



ПСИХОЛОГИЯ ВОСПРИЯТИЯ | PSYCHOLOGY OF PERCEPTION

Научная статья | Original paper

Восприятие бимодальных выражений эмоциональных состояний человека: механизмы интеграции

В.А. Барабанчиков^{1, 2} 

¹ Московский государственный психолого-педагогический университет,
Москва, Российская Федерация

² Московский институт психоанализа, Москва, Российская Федерация
 vladimir.barabanschikov@gmail.com

Резюме

Изучаются закономерности восприятия уни- и мультимодальных выражений аффективных состояний человека. Цель: поиск функциональных механизмов кроссмодальной интеграции голосовых и лицевых образующих бимодальных (голос + лицо) выражений. Согласно гипотезе, тип кроссмодальных отношений в восприятии аффективных состояний человека обусловлен содержанием (категорией) целевой эмоции. Сравнивались особенности восприятия 14 уни- и бимодальных выражений одних и тех же эмоциональных состояний актеров-натурщиков. Исследования проводились на платформе Женевского теста распознавания эмоций — GERT. Эксперимент состоял из трех серий. В первой наблюдателям предъявлялись короткие аудиовидеоролики аффективных состояний натурщиков, во второй — эти же видеоролики, но без звука, в третьей — интонации голоса без видеоизображения. В каждой серии участвовали независимые группы испытуемых: по 72 женщины в возрасте $22-27 \pm 5,6$ года. Требовалось определить эмоциональное состояние натурщика. Анализировались точность идентификаций и динамика структуры категориальных полей. Выявлены пять функциональных механизмов кроссмодальной интеграции, их соответствие содержанию эмоций, валентности и степени возбуждения, а также структуре категориального поля. Описан феномен когерентности унимодальных экспрессий одного и того же аффективного содержания. Раскрыта специфика организации бимодальных категориальных полей. Согласно полученным результатам, процесс кроссмодальной интеграции унимодальных состояний не ограничивается прямым взаимодействием сенсорных систем, в том числе зрительной и акустической. Важным является предметное (категориальное) содержание и логика объединения разномодальных впечатлений о состоянии человека.

Ключевые слова: межличностное восприятие, мультимодальность, бимодальные выражения аффективных состояний, вокальные экспрессии, лицевые экспрессии, точность выражения и идентификации эмоций, структура категориальных полей, кроссмодальная интеграция

Финансирование. Исследование выполнено при поддержке Российского научного фонда, проект № 24-18-00904 «Механизмы восприятия эмоционального состояния человека в процессах неверbalной коммуникации».

Благодарности. Автор благодарит за помощь в организации исследования Е.В. Суворову и А.В. Малионок.



Для цитирования: Барабанчиков, В.А. (2025). Восприятие бимодальных выражений эмоциональных состояний человека: механизмы интеграции. *Экспериментальная психология*, 2025. 18(2), 7–33. <https://doi.org/10.17759/exppsy.2025180201>

Perception of bimodal expressions of human emotional states: mechanisms of integration

V.A. Barabanschikov^{1,2} 

¹ Moscow State University of Psychology & Education, Moscow, Russian Federation

² Moscow Institute of Psychoanalysis, Moscow, Russian Federation

 vladimir.barabanschikov@gmail.com

Abstract

The patterns of perception of uni- and multimodal expressions of human affective states are studied. **Objective:** to search for functional mechanisms of cross-modal integration of vocal and facial forms of bimodal (voice + face) expressions. According to the **hypothesis**, the type of cross-modal relations in the perception of human affective states is determined by the content (category) of the target emotion. The features of perception of 14 unique and bimodal expressions of the same emotional states of actors-sitters were compared. The research was conducted on the platform of the Geneva Emotion Recognition Test – GERT. The experiment consisted of three series. In the first, the observers were presented with short audio videos of the affective states of the sitters, in the second – the same videos, but without sound, in the third – the intonation of the voice without a video image. Independent groups of subjects participated in each series: 72 women aged $22-27 \pm 5.6$. It was necessary to determine the emotional state of the sitter. The accuracy of identifications and the dynamics of the structure of categorical fields were analyzed. Five functional mechanisms of crossmodal integration have been identified, their correspondence to the content of emotions, valence and degree of arousal, as well as the structure of the categorical field. The phenomenon of coherence of unimodal expressions of the same affective content is described. The specifics of the organization of bimodal categorical fields are revealed. According to the **results** the process of cross-modal integration of unimodal states is not limited to the direct interaction of sensory systems, including visual and acoustic. The important thing is the subject (categorical) content and logic of combining diverse impressions about the human condition.

Keywords: interpersonal perception, multimodality, bimodal expressions of affective states, vocal expressions, facial expressions, accuracy of expression and identification of emotions, structure of categorical fields, cross-modal integration

Funding. The work was supported by the Russian Science Foundation, project No. 24-18-00904 “Mechanisms of perception of the emotional state of a person in the processes of non-verbal communication”.

Acknowledgements. The authors are grateful for assistance in data collection E.V. Suvorova and A.V. Malionok.

For citation: Barabanschikov, V.A. (2025). Perception of bimodal expressions of human emotional states: mechanisms of integration. *Experimental Psychology (Russia)*, 18(2), 7–33. (In Russ.). <https://doi.org/10.17759/exppsy.2025180201>

Введение

Поведение человека служит естественным ресурсом информации о его эмоциональном состоянии. В силу разнообразия органов движения и модальностей режимов активно-



сти — локомоций, предметных и коммуникативных действий, поз, жестов, мимики, интонаций голоса — аффективные состояния по-разному проявляются вовне, сохраняя содержательную определенность. Мы выражаем себя и воспринимаемся другими мультимодально, объединяя разномерные планы активности. Совокупность выразительных движений, развернутая во времени, играет роль внешней формы переживания. Ее организация подчиняется предметной детерминации и логике работы функциональных органов, чувствительных к ситуативным изменениям и сдвигам состояния индивида.

Центральным в изучении восприятия мультимодальных эмоциональных состояний является вопрос о кроссмодальных отношениях и механизмах интеграции разномодальной аффективной информации. Как многоплановая активность партнера по коммуникации объединяется наблюдателем в единую конструкцию, становится многомерным образом состояния человека? На площадке каких модальностей собираются подобные конструкты? Все ли модальности равнозначны? Если нет, то почему и как меняются их роли? Как выстраиваются кроссмодальные отношения в ходе межличностных взаимодействий? Зависит ли этот процесс от содержания эмоции, ее валентности, степени возбуждения партнеров и т. п.? В рамках формулируемой темы мы провели эксперименты, в которых сравнили показатели идентификации уни- и бимодальных экспрессий при выражении одних и тех же эмоций. Обнаружено, что голосовые (вокальные) и лицевые (мимические) экспрессии и соответствующие бимодальные (голос + лицо) проявления различаются между собой, точностью оценок и их конфигурацией, влиянием валентности и индивидуальных различий выражения, а также структурой категориальных полей. Обозначен круг возможностей, которыми по отдельности обладают вокальные или лицевые экспрессии, готовые к включению в кроссмодальные интеграции (Барабанщиков, Суворова, 2024; Барабанщиков, Суворова, Малионок, 2024). Данная работа является продолжением выполненных исследований. Ее цель — анализ функциональных механизмов кроссмодальной интеграции выражения и восприятия мультимодальных эмоциональных состояний человека.

Методика

Идея эксперимента состояла в том, чтобы: 1) разложить бимодальные выражения эмоциональных состояний натуралиста на относительно самостоятельные образующие — лицевые и голосовые (вокальные) экспрессии; 2) оценить функциональные возможности каждой из них; 3) соотнести их с оценками бимодального состояния как целого. Подобный подход предполагает сходство условий экспозиции уни- и мультимодальных выражений и соразмерность их идентификаций. Выражение эмоции и ее идентификация являются полюсами единого акта межличностного взаимодействия, благодаря которому экспрессия натуралистов обуславливает содержание воспринятого, а оценки наблюдателей/слушателей характеризуют состав и структуру проявления переживаний. Экспрессивный и информационный потенциалы эмоциональных состояний в данном случае совпадают.

Основу методики составляет Женевский тест распознавания эмоций (Geneva Emotion Recognition Test) — GERT, разработанный для изучения выражения и идентификации мультимодальных динамических эмоциональных состояний и их производных (Scherer, 2005; Schlegel, Scherer, 2017; Schlegel, Grandjean, Scherer, 2014). Особенно важным представляется расширенный набор аудиовидеоизображений ключевых эмоций. Он создавался при участии профессиональных актеров, которые проходили специальную подготовку под руководством известного режиссера. В работе с актерами использовался метод пережива-



ния К.С. Станиславского. Аудиовидеоклипы включали мимику, жесты, движение глаз и головы, а также псевдолингвистические фразы, содержащие интонации выражаемой эмоции.

По сравнению с похожими методиками (MERT, PONS, DANVA и др.) тест включает расширенный набор динамических эмоциональных экспрессий в нескольких модальностях, показывая более высокую точность идентификации. С его помощью эффективно изучалось восприятие эмоций отдельных модальностей (голоса, мимики лица, жестов, поз тела) (Schlegel, Grandjean, Scherer, 2012; Schlegel, Grandjean, Scherer, 2014; Schlegel, Fontaine, Scherer, 2017). Динамический характер экспрессий позволяет выявлять закономерности взаимодействия категорий эмоций в зависимости от степени возбуждения. Исследования динамических паттернов и мимической активности лица (Mortillaro, Mehu, Scherer, 2011; Mehu, Mortillaro, Banziger, Scherer, 2012; Mortillaro, Mehu, Scherer, 2015), а также положений и поз тела (Dael, Mortillaro, Scherer, 2012; Dael, Goudbeek, Scherer, 2013) подтверждают отсутствие четких границ эмоциональных категорий и более сложный механизм восприятия динамических экспрессий по сравнению с фотоизображениями.

Организация стимульного материала, средства его оценки и интерпретация данных построены на концепте «Швейцарского колеса эмоции» — объединении в круг 14 категорий эмоциональных состояний, проранжированных в зависимости от их валентности и степени возбуждения (Bänziger, Mortillaro, Scherer, 2012; Scherer, 1986) (рис. 1). Аффективные категории объединялись в условные группы. В группу А вошли эмоции с положительной валентностью: гордость, радость, развлечение, удовольствие и облегчение. Группы В, С, D образовали родственные эмоциональные состояния с отрицательной валентностью: тревога/страх (В), печаль/горе (С), раздражение/гнев (Д). Амбивалентные эмоции интереса и удивления, а также отвращение рассматривались как независимые.

Наше собственное исследование проводилось в рамках русскоязычной версии GERT, в которой воспроизведены схема и ключевые звенья оригинала (Барабанчиков, Суворова, 2020; 2021; 2022; 2023). Эксперимент состоял из трех серий. В первой участникам предъявлялись короткие аудио/видео ролики аффективных состояний натуралистов, во второй — видеоролики этих же состояний без звука, в третьей — интонации голоса без видеоизображений. Требовалось определить содержание воспринятой эмоции, указав на соответствующую



Рис. 1. Снимок экрана интерактивной версии Швейцарского колеса эмоций



Fig. 1. Screenshot of the interactive version of the Swiss Emotion Wheel

категорию интерактивного изображения Швейцарского колеса. Техническая реализация эксперимента осуществлялась на электронных платформах Vercel, Google Sheets.

Участники исследования. В каждой серии участвовали независимые группы испытуемых: по 72 женщины — студенты и аспиранты московских вузов в возрасте $22–27 \pm 5,6$ года.

