, populacji badawczej) służą weryfikacji hipotez wynikających z przeglądu literatury oraz wcześniej zidentyfikowanych luk badawczych. W efekcie, wrażenie spójnej i przemyślanej struktury badawczej zostaje zaburzone. Prezentowane Rozdziały od 1 do 4 sprawiają



wrażenie niezależnych części omawiających pokrewne zagadnienia z obszaru muzyki i moralności, ale niepowiązanych koncepcyjnie ani logicznie. Zamiast spójnie prowadzić czytelnika od ogólnej teorii przez konkretne pytania badawcze ku adekwatnie zaprojektowanej metodologii, każdy z rozdziałów zdaje się rozwijać osobny wątek, co utrudnia identyfikację głównej osi argumentacyjnej pracy.

W przypadku rozprawy doktorskiej, której celem jest wykazanie samodzielności badawczej oraz zdolności do tworzenia oryginalnych, empirycznie testowalnych modeli, brak tej spójności stanowi istotne ograniczenie. Dobrym rozwiązaniem byłoby wyraźniejsze zintegrowanie części teoretycznej i metodologicznej, na przykład poprzez wybranie lub zaproponowanie konkretnego modelu teoretycznego, który uwzględniałby zależności pomiędzy głównymi pojęciami teoretycznymi, zakładanymi mechanizmami psychologicznymi oraz przyjętymi decyzjami projektowymi. Taka struktura mogłaby pełnić funkcję mostu łączącego przegląd literatury z częścią empiryczną i znacząco poprawiłaby przejrzystość oraz spójność całej rozprawy.

Podsumowując, kryterium ogólnej wiedzy teoretycznej zostało spełnione w sensie zakresu i poprawności merytorycznej, jednak sposób jej przekształcenia w program badawczy jest niewystarczający. Najważniejszym ograniczeniem jest brak przedstawienia przez Doktorantkę zintegrowanego modelu teoretycznego, który łączyłby przegląd literatury z programem badawczym. Mimo że Doktorantka zna wiele teorii i badań, nie łączy ich w spójną koncepcję badawczą. Wiedza teoretyczna jest obecna, ale funkcjonalnie odseparowana od metodologii i hipotez, co w konsekwencji osłabia przejrzystość całej rozprawy.

## **2. Oryginalne rozwiązanie problemu naukowego**

Część empiryczna rozprawy doktorskiej mgr Katarzyny Pypno-Blajdy stanowi próbę odpowiedzi na pytanie o wpływ muzyki na moralne myślenie, zgodnie z ramą „moralnej triady” (thinking-feeling-doing). We wstępie Doktorantka pisze, że przeprowadziła dwa eksperyment,

jednak w rzeczywistości przeprowadziła cztery eksperymenty, ponieważ Badanie 2 podzielone jest na Badanie 2a, 2b i 2c.

W Badaniu 1 Doktorantka testowała czy ekspozycja na hymn narodowy wpływa na podejmowanie ocen moralnych, a w szczególności czy prowadzi do większego poparcia dla decyzji zgodnych z deontologią, a nie z utilitaryzmem.

W Badaniu 2a Doktorantka sprawdzała, czy muzyka charakteryzująca się wysoką walencją i pobudzeniem, określana dalej jako „*wesoła muzyka*”, wpływa na dwie kluczowe zmienne psychologiczne związane z moralnością: poczucie ważności tożsamości moralnej (mierzone za pomocą Self-Importance of Moral Identity Scale) oraz preferencje w zakresie fundamentów moralnych (mierzone Moral Foundations Questionnaire). Dodatkowo, Doktorantka badała, jakie emocje (afekty) są wywoływane przez ekspozycję na tę muzykę, co miało posłużyć zarówno jako weryfikacja skuteczności manipulacji, jak i potencjalny mechanizm pośredniczący.

W Badaniu 2b Doktorantka postawiła sobie cel odróżnienia wpływu emocji wywołanych przez muzykę od specyficznego działania samego bodźca muzycznego. Impulsem do tego było odkrycie w Badaniu 2a, że wśród uczestników z Chin „*wesoła*” muzyka wpłynęła na oceny fundamentu sprawiedliwości w podskali sądów MFQ. Doktorantka zasadnie stawia pytanie, czy efekt ten wynikał wyłącznie z emocji wywołanych przez muzykę, czy też z jakichś specyficznych właściwości samego bodźca muzycznego. To rozróżnienie jest ważne metodologicznie i teoretycznie, ponieważ pozwala na oddzielenie efektu afektywnego od możliwego wpływu formy bodźca (muzyka vs. inne modalności).

W Badaniu 2c Doktorantka przedstawia cel eksploracyjny, a dokładnie zbadanie różnic kulturowych w odbiorze muzyki wśród uczestników z Chin i USA. Celem badania było sprawdzenie, jak uczestnicy z obu krajów postrzegają i interpretują ten sam utwór muzyczny pod względem jego walencji, pobudzenia, znajomości i konotacji kulturowych. Cel ten jest ciekawy i zasadny w kontekście psychologii międzykulturowej oraz badań nad estetyką muzyczną.

Podsumowując, część empiryczna rozprawy doktorskiej mgr Katarzyny Pypno-Blajdy stanowi ambitną próbę zrealizowania programu badawczego mającego na celu zbadanie wpływu muzyki na moralne myślenie, zgodnie z przyjętą przez Doktorantkę koncepcję triady moralności („thinking-feeling-doing”). Niemniej ta koncepcja nie jest jasno i przekonująco przedstawiona w części teoretycznej. Program składa się z dwóch głównych badań: pierwszego, które koncentruje się na wpływie hymnu narodowego na wybory deontologiczne i utylitarne, oraz drugiego, podzielonego na trzy części, w którym analizowane są efekty „wesołej” muzyki na moralną tożsamość, fundamenty moralne, afekt oraz różnice kulturowe w odbiorze muzyki. Całość wpisana jest w nowoczesne standardy metodologiczne, obejmujące prerejestrację, analizę międzykulturową oraz otwarte dane. Poniżej przedstawiam ocenę tej części rozprawy, ze szczególnym uwzględnieniem konstrukcji badań, ich spójności teoretycznej oraz wartości naukowej.

## **2.1 Badanie pilotażowe**

W badaniu pilotażowym celem Doktorantki było wybranie muzyki dla grupy kontrolnej, która jest najmniej emocjonalna i znana, ponieważ Doktorantka chciała uniknąć wpływu emocji oraz rozpoznawalności na główne zmienne zależne w badaniu. Tutaj pojawia się moja pierwsza wątpliwość. Jeśli hymn narodowy jest specyficznym rodzajem muzyki, który przede wszystkim wpływa na poczucie dumy oraz aktywizuje konstrukty psychologiczne powiązane z tożsamością narodową, dlaczego w grupie kontrolnej utwór nie miał wywoływać emocji. Doktorantka tego nie wyjaśnia. Z metodologicznego punktu widzenia, utwór kontrolny także powinien wywoływać emocje, tak aby wykazać, że wpływ hymnu jest specyficzny, nie odnoszący się do emocji ogólnie, ale przede wszystkim do poczucia dumy.