Процедура. Во время основной сессии, в зависимости от серии эксперимента, участникам последовательно в случайном порядке предъявлялись 83 аудио-, видео- либо аудиоизоаписи эмоциональных экспрессий десяти актеров (5 мужчин и 5 женщин, средний возраст — 37 лет). Каждая экспозиция демонстрировалась один раз на 3–5 секунд без возможности повтора (рис. 2). При завершении показа на экране дисплея появлялось изображение Колеса эмоций. Переход к следующей экспозиции осуществлялся автоматически после выбора категории и касания на экране соответствующего круга.

Анализ данных. Определялись медианные показатели распознавания предъявляемых стимулов, их отношения и зависимости. При анализе категориальных полей учи-

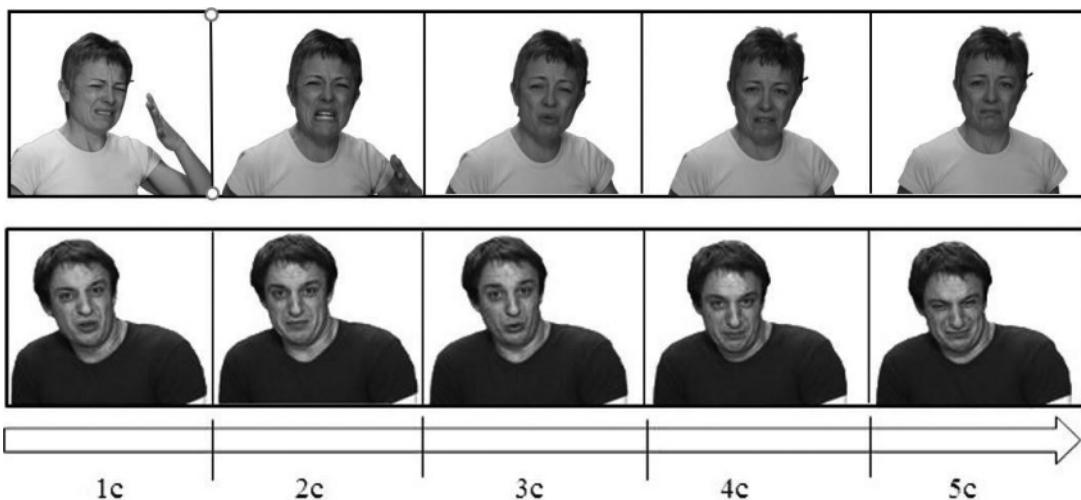


Рис. 2. Раскадровка видеоизображений экспрессии отвращения (база GERT).

Интервал между кадрами около 1 с.

Fig. 2. Storyboard of video images of disgust expression (GERT database).

The interval between frames is about 1 sec.



тывались оценки каждой из категорий с порогом узнавания выше 0,05. Использовался пакет программ SPSS 22.0. В качестве статистических критериев применялись: критерий c^2 Фридмана, критерий Краскела-Уоллиса с уровнем значимости $p < 0,01$, критерий Вилкоксона с уровнем значимости $p < 0,05$, G-критерий Макнемара с уровнем значимости $p < 0,05$.

Более подробно методика и ее производные описаны в работах (Барабанщиков, Суворова, 2020; 2024; Барабанщиков, Суворова, Малионок, 2024).

Основные результаты

Обобщенные оценки (частота выбора) аффективных состояний натуралистов в зависимости от модальности экспрессий и категории эмоций представлены в табл. 1. По горизонтали расположены целевые состояния, по вертикали — ответы испытуемых. Вдоль основной диагонали показаны релевантные оценки, характеризующие точность идентификации. Иррелевантные оценки, расположенные вертикально, указывают на воспринимаемое сходство целевой экспрессии с нецелевыми. Совокупность искомых значений каждого из столбцов отражает структуру категориальных полей целевого аффективного состояния.

Таблица 1 / Table 1

Матрицы оценок уни- и бимодальных экспрессий эмоциональных состояний натуралистов

Matrices of assessments of uni- and bimodal expressions of sitters' emotional states

Состояния / States	Вокальные экспрессии / Vocal expressions													
	Гордость / Pride	Радость / Joy	Развлечение / Amusement	Удовольствие / Pleasure	Облегчение / Relief	Интерес / Interest	Удивление / Surprise	Тревога / Anxiety	Страх / Fear	Горе / Despair	Печаль / Sadness	Отвращение / Disgust	Раздражение / Irritation	Гнев / Anger
Гордость / Pride	0,17	0,02	0,00	0,03	0,02	0,10	0,02	0,01	0,00	0,03	0,02	0,09	0,05	0,03
Радость / Joy	0,08	0,33	0,32	0,03	0,01	0,02	0,03	0,00	0,01	0,03	0,00	0,02	0,01	0,03
Развлечение / Amusement	0,04	0,09	0,42	0,01	0,00	0,02	0,09	0,00	0,00	0,01	0,01	0,02	0,00	0,01
Удовольствие / Pleasure	0,02	0,02	0,01	0,71	0,18	0,15	0,06	0,02	0,00	0,00	0,04	0,10	0,01	0,00
Облегчение / Relief	0,01	0,01	0,00	0,12	0,57	0,13	0,03	0,04	0,00	0,02	0,08	0,05	0,07	0,00
Интерес / Interest	0,15	0,00	0,00	0,03	0,02	0,17	0,33	0,24	0,00	0,00	0,06	0,02	0,15	0,00
Удивление / Surprise	0,17	0,04	0,00	0,00	0,01	0,03	0,33	0,17	0,01	0,01	0,02	0,05	0,10	0,01
Тревога / Anxiety	0,03	0,04	0,00	0,00	0,05	0,03	0,05	0,27	0,22	0,13	0,05	0,05	0,06	0,03
Страх / Fear	0,01	0,08	0,00	0,00	0,01	0,00	0,01	0,06	0,51	0,12	0,02	0,03	0,00	0,01
Горе / Despair	0,01	0,26	0,20	0,01	0,01	0,01	0,00	0,01	0,03	0,34	0,16	0,05	0,00	0,00
Печаль / Sadness	0,02	0,05	0,03	0,03	0,06	0,17	0,03	0,15	0,01	0,20	0,52	0,06	0,04	0,00
Отвращение / Disgust	0,07	0,01	0,00	0,01	0,02	0,05	0,01	0,01	0,02	0,04	0,02	0,22	0,10	0,03
Раздражение / Irritation	0,14	0,01	0,00	0,02	0,02	0,11	0,02	0,01	0,06	0,06	0,02	0,20	0,37	0,36
Гнев / Anger	0,08	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,13	0,01	0,00	0,03	0,04	0,50



Лицевые экспрессии / Facial expressions

Состояния / States	Гордость / Pride	Радость / Joy	Развлечение / Amusement	Удовольствие / Pleasure	Облегчение / Relief	Интерес / Interest	Удивление / Surprise	Тревога / Anxiety	Страх / Fear	Горе / Despair	Печаль / Sadness	Отвращение / Disgust	Раздражение / Irritation	Гнев / Anger
Гордость / Pride	0,63	0,05	0,00	0,15	0,06	0,01	0,00	0,00	0,01	0,00	0,01	0,00	0,01	0,00
Радость / Joy	0,10	0,67	0,41	0,06	0,02	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Развлечение / Amusement	0,04	0,07	0,38	0,06	0,01	0,02	0,01	0,00	0,00	0,00	0,00	0,00	0,01	0,00
Удовольствие / Pleasure	0,11	0,03	0,07	0,53	0,13	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Облегчение / Relief	0,01	0,03	0,02	0,10	0,61	0,01	0,01	0,01	0,00	0,00	0,03	0,00	0,01	0,00
Интерес / Interest	0,02	0,02	0,00	0,08	0,05	0,51	0,07	0,07	0,00	0,00	0,01	0,00	0,03	0,00
Удивление / Surprise	0,00	0,11	0,02	0,00	0,01	0,17	0,44	0,17	0,05	0,00	0,01	0,01	0,04	0,00
Тревога / Anxiety	0,01	0,00	0,01	0,00	0,01	0,06	0,12	0,39	0,19	0,12	0,11	0,03	0,07	0,01
Страх / Fear	0,01	0,00	0,00	0,00	0,00	0,01	0,13	0,21	0,38	0,10	0,02	0,07	0,04	0,01
Горе / Despair	0,00	0,00	0,07	0,00	0,01	0,01	0,01	0,02	0,05	0,48	0,13	0,15	0,04	0,00
Печаль / Sadness	0,00	0,00	0,01	0,00	0,02	0,01	0,01	0,04	0,01	0,14	0,57	0,11	0,01	0,00
Отвращение / Disgust	0,00	0,00	0,00	0,01	0,01	0,03	0,10	0,04	0,00	0,01	0,07	0,55	0,08	0,00
Раздражение / Irritation	0,03	0,01	0,00	0,01	0,05	0,13	0,07	0,04	0,11	0,10	0,02	0,07	0,57	0,29
Гнев / Anger	0,04	0,00	0,00	0,00	0,00	0,01	0,02	0,00	0,20	0,05	0,00	0,01	0,09	0,69

Бимодальные экспрессии / Bimodal expressions

Состояния / States	Гордость / Pride	Радость / Joy	Развлечение / Amusement	Удовольствие / Pleasure	Облегчение / Relief	Интерес / Interest	Удивление / Surprise	Тревога / Anxiety	Страх / Fear	Горе / Despair	Печаль / Sadness	Отвращение / Disgust	Раздражение / Irritation	Гнев / Anger
Гордость / Pride	0,47	0,09	0,00	0,04	0,02	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00
Радость / Joy	0,16	0,76	0,18	0,04	0,02	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Развлечение / Amusement	0,06	0,06	0,68	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Удовольствие / Pleasure	0,10	0,01	0,03	0,79	0,12	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Облегчение / Relief	0,01	0,03	0,01	0,06	0,77	0,02	0,01	0,02	0,00	0,00	0,02	0,00	0,02	0,00
Интерес / Interest	0,09	0,01	0,00	0,05	0,02	0,57	0,11	0,06	0,00	0,00	0,01	0,00	0,06	0,00
Удивление / Surprise	0,01	0,04	0,05	0,00	0,00	0,14	0,65	0,16	0,03	0,00	0,01	0,03	0,05	0,00
Тревога / Anxiety	0,00	0,00	0,00	0,00	0,00	0,02	0,10	0,58	0,12	0,11	0,02	0,03	0,04	0,00
Страх / Fear	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,13	0,60	0,08	0,01	0,02	0,00	0,00
Горе / Despair	0,00	0,00	0,02	0,00	0,01	0,02	0,02	0,02	0,13	0,61	0,22	0,08	0,02	0,01
Печаль / Sadness	0,00	0,00	0,00	0,00	0,00	0,03	0,01	0,03	0,00	0,12	0,66	0,07	0,01	0,00
Отвращение / Disgust	0,01	0,00	0,01	0,00	0,01	0,03	0,03	0,00	0,00	0,02	0,03	0,65	0,08	0,00
Раздражение / Irritation	0,05	0,00	0,00	0,00	0,02	0,12	0,06	0,00	0,03	0,03	0,02	0,10	0,67	0,22
Гнев / Anger	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,08	0,03	0,00	0,01	0,04	0,75

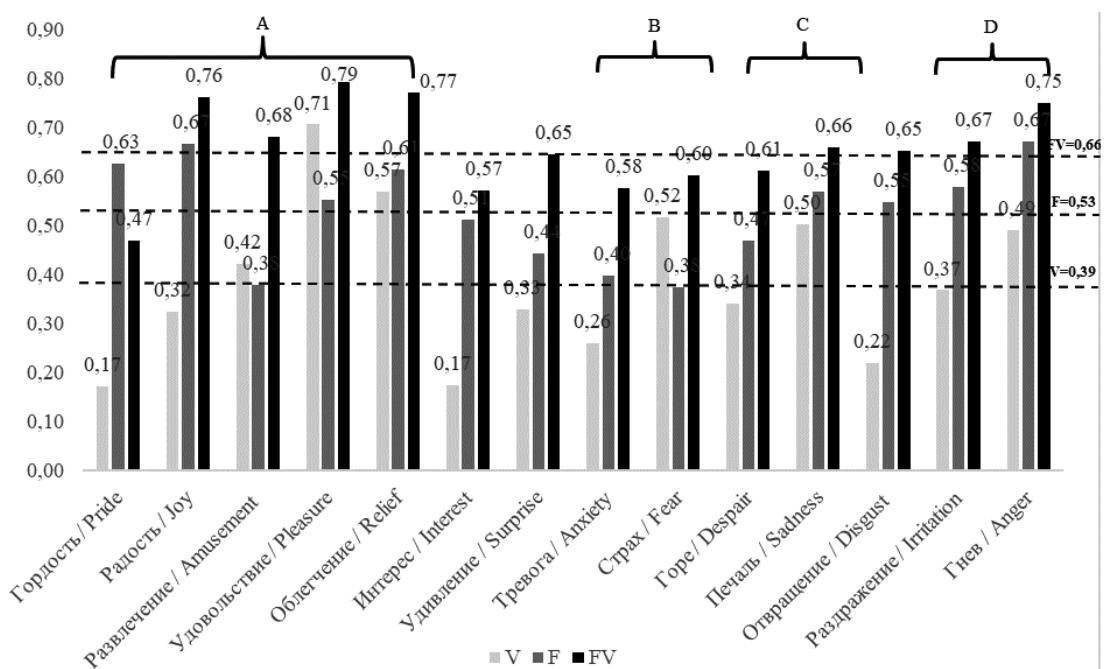


Рис. 3. Частота распределения релевантного выбора вокальных (V), лицевых (F) и бимодальных (FV) экспрессий

Fig. 3. The frequency of distribution of the relevant choice of vocal (V), facial (F) and bimodal (FV) expressions

Точность идентификации уни- и бимодальных экспрессий. Медианная точность распознавания вокальных, лицевых и бимодальных экспрессий показана на рис. 3.

При экспозиции *вокальных экспрессий* лучше всего распознается эмоция удовольствия (0,71), хуже — облегчение (0,57), страх (0,52), печаль (0,5) и гнев (0,49); низко оцениваются проявления тревоги (0,26), отвращения (0,22), гордости (0,17) и интереса (0,17). Общая средняя частота выбора релевантных ответов — 0,39 ± 0,16. Критерий $c^2 = 342,7$, $p < 0,1$.

При экспозиции *лицевых экспрессий* наиболее точно распознаются эмоции гнева (0,67), радости (0,67), гордости (0,63) и облегчения (0,61); наименее точно — удивление (0,44), тревога (0,4), развлечения (0,38) и страх (0,38). Общая средняя частота выбора релевантных ответов — 0,53 ± 0,09. Критерий $c^2 = 178,9$, $p < 0,1$.

При экспозиции *бимодальных экспрессий* хорошо распознаются эмоции удовольствия (0,79), облегчения (0,77), гнева (0,75), радости (0,76); сравнительно плохо — тревога (0,58), интерес (0,57) и гордость (0,47). Общая средняя частота выбора релевантных ответов — 0,66 ± 0,1. Критерий $c^2 = 140,2$, $p < 0,1$.

В группах аффективных состояний с разным уровнем активации (B, C и D) вокальных экспозиций отмечается линейная зависимость точности распознавания от степени возбуждения. Эмоции с более высоким возбуждением распознаются точнее: *тревога* (0,26) и *страх* (0,52), $Z = -3,861$, $p < 0,01$; *горе* (0,34) и *печаль* (0,5), $Z = -5,84$, $p < 0,01$; *раздражение* (0,37) и *гнев* (0,49), $Z = -3,816$, $p < 0,01$. При экспозиции лицевых экспрессий линейная



зависимость обнаружена в группах С и D: *горе* (0,47) и *печаль* (0,57), $Z = -2,184$, $p < 0,05$; *раздражение* (0,58) и *гнев* (0,67), $Z = -2,893$, $p < 0,01$. При совместной экспозиции лица и голоса линейная зависимость проявляется в группе D: *раздражение* (0,67) и *гнев* (0,75), $Z = -2,412$, $p < 0,05$.

Категориальные поля уни- и бимодальных экспрессий. Распределения релевантных и иррелевантных ответов испытуемых представлены на рис. 4. В зависимости от категории эмоций и способа их выражения каждое аффективное состояние идентифицируется по-разному.

На экспозициях бимодальных экспрессий **группы А** ярко выражены ядра *удовольствия* (0,79), *облегчения* (0,77) и *радости* (0,76). Удовольствие сопровождается экспрессиями облег-

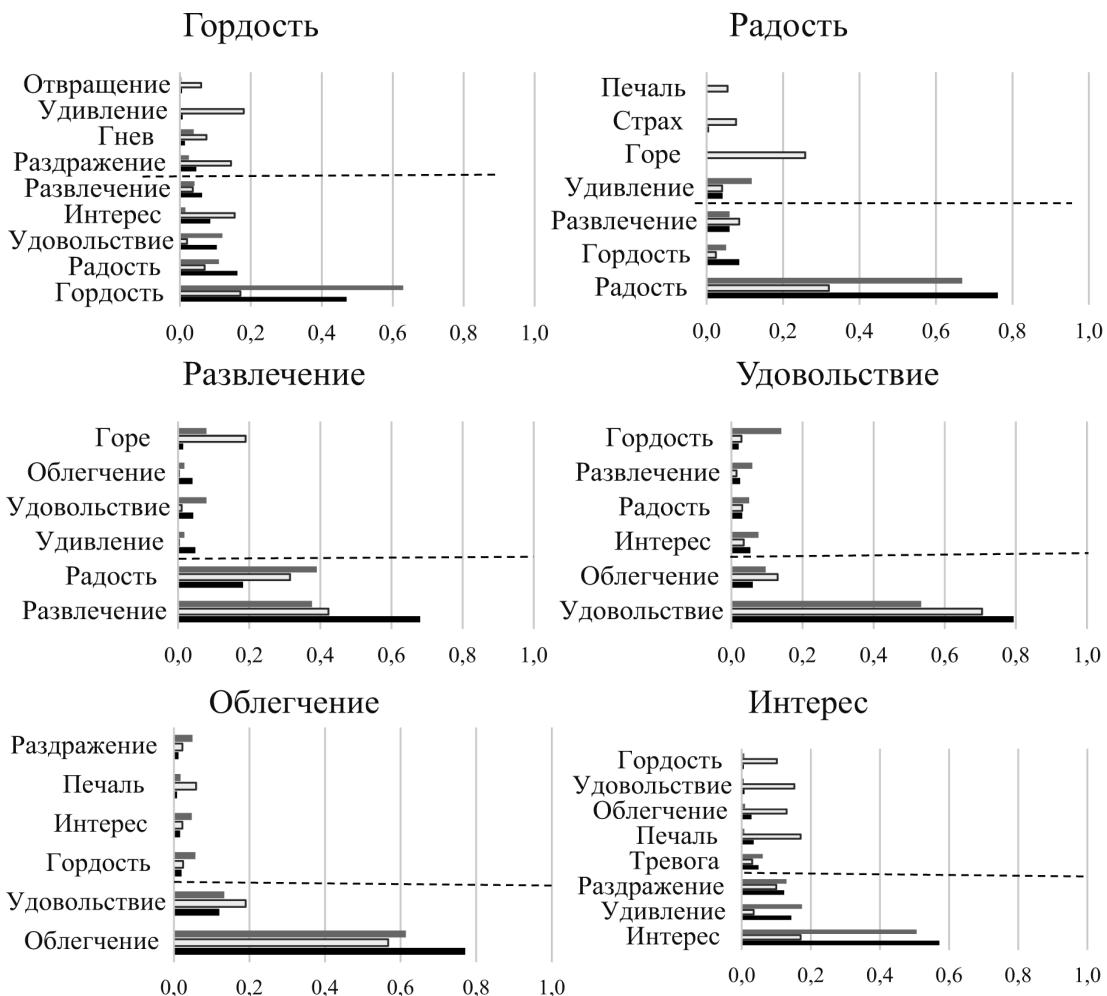


Рис. 4. Структуры категориальных полей уни- и бимодальных экспрессий голоса (V), лица (F) и лица и голоса (FV). Вверху указаны категории экспонируемых эмоций, слева – категории ответов испытуемых, внизу – частота выбора. Пунктирные линии разделяют когерентные (снизу) и некогерентные (сверху) экспрессии.

□ V ■ F ■■ FV

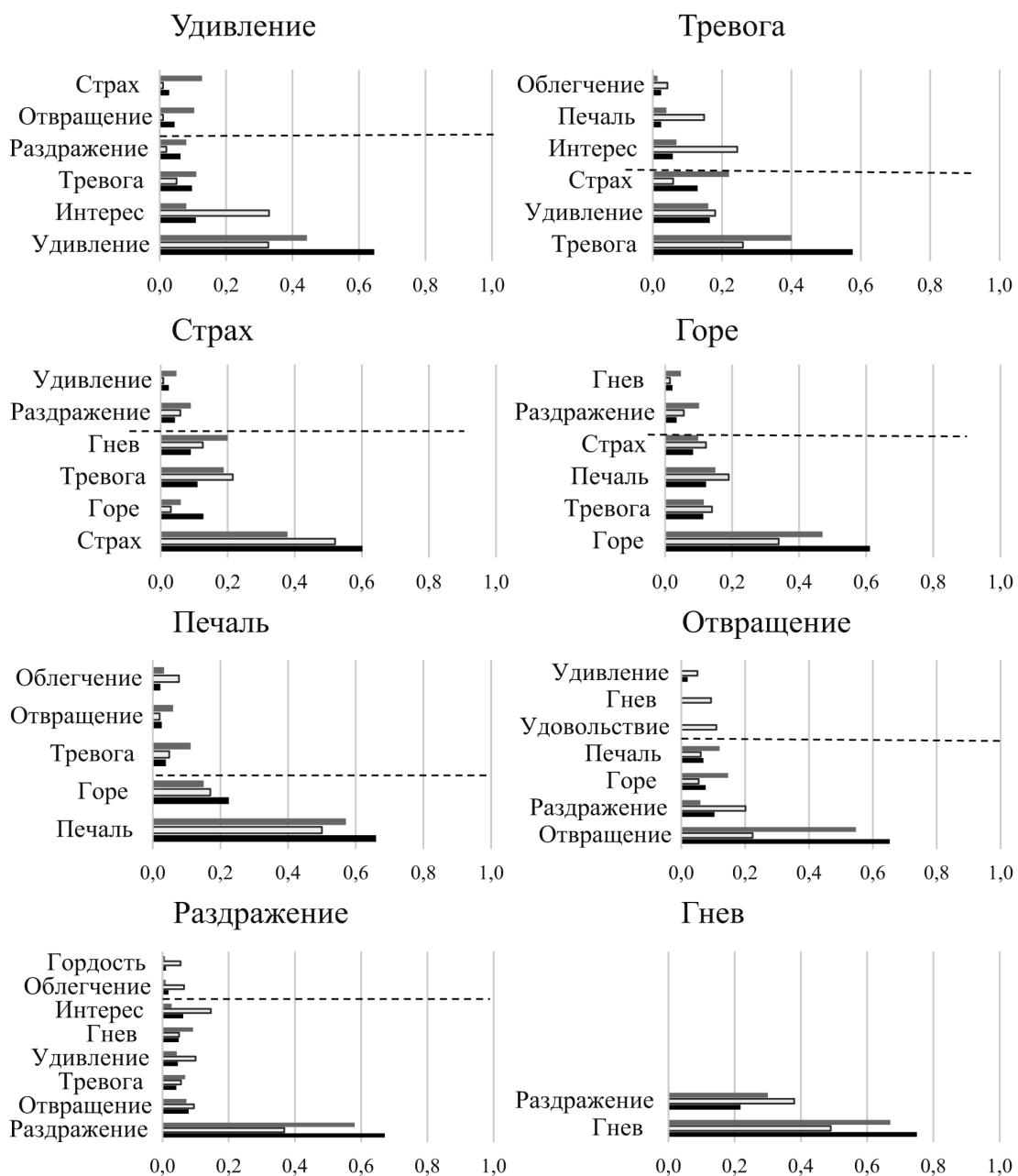


Рис. 4. Продолжение

чения (0,06); облегчение — удовольствием (0,12); радость — гордостью (0,09) и развлечением (0,06); развлечение — радостью (0,18). Наибольшее количество дополнительных экспрессий дает состояния *гордости* — радость (0,16), удовольствие (0,1), интерес (0,09) и развлечение (0,06). В отличие от объединенных и лицевых экспрессий, вокальная экспозиции гордости почти не распознается (0,17); дополнительными компонентами становятся удивление (0,18),