W analizie wyników badania pilotażowego, Doktorantka piszę, że nie stwierdzono różnic między wersją hymnu ze słowami i bez słów pod względem reakcji emocjonalnych ani stopnia znajomości. Jednocześnie dalej konkluduje: “ (...) że siła oddziaływania hymnu nie wynika wyłącznie z jego elementów muzycznych, lecz z nadawanych mu znaczeń. Oznacza to, że hymnu

*narodowego nie można analizować jedynie jako utworu muzycznego, niezależnie od tego, czy wykonywany jest z tekstem, czy bez. Dlatego zdecydowałam się na najbardziej popularną wersję hymnu, wykonywaną przez chór, z tekstem.”*

Ta konkluzja wydaje się logicznie niespójna, a decyzja o wyborze hymnu ze słowami nieuzasadniona. Jeśli obecność słów nie miała znaczenia, to można wnioskować, że same słowa nie odgrywają kluczowej roli w odbiorze emocjonalnym hymnu, przynajmniej w warunkach eksperymentalnych. Dlaczego zatem Doktorantka wybrała wersję ze słowami? Dodatkowo, jak sama zauważa w Dyskusji, ten wybór był poważnym ograniczeniem, ponieważ grupa kontrola słuchała utworu muzycznego bez tekstu. Warto więc przemyśleć, czy interpretacja wyników nie powinna zostać nieco zmodyfikowana, np. w kierunku podkreślenia znaczenia symboliki hymnu jako całości, a nie jego warstwy tekstowej per se.

## **2.2 Badanie 1**

Badanie 1 zawarte w rozprawie doktorskiej mgr Katarzyny Pypno-Blajdy stanowi ambitną próbę empirycznego przetestowania wpływu hymnu narodowego na decyzje moralne. Doktorantka stawia jasny cel: sprawdzenie, czy ekspozycja na hymn prowadzi do bardziej deontologicznych niż utylitarystycznych wyborów w moralnych dylematach oraz czy efekt ten jest moderowany przez przywiązanie narodowe, obejmujące takie zmienne jak kolektywny narcyzm, tożsamość narodowa i orientacja polityczna. Założenia teoretyczne zostały oparte na literaturze z zakresu psychologii moralności i symboliki narodowej, w tym na badaniach pokazujących, że decyzje deontologiczne są postrzegane jako bardziej moralne i społecznie pożądane (np. Everett et al., 2016; Bostyn & Roets, 2017; Brown & Sacco, 2019). Sama konstrukcja badania oraz zakres hipotez są ambitne i wpisują się w aktualne dyskusje nad rolą tożsamości grupowej i norm społecznych w ocenach moralnych.

Pomimo pewnej spójności teoretycznej (nadal brakuje modelu teoretycznego), metodologia Badania 1 nie jest spójna z przewidywaniami sformułowanymi przez Doktorantkę. Po pierwsze, Doktorantka pisze: “(...) *Given that national attachment emphasizes group solidarity and social*

*norms and that deontological choices align more closely with socially desirable traits, I predicted that listening to the national anthem would evoke the emotion of pride and strengthen national attachment and group identification. Consequently, I expected participants exposed to the anthem to make fewer judgments consistent with utilitarianism, reflecting an alignment with a socially favorable orientation consistent with deontology*". Założenie, że wysłuchanie hymnu narodowego prowadzi do nasilenia emocji dumy, co z kolei wzmacnia przywiązanie narodowe oraz identyfikację grupową, a to w konsekwencji wpływa na decyzje deontologicznych, sugeruje, że Doktorantka zamierza przetestować model mediacyjny, który miałby wyjaśniać efekt podstawowy.

Problem w tym, że Doktorantka takiego modelu nie testuje. Zamiast potraktować mierzone emocje jako mediator, Doktorantka używa ich jako wskaźnika skuteczności manipulacji. Swoją drogą emocje, oceniane za pomocą prostych wskaźników deklaracyjnych, nie zostały powiązane z żadną konkretną hipotezą. W rozprawie nie wskazano, jakie emocje (nawet często wspomniana przez Doktorantkę emocja dumy), miałyby stanowić wskaźnik skuteczności ani dlaczego akurat te emocje zostały uwzględnione. Brakuje zarówno uzasadnienia teoretycznego, jak i kierunkowych przewidywań. Choć w modelu LMM wykazano istotne interakcje między krajem, grupą i emocją, interpretacja tych wyników jest fragmentaryczna, momentami niespójna z danymi (np. odwrotnie zinterpretowany efekt szczęścia w interakcji trójczynnikowej), a ich znaczenie dla skuteczności manipulacji pozostaje niejasne. Co więcej, dane wskazują, że emocje takie jak duma i szczęście były wyżej oceniane przez uczestników z Chin, a nie ze Stanów Zjednoczonych, co stoi w sprzeczności z przedstawioną interpretacją.<sup>1</sup>

<sup>1</sup> Dla emocji szczęście i duma w grupie eksperymentalnej: Chińczycy ocenili te emocje wyżej niż Amerykanie. Przykładowo: Szczęście: Chiny = 3.8, USA = 2.5; Duma: Chiny = 5.1, USA = 4.2. To oznacza, że w reakcji na hymn, Chińczycy doświadczali silniejszych pozytywnych emocji niż Amerykanie. Interakcja Country x Group x Happiness ma dodatni współczynnik (Estimate = 0.92, p = .009). To znaczy, że różnica między grupami (eksperymentalna vs. kontrolna) jest większa w USA niż w Chinach – a konkretnie: Amerykanie silniej zareagowali na hymn wzrostem szczęścia względem swojej grupy kontrolnej. Jednak mimo tej statystycznej interakcji, wartości bezwzględne emocji nadal wskazują, że Chińczycy ocenili szczęście wyżej niż Amerykanie w grupie eksperymentalnej. A zatem efekt interakcji dotyczy zmiany względem grupy kontrolnej. Ale Doktorantka błędnie interpretuje to jako ogólną intensywność emocji, pisząc, że „Amerykanie w grupie eksperymentalnej ocenili szczęście wyżej niż Chińczycy”, co nie zgadza się z danymi z Tabeli 1. Wniosek o niespójności interpretacji opiera się na tym, że Doktorantka pomyliła efekt względny (interakcja) z bezwzględnymi średnimi emocji. To klasyczny błąd interpretacyjny, kiedy nie uwzględnia się, że różnice w ramach interakcji niekoniecznie przekładają się na wyższe wartości surowe.