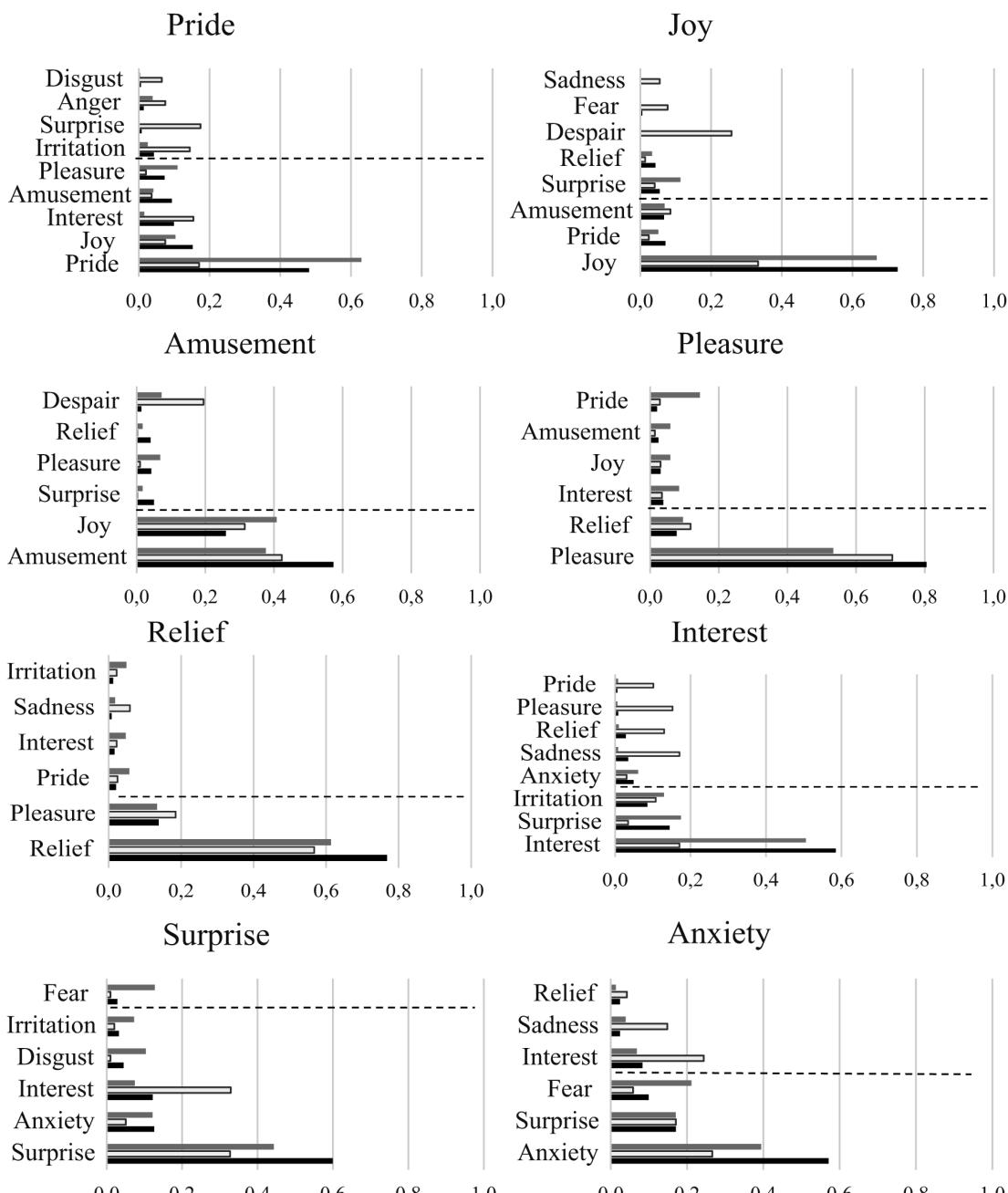


Fig. 4. Structures of categorical fields of uni- and bimodal expressions of voice (V), face (F) and face and voice (FV). The categories of emotions displayed are shown at the top, the categories of responses of the subjects are shown on the left, and the frequency of selection is shown at the bottom. Dotted lines separate coherent (bottom) and incoherent (top) expressions.

□ V ■ F ■■ FV

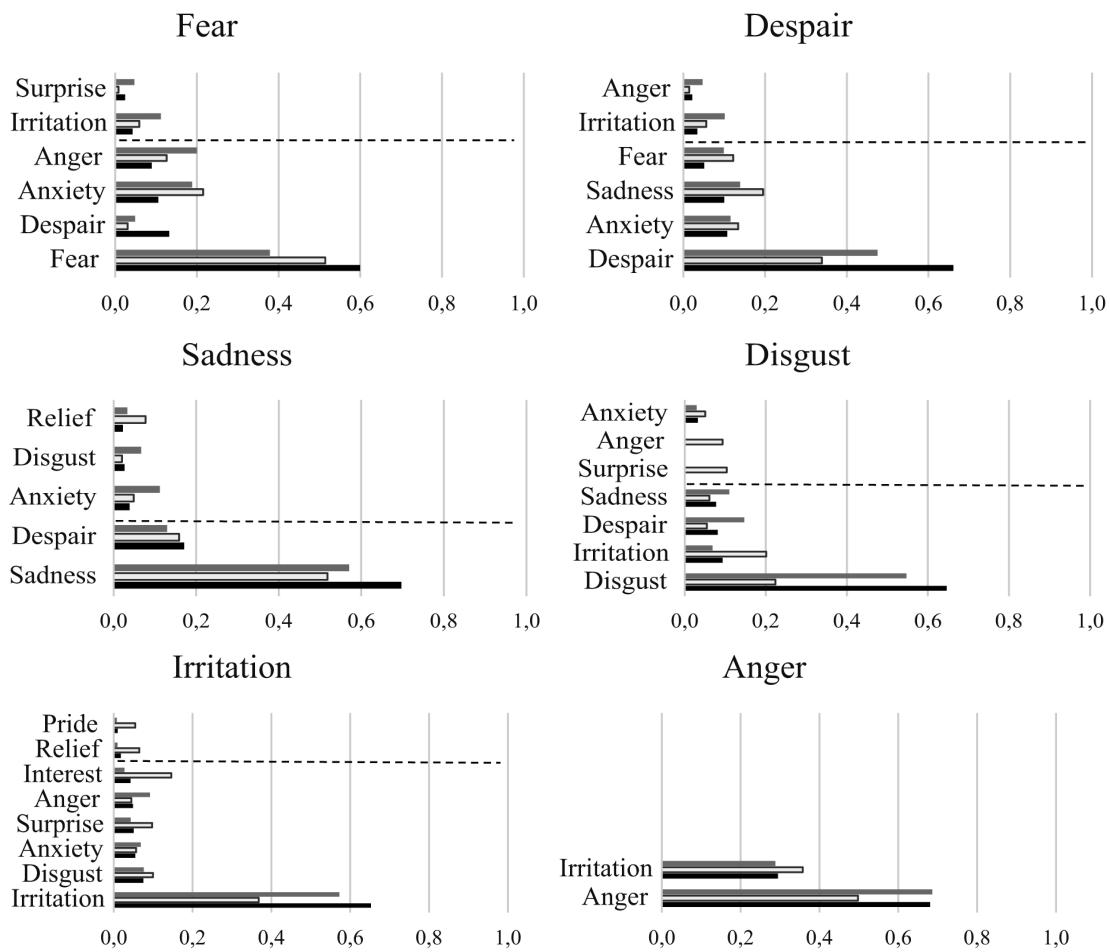


Fig. 4. Continuation

интерес (0,15), раздражение (0,14) и отвращение (0,06). На лицевых экспрессиях гордость распознается лучше и сопровождается радостью (0,11) и удовольствием (0,12).

Оценки вокальных экспрессий радости и развлечения включают эмоции противоположной валентности. Дополнительными компонентами выступают горе (0,26), развлечение (0,09), страх (0,08); на лицевой экспозиции — удивление (0,12) и развлечение (0,06). Оценки вокальных экспрессий развлечения содержат значительную долю радости (0,32) и горя (0,19); лицевых — радости (0,39), удовольствия (0,08) и горя (0,08). При распознавании видеоизображений доля эмоции радости как дополнительного компонента находится на одном уровне с оценками, относящимися к релевантной эмоции развлечения (0,38).

На вокальной экспозиции экспрессия удовольствия дополнительно включает только облегчение (0,12); на лицевой — гордость (0,15), облегчение (0,1), интерес (0,08), радость (0,06) и развлечение (0,06).

Бимодальные экспрессии облегчения включают компоненты удовольствия (0,14), вокальные — удовольствия (0,19) и печали (0,06), лицевые — удовольствия (0,13) и гордости (0,06).



Оценки бимодальных экспрессий *группы В* тесно связаны между собой и с соседними в пространстве Колеса экспрессиями горя и удивления; оценки вокальных и лицевых экспрессий смешены в сторону целевых экспрессий группы Д. На объединенных модальностях *страх* ассоциируется с эмоциями горя (0,13), тревоги (0,12) и гнева (0,09); на вокальных – тревоги (0,22), гнева (0,13) и раздражения (0,06); на лицевых – гнева (0,2), тревоги (0,19), раздражения (0,09) и горя (0,06). При экспозиции бимодальных экспрессий *тревоги* проявляется ее связь с эмоциями страха (0,1), удивления (0,16) и интереса (0,06). Вокальные экспрессии тревоги (0,27) распознаются на одном уровне с интересом (0,24), дополнительными компонентами являются эмоции удивления (0,17), печали (0,15) и страха (0,06), в оценках лицевых экспрессий присутствуют страх (0,21), удивление (0,17) и интерес (0,07).

При экспозиции состояний *группы С* в распознавании объединенных экспрессий выражены связи с родственными эмоциями той же аффективной группы и близкими эмоциями группы В. На бимодальной экспозиции эмоция *печали* дополняется состоянием горя (0,22), на вокальной – горя (0,16), облегчения (0,08), интереса (0,06) и тревоги (0,05); на лицевой – горя (0,13), тревоги (0,11) и отвращения (0,07). Эмоция *горя* в бимодальной экспозиции дополняется состояниями тревоги (0,11), печали (0,12) и страха (0,08); в вокальной – печали (0,2), тревоги (0,13), страха (0,12) и раздражения (0,06); в лицевой – печали (0,14), тревоги (0,12), страха (0,1), раздражения (0,1) и гнева (0,05).