Dalej, Doktorantka założyła, że emocja dumy wywołana przez odsłuchanie hymnu narodowego, powinna wzmocnić sposób identyfikować się z grupą. Na marginesie, Doktorantka mierzy dwa sposoby identyfikacji, bezpieczny oraz narcystyczny, więc nie jest jasne, który z nich powinien się wzmocnić po odsłuchania hymnu narodowego. To z kolei powinno wpłynąć na sposób podejmowania decyzji moralnych. Doktorantka zatem powinna przetestować model z dwoma mediatorami, gdzie rodzaj muzyki jest zmienną niezależną, duma pierwszym mediatorem, identyfikacja z grupą drugim, a decyzje moralne zmienną zależną. Ponieważ Doktorantka mierzy dwa rodzaje identyfikacji z grupą, model mediacyjny powinien zostać wykonany dwa razy.

Dalej, w opisie Badania 1 Doktorantka pisze: *“Moreover, I predicted that listening to the national anthem would substantially impact utilitarian-oriented moral judgments among participants with a higher level of national attachment (H5.1 - H.8.8).”* Pomijając fakt, że Doktorantka nie rozpisuje tych hipotez, co poważnie utrudnia ocenę rozprawy, z opisu wynika, swoją drogą słusznie, że identyfikacja z grupą będzie ważnym moderatorem podstawowego efektu. Tymczasem w analizie statystycznej (rozdział 5.3.3) zmienna ta została potraktowana jako kowariant, co oznacza, że uwzględniono jedynie jej główny efekt, nieuwzględniający interakcji z warunkiem eksperymentalnym. To istotne odstępstwo od zapowiedzianego modelu teoretycznego, ponieważ analiza kowariancji nie pozwala na przetestowanie hipotezy, zgodnie z którą efekt ekspozycji na hymn zależy od poziomu przywiązania narodowego. Brak analizy interakcji prowadzi więc do logicznej niespójności między pytaniami badawczymi a zastosowanym modelem statystycznym. W rezultacie nie można uznać, że hipotezy H5.1-H8.8 zostały formalnie przetestowane w sposób zgodny z ich treścią. Dla rzetelnego testu należałoby zastosować model z interakcją (np. LMM lub regresję moderacyjną), który bezpośrednio uwzględnia warunkową zależność między zmienną niezależną (warunkiem muzycznym) a moderatorem (przywiązaniem narodowym).

Analiza wyników nie przyniosła szerokiego potwierdzenia hipotez. Zgodnie z danymi zawartymi w rozprawie, jedynie hipoteza 4.2 dotycząca obniżonego poziomu instrumental harm u osób słuchających hymnu, została potwierdzona i utrzymała istotność po korekcji Benjamini-Hochberga. Efekt ten był jednak bardzo mały ( $\eta^2p = .012$ ), co ogranicza jego praktyczne znaczenie. Pozostałe hipotezy (H1–H4.1, H3.3, H3.4) nie uzyskały wsparcia empirycznego, a hipotezy moderacyjne (H5.1-H8.8) nie zostały poprawnie przetestowane. Warto również zaznaczyć, że jeden z kluczowych wskaźników, model CNI, nie spełnił wymagań rzetelności i dopasowania do danych (poor model fit). Doktorantka trafnie zauważyła, że model nie pasował do danych w analizach międzygrupowych i międzykulturowych ( $G^2$  istotne dla  $df = 2$ ), co uniemożliwiło dalsze wnioskowanie o parametrach C, N i I oraz uniemożliwiło testowanie hipotez H3.1-H3.4. W konsekwencji interpretacja wyników, choć prowadzona w duchu akademickiej ostrożności, momentami przybiera postać nadmiernej spekulacji. Sugerowane wyjaśnienie dla obniżonego poziomu instrumental harm opiera się na domniemanym wzroście dumy i przywiązania narodowego, ale nie przeprowadzono analiz mediacyjnych. Tym samym mechanizm działania hymnu pozostaje niewyjaśniony, a narracja interpretacyjna opiera się raczej na domniemaniach niż na empirycznych przesłankach.

Podsumowując Badanie 1 porusza ważne i aktualne zagadnienie z pogranicza psychologii moralności, tożsamości narodowej i wpływu symboli kulturowych. Jednak znaczące niespójności między teorią a analizą, błędy w operacjonalizacji zmiennych, ograniczona trafność pomiaru oraz brak testów interakcji, mediacji oraz moderacji uniemożliwiają uznanie tego badania za silne wsparcie dla postawionych hipotez. Jego wartość pozostaje głównie eksploracyjna, a zapowiedziane związki przyczynowe nie zostały poprawnie przetestowane.

### **2.3 Badanie 2a**

W Badaniu 2a Doktorantka postawiła trzy hipotezy: że słuchanie muzyki o pozytywnej walencji (1) zwiększy pozytywny afekt (H1), (2) zmniejszy znaczenie, jakie uczestnicy przypisują

swojej tożsamości moralnej (H2), oraz (3) obniży znaczenie przypisywane fundamentom moralnym (H3). Hipotezy te oparto na wynikach wcześniejszych badań wskazujących, że pozytywna muzyka wywołuje pozytywne emocje i może łagodzić oceny moralne (Ansani et al., 2019; Koelsch, 2020; Seidel & Prinz, 2013).

Choć H1 znajduje oparcie empiryczne jest ono raczej oczywiste: wesoła muzyka poprawia nam nastrój. Z kolei hipotezy H2 i H3 budzą poważne wątpliwości teoretyczne. Przede wszystkim Doktorantka ich nie uzasadnia. Zamiast tego pisze: *“based on past research on the impact of music on moral judgments, I suspected that music may also influence how relevant it is for someone to be a good”*. H3 natomiast zostaje przez Doktorantkę założona bez jakiegokolwiek odniesienia do literatury, nawet tak wątpliwego jak dla H2.

Podobnie jak w Badaniu 1, Doktorantka wskazuje na potencjalny moderator pisząc: (...) *In this study, I wanted to test it from a different perspective: whether the participants' level of empathy would interact with the potential power of music to change the self-importance of moral identity and moral foundations.*” Zatem, Doktorantka zamierza testować czy empatia wchodzi w interakcję z muzyką (poziom empatii jest zmienną, która może nasilać lub osłabiać efekt, czyli zgodnie z definicją jest moderatorem), jednak w analizie statystycznej używa empatii jako kowariant.

Analiza danych nie potwierdziła żadnej z zakładanych hipotez. Według H1 słuchanie muzyki o pozytywnej walencji powinno zwiększyć pozytywny afekt. Tak się nie stało, dlatego niejasne jest dla mnie, dlaczego Doktorantka w Dyskusji do Badania 2a pisze: *“I observed that music increased positive affect, which confirms H1. Moreover, happy music lowered the negative affect, with a stronger effect on Chinese participants.”* Jest to logicznym zaprzeczeniem wyników z Tabeli 8 i 9 oraz samego wniosku Doktorantki: *“The positive affect remained unaffected.”*

Hipotezy (H2 i H3) także nie uzyskały potwierdzenia. Ani moralna tożsamość, ani większość fundamentów moralnych nie uległy zmianie po manipulacji. Wyjątkiem jest jedynie pojedynczy



efekt interakcji dla oceny fundamentu sprawiedliwości: uczestnicy z Chin oceniali ją niżej po odsłuchaniu muzyki. Doktorantka próbuje wyjaśnić ten wynik różnicami kulturowymi w rozumieniu sprawiedliwości oraz różnicami w odbiorze emocjonalnym bodźca muzycznego. Są to ciekawe hipotezy, lecz ich weryfikacja wykracza poza zakres przedstawionego badania. Co więcej, wyciąganie wniosków na podstawie jednego niestabilnego efektu (przy braku głównych efektów) może prowadzić do błędów interpretacyjnych. Dodatkowo, pomimo zgromadzenia danych o empatii, nie wykorzystano ich do przetestowania przewidywanych interakcji (np. muzyka x empatii).

W tym miejscu, chciałbym zwrócić uwagę Doktorantce na to, że wyciągnęła wnioski na temat empatii, które nie znajdują pełnego potwierdzenia w zaprezentowanych wynikach statystycznych.. Doktorantka pisze w Dyskusji do Badania 2a: *“Empathy increased the internalization of the self-importance of moral identity and the relevance of fairness and loyalty, which suggests that for empathic participants, it was more important to be moral, and also to be fair and loyal. These findings are in line with previous studies (e.g., Dawson et al., 2021; Strupp-Levitsky et al., 2020).”* Zgodnie z przedstawionymi w tabelach wynikami modeli LMM, jedynie efekty dla ocen ważności fundamentów moralnych związanych z sprawiedliwością (fairness) oraz lojalnością (loyalty) okazały się statystycznie istotne po zastosowaniu korekty Benjamini-Hochberga (odpowiednio  $p = .034$  i  $p = .020$ ). Efekt dla internalizacji tożsamości moralnej miał wprawdzie  $p = .014$ , jednak po korekcji przekroczył poziom istotności ( $p = .057$ ), co nie uprawnia do formułowania wniosku o istnieniu efektu. Możliwe, że Doktorantka miała na myśli, efekt dla symbolizacji tożsamości moralnej ( $p < .001$ ). Dodatkowo należy zaznaczyć, że zastosowany model nie uprawnia do wnioskowania przyczynowego. Empatia została potraktowana jako kowariant cechowy, a więc obserwowany związek ma charakter korelacyjny, nie zaś kauzalny. Tymczasem sformułowanie *„Empathy increased (...)”* sugeruje relację przyczynową, co jest niewłaściwe w kontekście zastosowanego podejścia analitycznego. Wobec powyższego, proponuję zrewidować interpretację

wyników w sposób bardziej ostrożny i zgodny z danymi empirycznymi. Rekomenduję również unikanie języka przyczynowego w interpretacjach zmiennych niemodyfikowanych eksperymentalnie.

Na osobną uwagę zasługują dwie kwestie: brak grupy kontrolnej oraz spójności Badania 2a z poprzednim eksperymentem (Badaniem 1). Brak grupy kontrolnej w Badaniu 2a stanowi poważne ograniczenie metodologiczne, które znacząco osłabia siłę wnioskowania przyczynowego. Choć zastosowano schemat pomiaru powtórzonego (pre–post), jego interpretacyjna wartość jest ograniczona, ponieważ nie można bowiem wykluczyć alternatywnych wyjaśnień obserwowanych zmian. Efekty habituacji, zmęczenia uczestników, refleksji między pomiarami czy ogólnego wpływu sytuacji badawczej, które mogą zniekształcać wyniki niezależnie od oddziaływania muzyki.

Szczególnie problematyczny jest brak warunku porównawczego, np. odsłuchania utworu neutralnego lub brak jakiegokolwiek muzyki, który pozwoliłby odróżnić specyficzny wpływ muzyki o pozytywnej walencji od ogólnego wpływu słuchania czegośkolwiek. Co więcej, założenia badawcze całego programu (badania 2a–2c) sugerują, że kluczowe jest właśnie rozdzielenie wpływu emocji od wpływu muzyki per se, cel, który powinien zostać zrealizowany już na poziomie konstrukcji Badania 2a. W obecnym kształcie badanie to nie spełnia założeń niezbędnych do przetestowania hipotez H2 i H3 w sposób rzetelny i rozstrzygający. Dodatkowo, choć Doktorantka, według mnie błędnie, twierdzi, że hipoteza H1 została potwierdzona, brak grupy kontrolnej uniemożliwia stwierdzenie, czy zmiana ta wynika z działania muzyki, czy też np. z upływu czasu lub efektu ponownego pomiaru. W rezultacie, zarówno konstrukcja eksperymentalna, jak i logika wnioskowania, wymagają w tym miejscu istotnego dopracowania.

Spójność Badania 2a z poprzednim eksperymentem (Badaniem 1) nie jest jednoznaczna. Choć oba badania podejmują temat wpływu muzyki na sądy moralne, to ramy teoretyczne i koncepcyjne różnią się istotnie. W Badaniu 1 manipulacja opierała się na ekspozycji na symbol

narodowy (hymn), a głównym obiektem analizy były wybory moralne o charakterze deontologicznym vs. utylitarystycznym, modelowane m.in. za pomocą modelu CNI i dylematów moralnych. W Badaniu 2a badana jest z kolei ważność moralnej tożsamości i fundamentów moralnych pod wpływem pozytywnej muzyki. O ile oba badania mieszczą się w szerokiej dziedzinie psychologii moralności, o tyle brakuje wyraźnego teoretycznego pomostu, który łączyłby logikę przejścia od Badania 1 do Badania 2a. W szczególności nie zostało jasno określone, w jaki sposób wnioski z Badania 1 były podstawą do zaprojektowania Badania 2a, oraz dlaczego konstrukty psychologiczne użyte w Badaniu 2 (tożsamość moralna, fundamenty moralne) są odpowiednią kontynuacją analizy rozpoczętej wcześniej (deontologia/utyliaryzm, normy, konsekwencje).

Co więcej, różne typy muzyki (symboliczna vs. afektywna) mogą oddziaływać przez odmienne mechanizmy psychologiczne. W Badaniu 1 był to m.in. wzrost przywiązania narodowego i emocji grupowych (np. duma), w Badaniu 2a chodzi raczej o ogólny pozytywny afekt. Doktorantka nie przedstawia jednak wyraźnej koncepcji porównawczej, która wskazywałaby, że oba typy wpływu muzyki (symboliczny vs. hedonistyczny) są częścią jednej, spójnej koncepcji wpływu muzyki na moralność. Bez takiego uzasadnienia program badań sprawia wrażenie eksploracyjnego zestawu niezależnych eksperymentów, a nie logicznie rozwijanej struktury argumentacyjnej. Tym samym sugeruję, żeby Doktorantka wyraźnie podkreśliła koncepcyjny związek między Badaniem 1 i 2a, np. poprzez pokazanie, że oba badania testują różne ścieżki wpływu kontekstu muzycznego na komponenty moralności (normatywność vs. tożsamość vs. afekt), lub że Badanie 2a testuje alternatywny mechanizm (afektywny) dla efektów obserwowanych w Badaniu 1 (symboliczno-tożsamościowych).