При распознавании экспрессий *группы D* *гнев* во всех трех модальностях содержит раздражение: в контексте объединенных экспрессий гнева – 0,22; в лицевых экспрессиях – 0,29; в вокальных – 0,36. Эмоция *раздражения* при экспозиции бимодальных экспрессий дополняется отвращением (0,08), а также тревогой, удивлением и гневом на уровне случайных значений ($\leq 0,05$). При вокальных экспозициях структура категориальных полей расширяется: включаются эмоции интереса (0,15), отвращения (0,1), удивления (0,1), облегчения (0,07) и тревоги (0,06). Оценки лицевых экспрессий содержат компоненты гнева (0,09), отвращения (0,08) и тревоги (0,07).

При экспозиции бимодальных экспрессий *интерес* дополняется удивлением (0,14) и раздражением (0,12). Вокальные экспрессии ассоциируются с эмоциями печали (0,17), удовольствия (0,15), облегчения (0,13), раздражения (0,11) и гордости (0,1). В число дополнительных компонентов лицевых экспрессий входят удивление (0,17), раздражение (0,13) и тревога (0,06).

Оценки объединенных экспрессий *удивления* включают тревогу (0,11) и интерес (0,11); вокальных – в равной степени удивление и интерес (0,33), а также развлечение (0,09) и удовольствие (0,06). Оценки удивления лицевых экспрессий содержат страх (0,13), тревогу (0,12), отвращение (0,1), раздражение (0,07) и интерес (0,07).

Бимодальные экспрессии *отвращения* (0,22) включают раздражение (0,1), горе (0,08) и печаль (0,07); вокальные – в равной степени связаны с раздражением (0,2) и дополнительно включают удовольствие (0,11), гнев (0,09) и печаль (0,06). Дополнительными компонентами лицевых экспрессий являются горе (0,15), печаль (0,12), раздражение (0,06) и тревога (0,07).

Обсуждение результатов

Методика исследования позволила разложить многомерные выражения эмоциональных состояний натурщиков на лицевые и голосовые образующие, установить экспрессив-



ный потенциал каждой из них и в общем плане соотнести с оценками бимодального выражения как целого. Появилась возможность детальнее рассмотреть кроссмодальные отношения, которые обнаруживаются в созданных условиях.

Кроссмодальные отношения. На рис. 3 представлены гистограммы точности оценок экспонируемого состояния в зависимости от модальности выражения и категории эмоций. По совокупным ответам без учета аффективного содержания сравнительно низкая точность получена на просодических экспрессиях (частота выбора: $0,39 +/ - 0,16$), средняя — на лицевых (частота выбора: $0,52 +/ - 0,09$), высокая — на бимодальных выражениях (частота выбора: $0,63 +/ - 0,01$); различия показателей статистически значимы (критерий Краскела-Уоллиса $H = 19,7$; $p < 0,01$). Фиксируются следующие отношения между модальностями: точность идентификации бимодальных экспрессий (FV) превышает оценки как выражений лица (F), так и интонации голоса (V), демонстрируемых по отдельности ($FV > V$, $FV > F$). И в том, и в другом случае имеет место эффект кроссмодальной фасилитации, но точность идентификации экспрессий лица (F) превышает аналогичные оценки голоса (V): $V < F$. В структуре бимодальных состояний доминирующую роль играет мимика, дополненная динамикой взора и губ натуралиста, а также отдельными проявлениями жестов и поворотами головы. Контраст выразительности содержания и объема признаков унимодальных экспрессий усиливает эффект фасилитации (Dolan, Morris, Gelder, 2001; Gelder, Vroomen, 2000; Schirmer, Adolphs, 2017).

Совокупная точность идентификации унимодальных экспрессий ($V + F$) статистически значимо превышает точность идентификации бимодального состояния: $V + F > FV$. Восприятие объединенного выражения эмоции не сводится к сумме образующих, т. е. *неддитивно*, и использует совокупный экспрессивный потенциал в ограниченном объеме (в среднем около 70%).

Описанные тенденции поддерживаются соотношением коэффициентов конкордации (K_K) — степени согласованности оценок точности аффективных состояний: а) отдельно по интонациям голоса и выражениям лица ($K_{V/F} = -0,03$); б) по голосу и его объединением с мимикой ($K_{V/FV} = 0,69$, $p < 0,01$); в) по экспрессиям лица и их объединением с просодиями ($K_{F/FV} = 0,38$, $p < 0,0176$). Отсутствие корреляций в оценках экспрессий лица и голоса подчеркивает глубокое качественное различие источников информации.

В терминах когнитивной психологии, усредненные бимодальные выражения несут более полную информацию о состоянии воспринимаемого человека, чем унимодальные, причем возможности последних неравнозначны: мимика подвижного лица более эффективна, чем интонации голоса. В общей форме описанный результат соответствует данным, полученными разными авторами в сходных условиях эксперимента (Андреева, 2013; Барабанщиков, Королькова, 2020; Барабанщиков, Малионок, Суворова, 2024; Baart, Vroomen, 2018; Gelder, Vroomen, 2000; Massaro, Egan, 1996; Scherer, Scherer, 2011; Schirmer, Adolphs, 2017 и др.).

Разнородность воспринимаемых состояний. Дифференцированный анализ кроссмодальных отношений, выполненный по каждой целевой категории эмоций, показывает их неоднородность, избирательность и объединенность в группы. Зарегистрированы следующие статистические значимые различия в точности идентификации: при сравнении вокальных и лицевых экспрессий (V/F) — для всех категорий эмоций, кроме *развлечения, облегчения и печали*; при сравнении вокальных и объединенных экспрессий (V/FV) — для всех исследованных эмоций, кроме *удовольствия и страха*; при сравнении лицевых и объединенных экспрессий (F/FV) — для всех эмоций, кроме *радости, гнева и отвращения*.

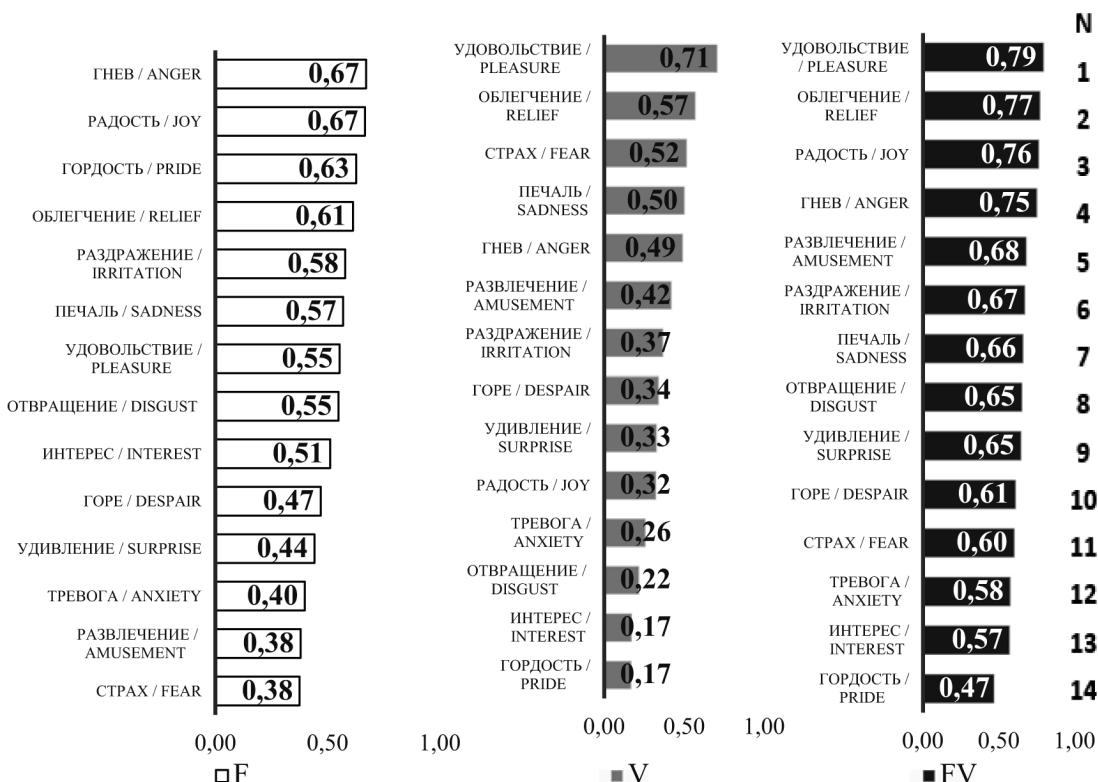


Рис. 5. Рейтинги точности распознавания эмоций по выражениям лица (F), по интонациям голоса (V) и их объединениям (FV); N – позиция воспринятой эмоции в рейтинге

Fig. 5. Ratings of emotion recognition accuracy by facial expressions (F), by intonation of voice (V) and their combinations (FV); N is the position of the perceived emotion in the rating

Весьма показательны распределения релевантных ответов наблюдателей в локальных системах отсчета, упорядочивающих оценки внутри каждой из модальностей. Рейтинги точности идентификации аффективных состояний по интонации голоса, выражению лица и их объединениям представлены на рис. 5. Можно заметить, что они имеют различный порядок следования от наивысшего значения к наименьшему, а одно и то же состояние при разных способах выражения воспринимается по-разному. Например, радость высоко оценивается по лицу, но весьма скромно по голосу. Гордость плохо определяется на слух даже при экспозиции лица, но достаточно точно — по проявлениям мимики. Страх оказывается и в начале рейтинга (интонации голоса), и в его конце (экспрессии лица), и в середине (объединенные экспрессии).

Точность идентификации категорий эмоций при переходе от одного способа выражения к другому меняется (табл. 3). Удивительно, но самые большие изменения в рейтингах (колебания 5–11 пунктов) связаны с аффективной группой А, образованной из эмоций, имеющих положительную валентность (гордость, радость, развлечение, удовольствие), а также с выражениями страха (группа В), амбивалентной эмоции интереса и отвращения. Вместе с тем, облегчение, удивление, тревога и горе практически сохраняют одни и те же позиции в разных списках т. е. не зависят от модальности. Если в первом случае имеет ме-



сто расхождение оценок уни- и бимодальных экспрессий, то во втором — стремление к постоянству, стабильности. Неменьшее разнообразие в сходстве и различии восприятия эмоций проявляется при попарном сравнении рейтингов идентификаций. Непредсказуемые, на первый взгляд, колебания оценок воспринимаемых состояний при изменении способов их выражения указывают на более сложную структуру кроссмодальных взаимодействий, не сводимую к отношениям категориально обезличенных средних по всей выборке.

Механизмы интеграции. Большинство гистограмм, иллюстрирующих точность оценок отдельных категорий эмоций в табл. 2, имеют вид двух- либо трехступенчатой «лесенки», что соответствует общим тенденциям, описанным выше: наименее адекватно воспринимаются отдельно экспонируемые интонации голоса, наиболее адекватно — их объединения с выражениями лица. Иначе идентифицируются гордость, удовольствие и страх. Состояние гордости по выражению лица оценивается точнее, чем по голосу и его объединению с мимикой; «лесенка» превращается в «вершину» ($V < FV < F$). Это единственный пример превосходства экспрессий лица не только над интонациями голоса, но и над их объединением с мимикой. При демонстрации состояний удовольствия и страха экспрессии лица, напротив, имеют наиболее низкие значения, а идентификации удовольствия и страха статистически неразличимы; «лесенка» сменяется «лощиной» ($V = FV > F$). Еще больше значимых отклонений обнаруживают парные сравнения точности оценок мимики лица (F) и интонаций голоса (V), с одной стороны, и экспрессий лица (F) и объединенных выражений (FV) — с другой. В первом случае при унимодальной демонстрации состояний развлечения, облегчения и печали точность идентификации сохраняется и остается ниже оценок объединенных экспрессий ($V = F < FV$); на «лесенке» образуется «нижнее плато». Во втором случае, в отличие от оценок интонаций голоса, состояния интереса, радости, раздражение, гнева и отвращения идентифицируются наблюдателями одинаково высоко ($V < F = FV$); «лесенка» приобретает «верхнее плато». В этой логике остаются лишь три состояния — удивление, тревога и горе, при идентификации которых усредненные отношения модальностей воспроизводятся буквально ($V < F < FV$).

С учетом статистического сходства и различий в распознавании уни- и бимодальных аффективных состояний одних и тех же категорий эмоций можно выделить типы конфигурации точности оценок, каждому из которых соответствует определенный *механизм* кроссмодальной интеграции, связывающий унимодальные экспрессии в единое целое (табл. 2). К типу I относится объединение путем *взаимного дополнения* унимодальных экспрессий в условиях их неравенства, когда больший вклад в оценку эмоции вносит выражение лица. Процесс генерируется на экспрессиях удивления, тревоги и горя. Нередко в публикациях подобный механизм рассматривается как основной или наиболее типичный, хотя по полученным данным это не всегда так. Чаще объединение происходит на основе *доминирования выражений лица* (тип II) и охватывает состояния интереса, радости, раздражения, гнева и отвращения. В этих случаях дополнительная модальность — интонация голоса — выступает в качестве фона, влияние которого на корректность общей оценки минимально. Тип III строится на *паритетете*, при равноправном участии как мимики, так и голоса. Реализуется при экспозиции состояний развлечения, облегчения и печали. В основе механизма типа IV лежит *доминирование интонаций голоса*; влияние выражений лица проявляется слабо. Действие механизма распространяется на состояния удовольствия и страха. Это зеркальное отражение объединений II типа на более узком поле аффективных категорий. Рассмотренные сочетания унимодальных экспрессий вызывают либо эффект фасилита-



ции — значимый рост точности оценок (типы I и III), либо его слабый след, намечающий тенденцию роста (типы II и IV). Из общей логики явно выпадает тип V, связанный с состоянием гордости. Доминированию экспрессий лица противостоит более слабое влияние голоса, достаточное, однако, для снижения точности идентификации бимодальных выражений. Возникает эффект противоположный фасилитации, указывающий на присутствие неконгруэнтного источника аффективной информации.

Во всех рассмотренных эпизодах унимодальные образующие используются в разной степени и ограничены содержанием эмоций. Каждая из них способна выполнять доминирующую, дополняющую, равноценную или сдерживающую функции в зависимости от возможностей альтернативы. Наиболее активную роль играют динамические экспрессии лица. Полученный результат, как правило, ведет к расширению и уточнению аффективной информации, а при наличии неконгруэнтных источников экспрессий вызывает снижение продуктивности восприятия аффективных состояний. Выделенные механизмы могут быть разбиты на подтипы, обладающие более тонкой спецификой. Например, для типа IV это — доминирование интонации голоса, по-разному проявляющееся при положительной (удовольствие) и отрицательной (страх) валентности. Возможна и более глубокая детализация по другим основаниям. Представление о множественности механизмов кроссмодальной интеграции подводит к вопросам организации их совместной работы, условиям актуализации, развития, взаимопереводов и др.

Таблица 2 / Table 2

Конфигурации точности оценок одних и тех же аффективных состояний натуралиста по интонациям его голоса (V), экспрессиям лица (F) и их совместного выражения (FV). А, Б, С, Д — группы состояний, объединенных на основе сходства валентности и уровня активации

Configurations of the accuracy of assessments of the same affective state of the sitter based on the intonation of his voice (V), facial expression (F) and their joint expression (FV). A, B, C, D — groups of states combined based on the similarity of valence and activation level

Тип конфигурации / Configuration type	Опорные аффективные состояния / Basic affective states	Вид на диаграмме / Type on the diagram	Кроссмодальные отношения / Crossmodal relationships	Механизм интеграции / Integration mechanism
I	удивление тревога (B) горе (C) / surprise, anxiety (B) despair (C)	<p>1,00 0,50 -</p> <p>V F FV</p> <p>“трехступенчатая лесенка” / “three-step ladder”</p>	V < F < FV	Взаимодополнение / Complementarity



Тип конфигурации / Configuration type	Опорные аффективные состояния / Basic affective states	Вид на диаграмме / Type on the diagram	Кроссмодальные отношения / Crossmodal relationships	Механизм интеграции / Integration mechanism
II	интерес радость (A) раздражение (D) гнев (D) отвращение / interest joy (A) irritation (D) anger (D) disgust	<p>“верхнее плато” / “upper plateau”</p>	V < F = FV	Доминирование / Dominance
III	развлечение (A) облегчение (A) печаль (C) / entertainment (A) relief (A) sadness (C)	<p>“нижнее плато” / “lower plateau”</p>	V = F < FV	Паритет модальностей / Parity of modalities
IV	удовольствие (A) страх (B) / pleasure (A) fear (B)	<p>“лощина” / “hollow”</p>	F < V = FV	Доминирование интонаций / Dominance of intonation
V	гордость (A) / pride (A)	<p>“вершина” / “top”</p>	V < F > FV	Сдержанное доминирование экспрессий лица / Restrained dominance of facial expressions

Роль валентности эмоций и степени активации. Строгого соответствия механизмов кроссмодальной интеграции определенным группам аффективных состояний, учитываемых методикой GERT, не обнаружено. Объединение унимодальных экспрессий группы А по-разному реализуется посредством механизмов II–IV типов, группы В – I–IV типов,

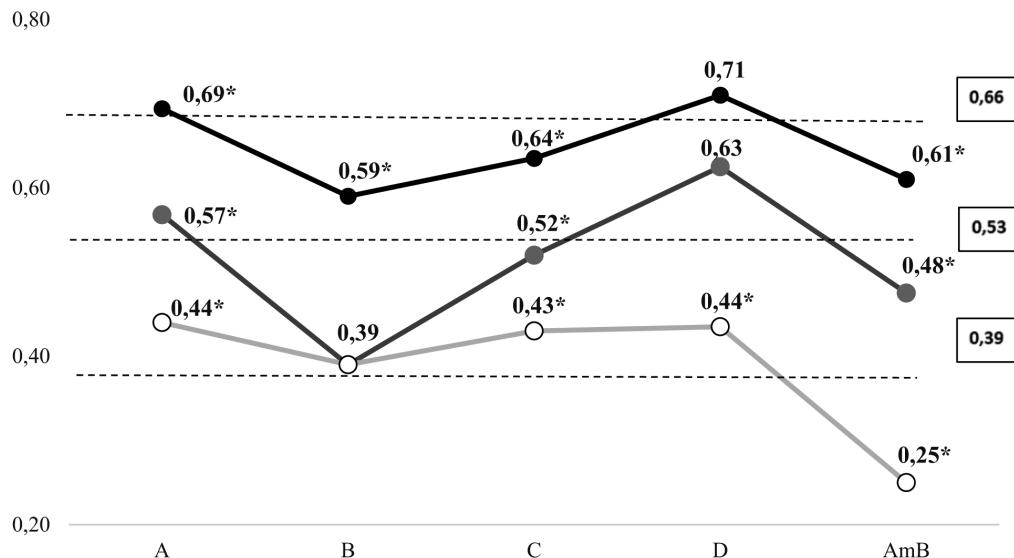


Рис. 6. Средняя точность идентификации аффективных групп вокальных, лицевых и бимодальных экспрессий. A, B, C, D – аффективные группы, Amb – амбивалентные эмоции; цифры сверху – значение частоты ответов; цифры в рамке справа – средние значения по каждой серии, звездочкой отмечены значимые различия.

■ – мультимодальные экспрессии, ■ – лицевые экспрессии, □ – вокальные экспрессии

Fig. 6. Average accuracy of identification of affective groups of vocal, facial, and multimodal expressions. A, B, C, D are affective groups, and Amb – ambivalent emotions; the numbers on top – the response rate; the numbers in the box on the right – the average values for each series, with an asterisk indicating significant differences. ■ – multimodal expressions, ■ – facial expressions, □ – vocal expressions

группы С – II и III типов, группы D – только II типа. По-разному происходит интеграция и амбивалентных состояний: удивления – путем взаимного дополнения (тип I), интереса – через доминирование выражений лица (тип II).

Особенности связи точности распознавания уни- и бимодальных выражений с валентностью эмоций показаны на рис. 6. Релевантные ответы наблюдателей распределены по статистически различным уровням точности и имеют похожие конфигурации. Нижний уровень образуют оценки вокальных экспрессий: для него характерна инвариантность средних значений распознавания как положительных (группа А), так и отрицательных (группы В, С, D) эмоций и резкое падение точности восприятия амбивалентных состояний. Идентификация лицевых и бимодальных экспрессий отличаются более высокими значениями и монотонным ростом оценок негативных состояний: В < С < D; снижение частоты релевантных выборов амбивалентных эмоций происходит более плавно. Усредненные идентификации бимодальных экспозиций отличаются от точности оценок унимодальных во всех аффективных группах, кроме D: для лицевых выражений $U_A = 456$; $U_B = 672,5$; $U_C = 644,5$; $U_{amb} = 401,1$, $p < 0,01$; для вокальных: $U_A = 335$; $U_B = 1096$; $U_C = 1132,5$; $U_{amb} = 244,5$, $U_D = 881,5$, $p < 0,01$. При экспозиции каждой из унимодальных экспрессий в отдельности усредненные оценки группы В совпадают. При экспозиции лицевых и бимодальных экспрессий статистически значимые различия оценок группы D не найдены. Выделяются три тенденции динамики восприятия



уни- и бимодальных состояний в зависимости от валентности: 1) средняя либо высокая точность распознавания положительных эмоций (группа А); 2) инвариантность либо монотонный рост точности оценок негативных состояний (В, С, D); 3) низкая и очень низкая точность оценок амбивалентных эмоций (Amb). Несмотря на очевидную связь способов кроссмодальной интеграции с семантическим содержанием (категорией) эмоций, их прямая зависимость от валентности не просматривается.

Более содержательной выглядит динамика точности идентификации аффективных групп, объединяющих эмоции с различной степенью активации: В (тревога/страх), С (печаль/горе), D (раздражение/гнев). В предшествующих работах (Барабанчиков, Суворова, 2021; 2022) мы обнаружили, что в условиях мультимодальной экспозиции прогнозируемый рост точности оценок с увеличением степени возбуждения действительно имеет место, но только по отношению к состояниям группы D. Этот же результат повторился в текущем исследовании при бимодальных экспозициях. Динамика восприятия унимодальных экспрессий оказалась иной. По интонациям голоса эмоции с высоким уровнем активации групп В и D воспринимаются более точно, а относящиеся к группе С – значимо хуже. По выражениям лица зарегистрирована прямая зависимость корректных оценок от уровня возбуждения групп С и D. Это означает, что в ходе кроссмодальной интеграции воспринимаемые отношения даже родственных эмоций подвергаются трансформациям, которые способны как усилить, так и ослабить то или иное унимодальное состояние, сохранить или копировать его эффективность. Открывается возможность использования оценок эмоции с разным уровнем возбуждения в качестве *индикаторов динамики* процесса кроссмодальной интеграции.

Когерентность унимодальных экспрессий. Как уже отмечалось, при демонстрации одного и того же эмоционального состояния суммарная точность идентификации унимодальных экспрессий превышает точность идентификации бимодального выражения ($V + F > FV$). Полнота аффективной информации в рамках общего целого увеличивается, но не настолько сильно, как это можно было бы ожидать. В процессе кроссмодальной интеграции часть специфических проявлений унимодальных эмоций не используется. Объединенный экспрессивный потенциал теряет до 40%. Неаддитивность отношений имеет место при идентификации каждой категории эмоции. В зависимости от их содержания диапазон «потерь» варьирует от 12% (развлечение) – до 59% (гордость). Все это говорит о том, что точность идентификации унимодальных состояний зависит не только от собственных проявлений разнокачественных экспрессий, но и от их совместности или *когерентности*.