Podsumowując, Badanie 2a stanowi próbę sprawdzenia, czy pozytywna muzyka wpływa na afekt, moralną tożsamość i fundamenty moralne. Choć hipoteza H1 (wzrost pozytywnego afektu) oparta jest na literaturze, nie została potwierdzona. Mimo to Doktorantka błędnie twierdzi, że efekt

wystąpił. Hipotezy H2 i H3 (spadek znaczenia tożsamości moralnej i fundamentów) są słabo uzasadnione teoretycznie, a wyniki analiz nie przyniosły ich potwierdzenia, poza pojedynczym efektem dla fundamentu sprawiedliwości u uczestników z Chin. Poważnym ograniczeniem jest brak grupy kontrolnej, co uniemożliwia wnioskowanie przyczynowe. Efekt zmiany może wynikać z upływu czasu lub samego udziału w badaniu. Mimo zapowiedzi, empatii nie potraktowano jako moderatora, lecz jako kowariant, co nie pozwala przetestować założonej interakcji. Wnioski na temat empatii formułowane są w sposób przyczynowy, mimo że dane mają charakter wyłącznie korelacyjny. Brakuje też wyraźnego pomostu między Badaniem 1 a 2a, które różnią się zarówno ramą teoretyczną, jak i zakładanymi mechanizmami. W obecnym kształcie Badanie 2a wymaga doprecyzowania w zakresie konstrukcji eksperymentalnej, interpretacji oraz powiązania z resztą rozprawy.

## 2.4 Badanie 2b

Wprowadzenie do Badania 2b stanowi logiczne rozwinięcie wyników Badania 2a, w którym Doktorantka stwierdziła, że wśród chińskich uczestników ekspozycja na „wesołą” muzykę wpłynęła na ocenę fundamentu sprawiedliwości w podskali osądów MFQ. Doktorantka zasadnie stawia pytanie, czy efekt ten wynikał wyłącznie z emocji wywołanych przez muzykę, czy też z jakichś specyficznych właściwości samego bodźca muzycznego. To rozróżnienie jest ważne metodologicznie i teoretycznie, ponieważ pozwala na oddzielenie efektu afektywnego od możliwego wpływu formy bodźca (muzyka vs. inne modalności).

Doktorantka formułuje następujące przewidywania dotyczące Badania 2b: *“I predicted that the picture would influence participants in the same way as a musical piece: lower negative affect (H1) but not influence positive affect (H2) and make participants agree less with the judgments subscale of MFQ (H3).”* Doktorantka zatem, zakłada że obrazek zadziała „tak samo jak muzyka” w Badaniu 2a, czyli obniży negatywny afekt, nie wpłynie na pozytywny afekt oraz obniży zgodność z ocenami moralnymi w subskali MFQ (zakładam, że chodzi o oceny sprawiedliwości).

Choć hipotezy te są zwięźle sformułowane, niektóre kwestie wymagają doprecyzowania.

Przede wszystkim niejasna jest logika hipotezy H3. Skoro obrazek nie miałby wpływać na pozytywny afekt (H2), to na jakiej podstawie zakłada się, że wpłynie na oceny moralne (H3)? Jeśli mechanizm miałby być wyłącznie afektywny, to brak zmiany w pozytywnym afekcie powinien wskazywać również na brak efektu moralnego. Tę niekonsekwencję warto byłoby wyjaśnić (np. przez przypuszczenie, że kluczowy jest spadek afektu negatywnego, a nie wzrost pozytywnego). Ale to nie zostało wyartykułowane lub przetestowane w Badaniu 2a, na przykład za pomocą modelu mediacyjnego.

Dobór bodźca, zdjęcie pieska z zestawu OASIS, jest dobrze uzasadniony: posiada wysoką walencję i pobudzenie, nie zawiera treści kulturowo nacechowanych, i był dobrany w sposób zbliżony do muzyki użytej w Badaniu 2a. Zaletą Badania 2b jest jego konceptualna spójność z Badaniem 2a, ponieważ Doktorantka w logiczny sposób próbuje rozdzielić mechanizmy emocjonalne od specyfiki modalności bodźca. Wciąż jednak brakuje wyraźnego modelu przyczynowego, np. czy wpływ na osądy moralne ma nastąpić przez redukcję negatywnego afektu? A może przez inne mechanizmy (np. wzrost bezpieczeństwa, rozluźnienia, itp.)

Badanie 2b, w mojej opinii, ma jedno poważne ograniczenie, które w konsekwencji mogło mieć związek z brakiem potwierdzenia H3 oraz, co ważniejsze, wyciągnięcia nieprawidłowych wniosków przez Doktorantkę. Doktorantka wprost deklaruje, że chce rozstrzygnąć, czy efekt z Badania 2a był wynikiem samej muzyki, czy afektu wywołanego przez muzykę: *“I aimed to test whether this effect was driven only by affect (and the music piece was just a stimulus that smoothly evoked it) or if there was something special about the music itself.”* ale w Badaniu 2b nie testuje muzyki. Prowadzi to do logicznej sprzeczności, ponieważ Doktorantka pisze, że chciała sprawdzić czy efekt wynika z muzyki czy z afektu, ale usuwa muzykę z badania. Aby móc twierdzić, że bodźce (muzyczny i wizualny) są funkcjonalnie równoważne i że ewentualne różnice wyników wynikają ze specyfiki muzyki per se, konieczne byłoby przeprowadzenie bezpośredniego

porównania warunków w jednym badaniu, z losowym przydziałem do grup (muzyka vs. obraz) i testem interakcji bodziec x pomiar. Tymczasem w Badaniu 2b testowany jest tylko jeden warunek, obraz bez grupy muzycznej, co uniemożliwia logiczne i statystyczne porównanie z Badaniem 2a. W efekcie, brak różnicy w subskalach MFQ po ekspozycji na obraz nie uprawnia Doktorantki do wniosku, że „*muzyka działa inaczej niż obraz*”, ponieważ może ona po prostu działać silniej, być lepiej zakotwiczona kulturowo, lub bardziej znacząca symbolicznie. Konkluzja o specyfice oddziaływania muzyki nie została zatem empirycznie potwierdzona, a projekt badania nie pozwala na rozstrzygające testowanie hipotezy H3.

Podsumowując, Badanie 2b stanowi logiczną próbę rozdzielenia wpływu afektu od specyfiki muzyki jako bodźca. Choć konceptualnie dobrze osadzone, jego konstrukcja metodologiczna nie pozwala na jednoznaczne wnioskowanie. Kluczowy problem polega na braku grupy muzycznej. Mimo że Doktorantka deklaruje chęć porównania wpływu muzyki i afektu, bada wyłącznie wpływ afektu za pomocą obrazu. To uniemożliwia testowanie zakładanej hipotezy o specyfice działania muzyki. W efekcie brak zmian w ocenach moralnych po ekspozycji na obraz nie uprawnia do konkluzji, że muzyka działa inaczej niż bodziec wizualny. Być może po prostu działa silniej lub głębiej kulturowo. W obecnej formie Badanie 2b nie dostarcza rozstrzygających danych na temat mechanizmów obserwowanego wcześniej efektu.

## **2.5 Badanie 2c**

We wstępie do Badania 2c Doktorantka przedstawia cel eksploracyjny: zbadanie różnic kulturowych w odbiorze muzyki wśród uczestników z Chin i USA. Zgodnie z deklaracją, intencją Doktorantki było sprawdzenie, jak uczestnicy z obu krajów postrzegają i interpretują ten sam utwór muzyczny pod względem jego walencji, pobudzenia, znajomości i konotacji kulturowych. Cel ten jest ciekawy i zasadny w kontekście psychologii międzykulturowej oraz badań nad estetyką muzyczną. Problem pojawia się jednak w logicznym uzasadnieniu badania.

Doktorantka stwierdza, że Badanie 2c wynika bezpośrednio z wcześniejszych ustaleń, a mianowicie: *„Because I found an effect of listening to happy music on judgments about fairness foundations only among Chinese participants and not Americans (Study 2a), and that effect was driven solely by the music itself, not by visual stimuli (Study 2b), I aimed to test the differences in the cultural connotations of music perception”*. To twierdzenie stanowi podstawę logiczną Badania 2c, ale jednocześnie ujawnia kluczowy błąd w rozumowaniu Doktorantki. W Badaniu 2b nie użyto warunku z muzyką, przez co nie można wnioskować, że efekt zaobserwowany w Badaniu 2a był *„specyficzny dla muzyki”*, a nie wynikał ogólnie z afektu. Brak bezpośredniego porównania dwóch typów bodźców (muzycznego i wizualnego) w ramach jednego eksperymentu uniemożliwia wyciągnięcie takiej konkluzji. Tym samym fundament, na którym zbudowano Badanie 2c, jest niewystarczająco umocowany empirycznie i logicznie wątpliwy.

Ponadto, jeśli punktem wyjścia Badania 2c jest rzekoma *„specyfika oddziaływania muzyki”*, należałoby najpierw ją wykazać empirycznie, czego nie zrobiono. Zamiast tego, interpretacja Badania 2b posłużyła jako nieuprawnione potwierdzenie tej tezy, mimo że brak efektu poznawczego po obrazku może wynikać z wielu alternatywnych przyczyn (np. braku znaczenia społecznego bodźca, mniejszego zaangażowania, różnicy sensorycznej). Podsumowując, choć sam cel Badania 2c jest interesujący i potencjalnie wartościowy, jego uzasadnienie w strukturze rozprawy jest nieprzekonujące. Opiera się na błędnej przesłance metodologicznej i logicznej, przez co narusza spójność narracyjną całego bloku badawczego (Badania 2a–2c).

Badanie 2c miało charakter eksploracyjny i jego głównym celem było zbadanie kulturowych różnic w odbiorze muzyki między uczestnikami z Chin i USA. Założeniem (według mnie błędnym) było, że wcześniejsze wyniki (Badanie 2a i Badanie 2b) sugerują specyficzne oddziaływanie muzyki jako bodźca emocjonalnego, co miało uzasadniać potrzebę zbadania, jak muzyka jest interpretowana i oceniana w różnych kontekstach kulturowych. Z metodologicznego punktu widzenia badanie zostało zaprojektowane poprawnie, jeśli przyjąć jego eksploracyjny charakter.

Mimo powyższych zalet, projekt ma także istotne ograniczenia. Zwłaszcza w kontekście jego roli w strukturze całego programu badawczego.

Po pierwsze, jak wskazałem wcześniej, wniosek o „*specyfice muzyki*” nie został empirycznie uzasadniony. Tym samym, Badanie 2c rozwija założenie, które nie zostało wiarygodnie potwierdzone, co osłabia logikę ciągłości między badaniami. Po drugie, brakuje zmiennych porównawczych z Badania 2a. Mimo że Doktorantka użyła tego samego utworu, nie przeprowadziła porównania z wcześniejszymi pomiarami z Badań 2a i 2b. W konsekwencji Badanie 2c funkcjonuje jako odrębny blok danych, co ogranicza możliwość triangulacji wyników. W końcu brakuje hipotez kierunkowych, co samo w sobie nie jest wadą w badaniach eksploracyjnych, ale czyni interpretację wyników bardziej podatną na błędy confirmacyjne, szczególnie gdy mają one wspierać wcześniej założone tezy.

Wyniki Badania 2c sugerują, że mimo braku różnic w deklarowanej identyfikacji pochodzenia kompozytora, występują istotne różnice międzygrupowe w jakościowych skojarzeniach i wrażeniach, a także w ocenie walorów muzycznych utworu (lubienie, znajomość, pobudzenie, walencja). Wyniki te są interesujące i spójne z szerszymi obserwacjami dotyczącymi ekspresji emocji i skojarzeń kulturowych. Niemniej jednak ich funkcja wyjaśniająca w kontekście wcześniejszych badań (zwłaszcza Badanie 2a) pozostaje ograniczona ze względu na brak integracji wyników i kontrowersyjne założenia płynące z Badania 2b. Z tych względów, w obecnej postaci, Badanie 2c można traktować jako interesujące uzupełnienie projektu, ale nie jako silny dowód na mechanizm specyficznego wpływu muzyki na procesy moralne.

Doktorantka w podsumowaniu Badania 2c formułuje kilka wniosków, z których część jest trafna, a część budzi istotne zastrzeżenia interpretacyjne. Przede wszystkim należy docenić konsekwencję analityczną. Doktorantka interpretuje wyniki zgodnie z uzyskanymi danymi i wyraźnie wskazuje różnice kulturowe w zakresie skojarzeń, długości wypowiedzi oraz ocen poznawczych i afektywnych bodźca muzycznego. Trafne jest również odniesienie do znanych



efektów psychologicznych, takich jak efekt czystej ekspozycji, czy znaczenia znajomości muzyki dla jej odbioru. Zasluguje również na uznanie ostrożność interpretacyjna. Doktorantka nie przecenia wielkości efektów, jasno wskazując, że różnice w znajomości utworu miały średnią siłę efektu, a różnice kulturowe raczej niewielką. To poprawny i rzetelny sposób prezentacji wyników eksploracyjnych.