Проявления разномодальных выражений одной и той же эмоции рассматриваются как когерентные, если непосредственно участвуют в кроссмодальной интеграции, а их влияние превышает 5% – порог частоты выбора. Характер совокупности соотношений когерентных и некогерентных экспрессий можно проследить по гистограммам структур категориальных полей уни- и бимодальных состояний (рис. 4). Они разделены пунктирной линией: сверху располагаются оценки некогерентных, снизу – когерентных выражений. На уровне дополнительных эмоций в большинстве случаев некогерентными являются голосовые экспрессии, даже когда показывают относительно высокую частоту выбора. При идентификации, например, радости это горе, страх, печаль; отвращения – гнев и удовольствие. Содержательно преобладают отрицательные либо амбивалентные эмоции. Некогерентные лицевые экспрессии в целом представлены более низкими оценками. При идентификации



радости это удивление; печали — облегчение, отвращение, тревога. Здесь также преобладают отрицательные и амбивалентные состояния.

Плотность когерентных связей обусловлена категорией эмоции (табл. 3). Высокая плотность когерентных связей отмечена в бимодальных экспрессиях тревоги, развлечения, радости, удивления, интереса и горя, низкая « тревоги, гнева, печали и раздражения; предельно низкая — гордости. Среднее значение коэффициента когерентности по оценке точности целевых экспрессий — 73%. Зарегистрирована тенденция роста плотности когерентных связей с увеличением степени возбуждения эмоций группы С и обратная тенденция для группы В.

Таблица 3 / Table 3

Плотность когерентных связей бимодальных (лицо/голос) состояний (K_{cog})
The density of coherent connections of bimodal (face/voice) states (K_{cog})

Эмоция / Emotion	FV	V	F	V + F	K_{cog} , %
Гордость / Pride	0,47	0,17	0,63	0,46*	59%
Радость / Joy	0,76	0,32	0,67	0,99	83%
Развлечение / Amusement	0,68	0,42	0,38	0,80	88%
Удовольствие / Pleasure	0,79	0,71	0,55	1,26	68%
Облегчение / Relief	0,77	0,57	0,61	1,18	75%
Интерес / Interest	0,57	0,17	0,51	0,69	84%
Удивление / Surprise	0,65	0,33	0,44	0,77	85%
Тревога / Anxiety	0,58	0,26	0,40	0,66	62%
Страх / Fear	0,60	0,52	0,38	0,89	71%
Горе / Despair	0,61	0,34	0,47	0,81	85%
Печаль / Sadness	0,66	0,50	0,57	1,07	65%
Отвращение / Disgust	0,65	0,22	0,55	0,77	77%
Раздражение / Irritation	0,67	0,37	0,58	0,95	65%
Гнев / Anger	0,75	0,49	0,67	1,16	63%
Среднее / Average	0,66	0,39	0,53	0,91	73%

Примечание: $K_{cog} = 1 - \frac{(V+F)-VF}{V+F} 100\%$, где V, F, FV — точность идентификации вокальной, лицевой и бимодальной экспрессий.

Note: $K_{cog} = 1 - \frac{(V+F)-VF}{V+F} 100\%$, V, F, FV — accuracy of identification of vocal, facial and bimodal expressions.

Бимодальная структура категориальных полей. Воспринимаемое содержание унимодальных эмоциональных состояний человека представлено структурами категориальных полей, которые наряду с целевой экспрессией — аффективным ядром — включают оценки ассоциированных выражений, имеющих с ней общие признаки. В среднем на каждую категорию приходится 4–5 дополнительных эмоций, хотя в отдельных случаях их может быть семь (гордость) или одна (гнев) (рис. 4). Категориальные поля биполярных состояний отличаются от унимодальных значительным сокращением состава дополнительных признаков: в среднем до 2,28 аффективных категорий на каждую целевую эмоцию. Объем сокращений неравномерен и зависит от валентности. Периферия категориальных



полей аффективной группы А (с учетом идентификации гордости) сужается в 3,2 раза по отношению к совокупному объему категорий, аффективных групп В, С, D – в 1,7 раза, амбивалентных эмоций – в 2,25 раза.

Оценки иррелевантных состояний кратко ниже релевантных и тесно связаны с расположением на круговой шкале Швейцарского колеса относительно целевой категории. Как правило, выбираются состояния смежные или соседствующие с целевыми. Например, для тревоги это страх и удивление (первый уровень сходства), горе и интерес (второй уровень) и т. п. Соотношение оценок предшествующих и следующих за данной эмоцией по часовой стрелке может носить как симметричный, так и асимметричный характер.

Среди особенностей идентификации иррелевантных (дополнительных) состояний целесообразно отметить следующие: 1) чем ближе к целевой категории локализуются экспонируемые эмоции, тем выше вероятность ее выбора (гнев, печаль, тревога); 2) на отдельных участках Колеса средняя частота иррелевантных ответов не зависит от расположения цели (гордость, удивление, горе) и может носить циклический характер: подъем частоты выбора соседних или смежных эмоций, ее спад и снова подъем (тревога, страх, отвращение); 3) для единичных категорий дополнительных эмоций возможен значительный рост частоты ответов на фоне низких показателей смежных и/или соседних состояний (радость, удивление, тревога); 4) динамика оценок дополнительных эмоций по выражениям лица чаще соответствует оценкам бимодальных состояний и противоречит динамике вокальных проявлений. Перечисленные особенности и их сочетания обнаруживаются на всех модальностях аффективных состояний, но в каждой из них имеют свою специфику. В отличие от унимодальных, категориальные поля бимодальных состояний имеют четко выраженный устойчивый анизотропный контур. Различимость сигнала по отношению к шуму (источникам дополнительной аффективной информации) максимальна. Очевидно, что мультимодальные выражения и восприятия эмоциональных состояний человека наиболее полно соответствуют эволюционным и социокультурным требованиям эффективности межличностного взаимодействия и по отношению к унимодальным экспрессиям первичны.

Использование голоса либо лица в качестве основной опоры кроссмодальной интеграции носит функциональный характер. Это касается и целевых, и дополнительных экспрессий (рис. 4). Идентификация демонстрируемого страха как страха (целевое выражение) больше опирается на голосовые интонации. Проявление страха в контексте удивления полностью базируется на экспрессиях лица, а в контексте состояния горя в равной степени на обеих образующих. Наряду с содержанием воспринимаемой эмоции, важным является место, занимаемое ею в мультимодальной структуре категориального поля.

Несмотря на общность состава унимодальных категориальных полей одного и того же аффективного состояния, их структуры качественно различны и по-разному участвуют в процессах *периферической кроссмодальной интеграции*. При разнообразии рассмотренных отношений и наличии статистически значимых различий, они не подходят под типы интеграции, описанные выше. Более того, возможен диссонанс между восприятием целевых и нецелевых экспрессий одной и той же категории.

На рис. 7 показаны соотношения частоты выбора целевого и дополнительного эмоциональных состояний группы D (раздражение/гнев) при экспозициях выражений лица, интонаций голоса и объединенных экспрессий в зависимости от расположения цели. В первых двух фрагментах слева представлены отношения целевой и смежной (1-й уровень сход-



Целевые экспрессии / Target expressions

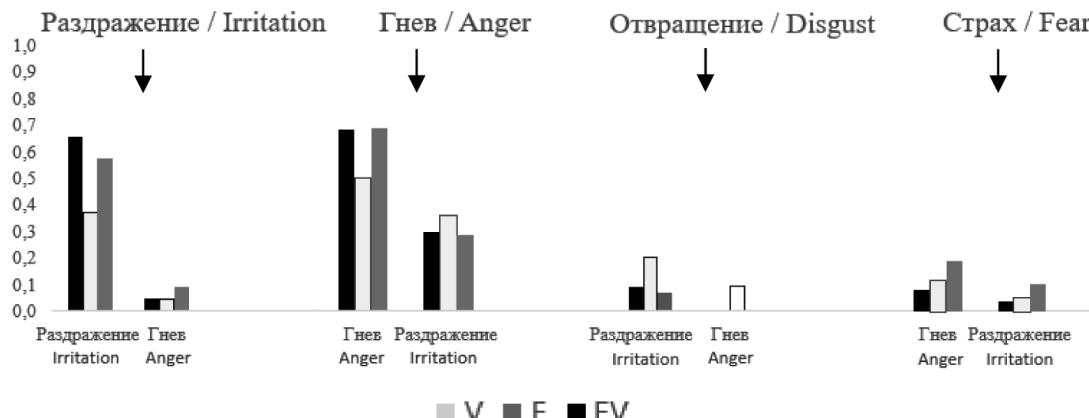


Рис. 7. Частота дополнительных выборов состояний группы D в зависимости от категории целевой экспрессии в условиях уни- и бимодальной экспозиции голоса (V), лица (F) и лица и голоса (FV)

Fig. 7. The frequency of additional selections of group B states depending on the category of target expression under conditions of unique and bimodal exposure of voice (V), face (F) and face and voice (FV)

ства) дополнительной эмоции, в двух последующих — дополнительные эмоции в структуре категориальных полей отвращения и страха. За исключением релевантных идентификаций (совпадения оценок раздражения и гнева с целевыми категориями) интеграция унимодальных экспрессий не усиливает, а ослабляет или исключает бимодальные проявления, а характерная для оценок целевых эмоций зависимость от степени возбуждения пропадает.

Согласно экспериментальным данным мультимодальный образ эмоции каждый раз создается заново. События на периферии категориального поля определяются целевым состоянием. Меняется ядро — меняется и ассоциированное содержание и его отношения на периферии категориальных полей. Действие механизмов кроссмодальной интеграции, обнаруживаемых при идентификации целевых эмоций, на периферии не зарегистрированы. Место сборки и оператор объединенных аффективных состояний соотносимы с ядрами категориальных полей. Дополнительное содержание, за исключением сходства экспрессий первого уровня, воспринимается как фоновое.

Совокупные результаты указывают на высокую избирательность и вариативность кроссмодальной интеграции, гибкое использование разнокачественных идентификационных признаков и динамики их функциональных объединений, задаваемых целевой категорией. Безусловным достижением кроссмодальной интеграции является не только увеличение точности восприятия аффективных состояний, но и упрощение его содержания, снижение уровня общей экспрессивности. Образ эмоции очищается от второстепенных некогерентных проявлений, включая в свою конфигурацию все то, что приводит к оптимально возможному перцептивному результату.

Процесс кроссмодальной интеграции аффективных выражений представляется более дифференцированным и тонким, чем можно было бы ожидать. Он не сводится к прямому взаимодействию сенсорных систем, в том числе зрительной и акустической, хотя и развертывается на их основе. Важным является предметное (категориальное) содержание



экспрессий и логика объединения разномодальных впечатлений о состоянии человека, реализующаяся в процессах коммуникации (Барабанщиков, 2009; 2012а; 2012б; 2016).

Выводы

1. На уровне медианных значений точность идентификаций бимодальных экспрессий эмоциональных состояний натурщиков превышает оценки как выражений лица, так и интонации голоса, демонстрируемых в отдельности, но остается меньше их совокупности. Наиболее активную роль в структуре бимодальных состояний играет мимика лица, дополненная динамикой взора и губ. В целом этот результат соответствует данным, полученным разными авторами в сходных условиях эксперимента. Доминирование выражений лица над интонациями голоса выглядит как универсальный способ объединения разнокачественных образующих.

2. Дифференцированный анализ кроссмодальных отношений, выполненный на уровне категорий эмоций, показал их неоднородность, высокую избирательность и вариативность. Конфигурации точности идентификаций одного и того же состояния при смене модальности меняются. Наибольшие изменения связаны с эмоциями, имеющими положительную валентность. Колебания оценок точности не сводятся к отношениям категориально обезличенных средних, указывая на более сложную структуру межмодальных отношений.