Zastrzeżenia budzi jednak sposób łączenia wyników Badania 2c z wnioskami z Badań 2a i 2b. Doktorantka sugeruje, że efekt zaobserwowany w Badaniu 2a (słabsza ocena fundamentu fairnes wśród Chińczyków po wysłuchaniu muzyki) był „napędzany” przez znajomość i kulturową przynależność utworu i w tym sensie to nie kultura per se, lecz znajomość muzyki miała kluczowe znaczenie. Takie wnioskowanie byłoby jednak zasadne tylko wtedy, gdyby znajomość muzyki była bezpośrednio testowana jako moderator efektu moralnego w Badaniu 2a, a nie była. Co więcej, sugerowanie, że „muzyczny bodziec miał większy wpływ niż różnice kulturowe” jest zbyt daleko idącym uogólnieniem, biorąc pod uwagę charakter Badania 2c. Było to badanie eksploracyjne, które nie testowało bezpośrednio efektów moralnych ani nie umożliwiało rozstrzygnięcia, co było głównym czynnikiem sprawczym w Badaniu 2a (muzyka, kultura, czy ich interakcja). Wnioskowanie o hierarchii wpływu (bodziec > kultura) jest w tym kontekście statystycznie nieuprawnione i logicznie ryzykowne.

Podobnie problematyczne jest traktowanie różnic w znajomości utworu jako samodzielnej przyczyny różnic w reakcjach moralnych. Dane z Badania 2c pokazują tylko, że Amerykanie ocenili muzykę jako bardziej znajomą, ale nie wiadomo, czy to ta znajomość wpłynęła na brak efektu moralnego w tej grupie (brak testu mediacji lub moderacji). Tym samym, argumentacja o „decydującej roli bodźca” opiera się na nieprzetestowanej hipotezie, a nie bezpośrednich wynikach. Zamiast konkluzji przyczynowych, bardziej zasadne byłoby sformułowanie hipotez do dalszych badań. W obecnej postaci Badanie 2c nie dostarcza wystarczających podstaw do jednoznacznych

wniosków o mechanizmach działania muzyki na oceny moralne, ale wnosi interesujący wkład eksploracyjny w postaci danych kulturowych i percepcyjnych.

Dyskusja ogólna do Badań 2a, 2b i 2c dostarcza czytelnikowi ogólnego podsumowania wyników i prób ich zinterpretowania, ale zawiera nieuprawnione uogólnienia i brak spójności logicznej w łączeniu wniosków z trzech badań. Kluczowy problem to twierdzenie, że zaobserwowany efekt (spadek znaczenia sprawiedliwości) był wynikiem działania bodźca muzycznego, a nie afektu czy różnic kulturowych, mimo że dane nie pozwalają na takie rozstrzygnięcie. Dodatkowo, brak testu interakcji (muzyka vs. obraz) w Badaniu 2b, a także brak bezpośredniego testu moderacyjnego znajomości/oceny muzyki w Badaniu 2a osłabia wiarygodność konkluzji. Doktorantka poprawnie identyfikuje kierunki dalszych badań, ale nie zachowuje wystarczającej ostrożności interpretacyjnej w odniesieniu do własnych danych. W obecnym kształcie, dyskusja do Badań 2a, 2b i 2c nie spełnia w pełni wymogów rzetelnej analizy integracyjnej i wymaga dopracowania, zarówno w zakresie spójności logicznej, jak i rozdzielenia danych empirycznych od spekulacji teoretycznych.

## **2.6 Dyskusja ogólna**

Dyskusja ogólna w obecnym kształcie prezentuje ambitną próbę powiązania wyników empirycznych z szeroką refleksją teoretyczną na temat roli muzyki w kształtowaniu przekonań moralnych. Doktorantka stara się osadzić uzyskane wyniki w kontekście klasycznych i współczesnych ujęć moralności oraz przywołuje liczne nurty badawcze od filozofii po psychologię eksperymentalną. Mimo tych zalet, należy jasno wskazać, że dyskusja ogólna nie spełnia standardów merytorycznych wymaganych dla rozprawy doktorskiej w zakresie psychologii, a główne wnioski interpretacyjne są zbyt daleko idące, niespójne z danymi i logicznie nieuprawnione. Poniżej przedstawię główne argumenty na których oparłem swoje wnioski.

Po pierwsze, najważniejszym problemem tej sekcji jest sformułowanie przez Doktorantkę konkluzji: „*Yes, music can influence moral beliefs.*” Takie stwierdzenie jest nieproporcjonalne

wobec siły i zakresu danych empirycznych. Efekt dotyczący wpływu muzyki dotyczy tylko jednej grupy kulturowej (Chińczyków), jednego typu bodźca (utwór z filmu) i jednej składowej fundamentów moralnych (sprawiedliwość). Co więcej, efekt ten nie został potwierdzony replikacyjnie ani sprawdzony w warunku kontrolnym. Próba rozciągnięcia tego wyniku na ogólne stwierdzenie o wpływie muzyki na moralność jest epistemologicznie nieuprawniona i podważa rzetelność naukową interpretacji.

Po drugie, w dyskusji brakuje spójność w modelu teoretycznym i testowanych hipotezach. Doktorantka przedstawia rozprawę jako logicznie powiązaną strukturę eksperymentów badających wpływ muzyki na moralność, ale, jak wykazałem w recenzji badań, brakuje jasnego pomostu koncepcyjnego łączącego wszystkie badania, a w szczególności Badanie 1 (symboliczny wpływ hymnu na sądy deontologiczne) z Badaniem 2a (afektywny wpływ „wesołej” muzyki na tożsamość moralną i fundamenty moralne). Nie zostało przekonująco uzasadnione, dlaczego te dwa podejścia, symboliczne i hedonistyczne, można traktować jako części jednego spójnego programu badawczego.

Po trzecie, choć Doktorantka przywołuje koncepcję „*moralnej triady*” (myślenie, odczuwanie, działanie) oraz różne podejścia do moralności, nie tworzy koherentnego modelu teoretycznego, który byłby testowany w badaniach. Hipotezy są deklaratywne, ale słabo uzasadnione, a ścieżki przyczynowe nie zostały poddane analizie mediacyjnej ani moderacyjnej. Tak więc empiryczna część nie realizuje ambicji teoretycznej nakreślonej w dyskusji ogólnej.

Po czwarte, wnioski w tej sekcji wykraczające poza zebrane dane. Doktorantka stwierdza m.in., że wpływ muzyki jest „*niezależny od emocji*” (w odniesieniu do Badania 2b), że efekt występuje tylko u uczestników z Chin (Badanie 2a), oraz że wynika ze specyfiki muzyki jako bodźca. Tymczasem w Badaniu 2b nie przeprowadzono bezpośredniego porównania bodźców (muzyka vs. obraz), więc twierdzenie o „*specyfice muzyki*” jest logicznie i empirycznie nieuprawnione. Brak grupy kontrolnej w Badaniu 2a uniemożliwia rozróżnienie wpływu muzyki od

wpływu czasu czy habituacji. Hipotezy moderacyjne w Badaniu 1 nie zostały właściwie przetestowane (brak interakcji z poziomem przywiązania narodowego), mimo że stanowiły centralny punkt projektu.

W końcu, Doktorantka przecenia wartość empiryczną rozprawy. Dyskusja ogólna akcentuje techniczne zalety pracy (prerejestracje, duże próby, międzykulturowość), co jest cenne. Jednak wysoka jakość proceduralna nie rekompensuje problemów konceptualnych i analitycznych. Rzetelność narzędzi (np. CNI), niepełna operacjonalizacja zmiennych (np. emocji), a także nieadekwatne testy hipotez (np. brak modeli mediacyjnych czy moderacyjnych tam, gdzie były zapowiedziane) znacznie ograniczają wartość poznawczą wyników.

Podsumowując, dyskusja ogólna zgrabnie syntetyzuje ambicje projektu, ale nie w pełni oddaje jego ograniczenia. Zbyt optymistyczna narracja, nieuwzględniająca metodologicznych niedociągnięć i interpretacyjnych nieścisłości z wcześniejszych rozdziałów, może wprowadzać czytelnika w błąd co do rzeczywistej siły dowodów. Z uwagi na to, że praca aspiruje do nadania stopnia naukowego, należałoby oczekiwać większej konsekwencji między celem, metodą i interpretacją wyników. Sugeruję, by Doktorantka dokonała rewizji języka interpretacyjnego (szczególnie w zakresie wniosków przyczynowych), doprecyzowała model teoretyczny łączący Badania 1 i 2 oraz wyraźniej wskazała ograniczenia wynikające z konstrukcji eksperymentalnej.

### **3. Umiejętność samodzielnego prowadzenia pracy naukowej**

Ocena zdolności do samodzielnego prowadzenia pracy naukowej stanowi jeden z najważniejszych elementów ewaluacji rozprawy doktorskiej. Wymaga ona nie tylko zaprezentowania wyników badań empirycznych, lecz także umiejętności ich samodzielnego zaprojektowania, przemyślenia założeń teoretycznych, dobrania odpowiednich metod analitycznych, wyciągania trafnych wniosków oraz adekwatnego osadzenia całej pracy w ramach aktualnych debat naukowych.

W rozprawie Doktorantka podjęła ambitny problem badawczy dotyczący wpływu muzyki na moralność oraz zaprojektowała dwa badania empiryczne, z czego jedno zostało rozwinięte w serię trzech powiązanych eksperymentów (2a, 2b, 2c). Zgromadziła relatywnie dużą próbę (łącznie ponad 900 uczestników), zastosowała metody eksperymentalne, modele mieszane (LMM), analizę współzmiennych oraz narzędzia mierzące różne aspekty moralności (CNI, MFQ, MIQ). Praca została również zrealizowana w duchu praktyk otwartej nauki, zawiera odwołania do prerejestracji i publicznych repozytoriów danych. Wszystko to świadczy o dużym nakładzie pracy oraz znajomości współczesnych standardów metodologicznych w psychologii empirycznej. Jednocześnie jednak, analiza szczegółowa badań empirycznych oraz interpretacji wyników ujawnia poważne zastrzeżenia, które rzutują na ocenę stopnia samodzielności i dojrzałości badawczej Doktorantki. Wśród najistotniejszych z nich należy wymienić: rozbieżności między modelem teoretycznym a testowanymi analizami, brak podstawowych warunków metodologicznych w kluczowych eksperymentach, nadmiernie optymistyczna interpretacja wyników, problemy ze spójnością programu badawczego, problemy logiczne i brak krytycznej autorefleksji.

W świetle powyższych obserwacji trudno jednoznacznie uznać, że Doktorantka opanowała w pełni i samodzielnie wszystkie etapy procesu badawczego. Z jednej strony zrealizowała złożony projekt badawczy, zapoznała się z literaturą, zgromadziła duże próby i stosowała nowoczesne techniki analityczne. Z drugiej strony liczne uchybienia koncepcyjne, nieadekwatne testy hipotez, nieuprawnione wnioski i niespójność między teorią a analizą prowadzą do ostrożnego wniosku, że mimo formalnej poprawności wykonania projektu, jego opracowanie, zrozumienie i spójność teoretyczna mogą świadczyć o pewnym braku samodzielności lub dojrzałości naukowej Doktorantki.

#### **4. Konkluzja**

Na podstawie szczegółowej analizy rozprawy doktorskiej mgr Katarzyny Pypno-Blajdy stwierdzam, że praca nie spełnia ustawowych kryteriów wymaganych do nadania stopnia doktora, określonych w art. 186 ust. 1 pkt 3 ustawy Prawo o szkolnictwie wyższym i nauce.

Mimo że Doktorantka wykazuje dużą erudycję i znajomość literatury z zakresu psychologii moralności, psychologii emocji oraz badań nad muzyką, wiedza ta nie została przekształcona w spójną i logiczną konstrukcję teoretyczno-empiryczną. Poszczególne rozdziały pracy, teoretyczne, metodologiczne i empiryczne, funkcjonują obok siebie bez wyraźnych powiązań koncepcyjnych. Brakuje modelu teoretycznego, który łączyłby przedstawione pojęcia z zaproponowanym programem badawczym, a w konsekwencji także przejrzystych i wywiedzionych z teorii hipotez.

Część empiryczna opiera się na czterech badaniach (1, 2a, 2b i 2c), które jednak nie tworzą spójnego programu, a ich założenia oraz metodologia nie pozwalają na jednoznaczne wnioskowanie o efektach przyczynowych. W badaniach wielokrotnie brakuje grup kontrolnych, testów porównawczych, jasnego operacjonalizowania zmiennych, a interpretacje wyników często przekraczają zakres danych i są logicznie nieuprawnione. Badania mają raczej charakter eksploracyjny niż confirmacyjny, a ich wartość dowodowa, również ze względu na ograniczenia metodologiczne, jest niewystarczająca, by uzasadnić postawione w pracy tezy. Pominięto także kluczowe alternatywne modele wyjaśniające (np. wpływ nastroju jako mediatora działania muzyki), co podważa rzetelność analizy i projektowania badań.

W mojej ocenie, aby praca mogła spełnić wymagania stawiane rozprawie doktorskiej, należałoby: a) doprecyzować i uzasadnić przyjętą ramę teoretyczną oraz w sposób jasny wyprowadzić z niej hipotezy badawcze; b) przeformułować pytania badawcze i zaprojektować badania w sposób umożliwiający ich jednoznaczne testowanie; c) uwzględnić alternatywne wyjaśnienia i zaprojektować procedury pozwalające je rozróżnić; d) zapewnić spójność między częścią teoretyczną, metodologiczną i empiryczną oraz ograniczyć spekulatywność interpretacji wyników.

W związku z powyższym wnioskuję o niedopuszczenie Doktorantki do dalszych etapów  
przewodu doktorskiego.

Sopot, 10.07.2025

dr hab. Konrad Bocian, prof. USWPS