3. Выделены пять типов локальных конфигураций точности идентификации уни- и бимодальных аффективных состояний, зависящие от содержания (категории) эмоций, и соответствующие им механизмы кроссмодальной интеграции: 1) взаимодополнение унимодальных экспрессий; 2) доминирование выражений лица; 3) паритет — равноправное участие как мимики, так и голоса; 4) доминирование интонации; 5) сдержанное доминирование экспрессий лица. Строгого соответствия выявленных механизмов аффективным состояниям разной валентности не обнаружено.

4. В процессе кроссмодальной интеграции часть унимодальных проявлений не используется. В зависимости от содержания эмоции объединенный экспрессивный потенциал теряет до 40%. Идентификацию бимодальных состояний обеспечивают когерентные разномодальные экспрессии одного и того же состояния. Наиболее часто некогерентными оказываются вокальные экспрессии с отрицательной валентностью.

5. Бимодальные категориальные поля отличаются от унимодальных более высокой точностью идентификации целевой эмоции, кратным сокращением состава и интенсивности аффективной периферии. Бимодальный образ эмоции ограничен содержанием, которое ведет к оптимально возможному перцептивному результату. Действия механизмов кроссмодальной интеграции, обнаруживаемые при идентификации целевых эмоций, на периферии категориальных полей не зафиксированы.

Список источников / References

1. Андреева, Е.В. (2013). Феномен кроссмодальных взаимодействий: современное состояние проблемы в зарубежной психологии. *Российский научный журнал*, 1, 219–225.
Andreeva, E.V. (2013). The phenomenon of cross-modal interactions: Current state of the problem in foreign psychology. *Russian Scientific Journal*, 1, 219–225. (In Russ.).
2. Барабанщиков, В.А. (2009). *Восприятие выражений лица*. М.: Институт психологии РАН.
Barabanschikov, V.A. (2009). *Perception of facial expressions*. Moscow: Institute of Psychology, RAS. (In Russ.).



3. Барабанщиков, В.А. (2012а). Психологические механизмы восприятия выражений лица. В: В.А. Барабанщиков, А.А. Демидов, Д.А. Дивеев (Ред.), *Лицо человека как средство общения* (с. 13–31). М.: Когито-Центр.
Barabanshchikov, V.A. (2012a). Psychological mechanisms of facial expression perception. In: V.A. Barabanshchikov, A.A. Demidov, D.A. Diveev (Ed.), *The human face as a means of communication* (pp. 13–31). Moscow: Cogito-Center. (In Russ.).
4. Барабанщиков, В.А. (2012б). *Экспрессии лица и их восприятие*. М.: Институт психологии РАН.
Barabanshchikov, V.A. (2012b). *Facial expressions and their perception*. Moscow: Institute of Psychology, RAS. (In Russ.).
5. Барабанщиков, В.А. (2016). Динамика восприятия выражений лица. М.: Когито-Центр.
Barabanshchikov, V.A. (2016). *Dynamics of facial expression perception*. Moscow: Cogito-Center. (In Russ.).
6. Барабанщиков, В.А., Королькова, О.А. (2020). Восприятие экспрессий «живого» лица. *Экспериментальная психология*, 13(3), 55–73.
Barabanshchikov, V.A., Korolkova, O.A. (2020). Perception of expressions of the “live” face. *Experimental Psychology (Russia)*, 13(3), 55–73. (In Russ.).
7. Барабанщиков, В.А., Суворова, Е.В. (2020). Оценка эмоционального состояния человека по его видеоизображению. *Экспериментальная психология*, 13(4), 4–24. <https://doi.org/10.17759/exppsy.2020130401>
Barabanshchikov, V.A., Suvorova, E.V. (2020). Human Emotional State Assessment Based on a Video Portrayal. *Experimental Psychology (Russia)*, 13(4), 4–24. (In Russ.). <https://doi.org/10.17759/exppsy.2020130401>
8. Барабанщиков, В.А., Суворова, Е.В. (2021). Оценка мультимодальных экспрессий лица в лаборатории и онлайн. В: К.И. Ананьева, В.А. Барабанщиков, А.А. Демидов (Ред.), *Лицо человека в контекстах природы, технологий и культуры* (с. 310–322). М.: Когито-Центр.
Barabanshchikov, V.A., Suvorova, E.V. (2021). Assessment of multimodal facial expressions in the laboratory and online. In: K.I. Ananyeva, V.A. Barabanshchikov, A.A. Demidov (Ed.), *The Human Face in Contexts of Nature, Technology, and Culture* (pp. 310–322). Moscow: Kogito-Tsentr. (In Russ.).
9. Барабанщиков, В.А., Суворова, Е.В. (2022). Гендерный фактор в распознавании эмоционального состояния человека по его аудио-видеоизображениям. *Российский психологический журнал*, 19(2), 6–20. <https://doi.org/10.17759/grj.2022190201>
Barabanshchikov, V.A., Suvorova, E.V. (2022). The gender factor in recognizing a person's emotional state from their audio and video images. *Russian Psychological Journal*, 19(2), 6–20. (In Russ.). <https://doi.org/10.17759/grj.2022190201>
10. Барабанщиков, В.А., Суворова, Е.В. (2022). Индивидуальные формы выражения и идентификация мультимодальных динамических состояний человека. *Познание и переживание*, 3(2), 6–35. <https://doi.org/10.17759/kogito.202203220>
Barabanshchikov, V.A., Suvorova, E.V. (2022). Individual forms of expression and identification of multimodal dynamic human states. *Cognition and Experience*, 3(2), 6–35. (In Russ.). <https://doi.org/10.17759/kogito.202203220>
11. Барабанщиков, В.А., Суворова, Е.В. (2023). Выражение и восприятие мультимодальных эмоциональных состояний. *Национальный психологический журнал*, 17(3), 106–127. <https://doi.org/10.11621/npj.2023.0311>
Barabanshchikov, V.A., Suvorova, E.V. (2023). Expression and perception of multimodal emotional states. *National Psychological Journal*, 17(3), 106–127. (In Russ.). <https://doi.org/10.11621/npj.2023.0311>
12. Барабанщиков, В.А., Суворова, Е.В. (2024). Восприятие подвижного лица как образующей мультимодальных аффективных состояний. *Экспериментальная психология*, 17(4), 4–27. <https://doi.org/10.17759/exppsy.2024170401>
Barabanshchikov, V.A., Suvorova, E.V. (2024). Vivid Face Perception as a Constructive Component of Multimodal Affective States. *Experimental Psychology (Russia)*, 17(4), 4–27. (In Russ.). <https://doi.org/10.17759/exppsy.2024170401>



13. Барабанщиков, В.А., Суворова, Е.В., Малионок, А.В. (2024). Восприятие просодической образующей мультимодальных аффективных состояний. *Экспериментальная психология*, 17(3), 30–51. <https://doi.org/10.17759/exppsy.2024170303>
- Barabanschikov, V.A., Suvorova, E.V., Malionok, A.V. (2024). Perception of the Prosodic Formative of Multimodal Affective States. *Experimental Psychology (Russia)*, 17(3), 30–51. (In Russ.). <https://doi.org/10.17759/exppsy.2024170303>
14. Baart, M., Vroomen, J. (2018). Recalibration of vocal affect by a dynamic face. *Experimental Brain Research*, 236(7), 1911–1918.
15. Bänziger, T., Mortillaro, M., Scherer, K.R. (2012). Introducing the Geneva Multimodal expression corpus for experimental research on emotion perception. *Emotion*, 12(5), 1161. <https://doi.org/10.1037/a0025827>
16. Dael, N., Goudbeek, M., Scherer, K.R. (2013). Perceived gesture dynamics in nonverbal expression of emotion. *Perception*, 42(6), 642–657. <https://doi.org/10.1088/p7364>
17. Dael, N., Mortillaro, M., Scherer, K.R. (2012). Emotion expression in body action and posture. *Emotion*, 12(5), 1085. <https://doi.org/10.1037/a0025737>
18. Dolan, R.J., Morris, J.S., de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proceedings of the National Academy of Sciences*, 98(17), 9465–9470. <https://doi.org/10.1073/pnas.171288598>
19. Gelder, B. de, Vroomen, J. (2000). The Perception of Emotions by Ear and by Eye. *Cognition and Emotion*, 14(3), 289–311. <https://doi.org/10.1080/026999300378824>
20. Massaro, D.W., Egan, P.B. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin & Review*, 3(2), 215–221. <https://doi.org/10.3758/BF03212421>
21. Mehu, M., Mortillaro, M., Banziger, T., Scherer, K.R. (2012). Reliable facial muscle activation enhances recognizability and credibility of emotional expression. *Emotion*, 12(4), 701–715. <https://doi.org/10.1037/a0026717>
22. Mortillaro, M., Mehu, M., Scherer, K.R. (2011). Subtly different positive emotions can be distinguished by their facial expressions. *Social Psychological and Personality Science*, 2(3), 262–271.
23. Scherer, K.R. (1986). Vocal affect expression: a review and a model for future research. *Psychological bulletin*, 99(2), 143–165. <https://doi.org/10.1037/0033-2909.99.2.143>
24. Scherer, K.R. (2005). What are emotions? And how can they be measured? *Social science information*, 44(4), 695–729. <https://doi.org/10.1177/0539018405058216>
25. Scherer, K.R., Scherer, U. (2011). Assessing the ability to recognize facial and vocal expressions of emotion: Construction and validation of the Emotion Recognition Index. *Journal of Nonverbal Behavior*, 35, 305–326.
26. Schirmer, A., Adolphs, R. (2017). Emotion perception from face, voice, and touch: comparisons and convergence. *Trends in cognitive sciences*, 21(3), 216–228.
27. Schlegel, K., Fontaine, J.R., Scherer, K.R. (2017). The nomological network of emotion recognition ability. *European Journal of Psychological Assessment*, 35(3), 352–363
28. Schlegel, K., Grandjean, D., Scherer, K.R. (2012). Emotion recognition: Unidimensional ability or a set of modality-and emotion-specific skills? *Personality and Individual Differences*, 53(1), 16–21.
29. Schlegel, K., Scherer, K.R. (2018). The nomological network of emotion knowledge and emotion understanding in adults: evidence from two new performance-based tests. *Cognition and Emotion*, 32(8), 1514–1530. <https://doi.org/10.1080/02699931.2017.1414687>

Информация об авторе

Владимир Александрович Барабанщиков, доктор психологических наук, профессор, член-корреспондент РАО, директор Института экспериментальной психологии, Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ); профессор, Московский институт психоанализа (НОЧУ ВО МИП), Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-5084-0513>, e-mail: vladimir.barabanschikov@gmail.com



Information about the author

Vladimir A. Barabanschikov, Dr. Sci. (Psychology), Professor, Corresponding Member of Russian Academy of Education, Director, Institute of Experimental Psychology, Moscow State University of Psychology & Education; Professor, Moscow Institute of Psychoanalysis, Moscow, Russian Federation, ORCID: <https://orcid.org/0000-0002-5084-0513>, e-mail: vladimir.barabanschikov@gmail.com

Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.

Conflict of interest

The author declares no conflict of interest.

Декларация об этике

Исследование было рассмотрено и одобрено Этической комиссией Института экспериментальной психологии МГППУ (протокол от 26.02.2025).

Ethics statement

The study was reviewed and approved by the Ethics Commission of the Institute of Experimental Psychology of MSUPE (report 2025/02/26).

Поступила в редакцию 05.05.2025

Received 2025.05.05

Поступила после рецензирования 28.05.2025

Revised 2025.05.28

Принята к публикации 30.05.2025

Accepted 2025.05.30

Опубликована 30.06.2025

Published 2025.06.30