[Geriatronics-Summit] [Paper Submission]

Design evaluation of service robot GARMI Survey in the Pinakothek d. Moderne & Deutsches Museum (project KoBo34)

1st Eva Theresa Jahn
MIRMI - Munich Institute of Robotics and Machine Intelligence
Technical University Munich
Munich, Germany
eva.jahn@tum.de

2nd Abdeldjallil Naceri
MIRMI - Munich Institute of Robotics and Machine Intelligence
Technical University Munich
Munich, Germany
djallil.naceri@tum.de

3rd Luis F. C. Figueredo
MIRMI - Munich Institute of Robotics and Machine Intelligence
Technical University Munich
Munich, Germany
luis.figueredo@tum.de

4th Sami Haddadin
MIRMI - Munich Institute of Robotics and Machine Intelligence
Technical University Munich
Munich, Germany
haddadin@tum.de

***Abstract*— In the KoBo34 project, intuitive interaction with a cooperative service robot for the 3rd and 4th age, we investigated ways to provide support to older people in their daily lives. In addition to integrated capabilities, design plays an important role in the acceptance of a technical devices. To get feedback on the current design of the humanoid robot, we conducted two broad surveys between September 2021 and March 2023. In the first survey, we focused explicitly on perceptions of GARMI; in the second, we compared the results with six selected systems from a dataset of 500 robots. We conducted the data collection in two Munich museums and in the second case also used an online version. We were interested in likeability ratings in general and in the influence of independent variables of age, gender, and experience in care and robotics. The results of the first survey with 250 participants show that GARMI is perceived with a positive tendency. Significant differences by age and experience in care emerge in the evaluation of design, tasks and areas of use. We did not find any significant differences in gender and experience in robotics. This is also the case by the selected robot systems from the second survey with 2712 participants. We use the results for development of next version of robot GARMI and further research on the acceptance of service robots.**

Keywords—survey, design, evaluation, assistive robots, likeability, tasks, age, gender, experience in care and robotics

I. Introduction

Demographic changes are leading to increased research and development in service robotics [1]. The focus is on tasks for support in formal and informal care, in both public and private spaces [2, 3]. Robotic systems, which are developed in the laboratories, concentrate in the first step on technical requirements and their implementation on the hardware and software level [4]. In the second step, the design of the systems is important, which significantly influences the acceptance for use. This is evident not only in the first encounter with the robot, but in the trust and willingness for long-term use on the part of the elderly person or caregiver [5]. When deciding on the design, the chosen solution should match the intended tasks and area of use [6].

In the KoBo34 project, we placed a high value on participatory development. In a broad-based needs analysis with 71 participants and subsequent pilot test with 19 participants, we qualitatively examined the attitudes of all stakeholders involved [7, 8]. In the parallel quantitative study, we were interested in the evaluation of robot GARMI (Fig. 1) by a broad public.

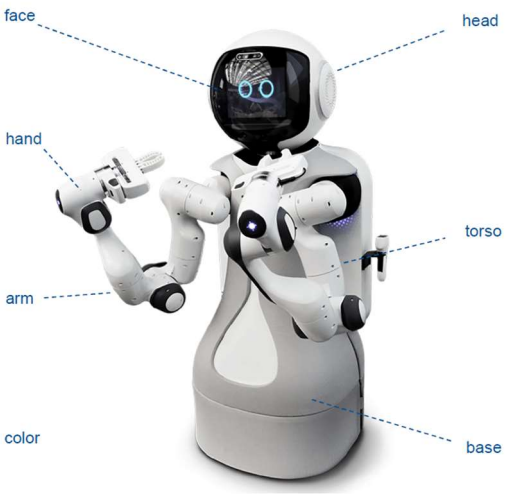

Fig. 1. Robot GARMI, Franka Emika.

Funding agency: BMBF project no. V5ARA202, grant no. 16SV7985
https://www.interaktive-technologien.de/projekte/kobo34

II. MATERIALS AND METHODS

A. Research subject

Our object of investigation is the robot prototype GARMI, which is being developed by Franka Emika in cooperation with Technical University Munich. GARMI is a humanoid robot for safe two-armed manipulation and interaction with humans in direct contact. The robot is 164 cm tall, 57 cm wide, and weighs 120 kg. The body has 16 degrees of freedom, 7 of which are on the arms and 2 on the head. The payload per arm is limited to 3 kg. The robot moves on a platform with 4 wheels and is driven by brushless DC motors. The platform is powered by a butteries with current 4 hours durability.

B. Quantitative survey

For data collection, we prepared two surveys in the field and implemented them in two museums in Munich.

1) Survey GARMI

We conducted the first survey to accompany the KI.ROBOTIK.DESIGN exhibition at the Pinakothek der Moderne from September 2021 to September 2022. The questionnaire in German language consisted of a double-sided sheet (Fig. 2) and included questions on a 5-Likert scale about design by six individual body parts, color and overall impression, four tasks categories as well four areas of use [9].

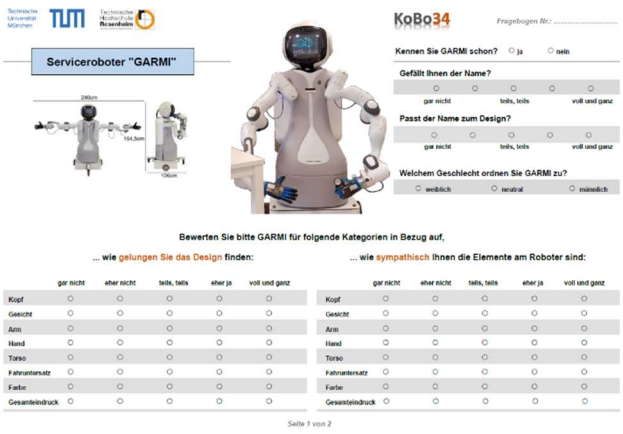

Fig. 2. GARMI design questionnaire.

Body parts included design and likeability of head, face, hand, arm, torso and base. The task categories were:

- **serve** (transport, help im household, opening of doors, pick up ob objects)
- **needs** (food & drinks, hygiene, dressing, massage)
- **mobility** (help with get up, bed and wheelchair transfer)
- **multimedia** (video telephony, internet, alarm function, entertaintment, games, music, reminder).

The four areas of operation are divided into:

- **hospital** (mainly short term care in public space, patients are independent of age)
- **nursing home** (long term care in public space, residents are dependent on age, usually older 65 years)
- **at home** (long term care in private room, users are independent of age)
- **own person** (the willingness to use the robot itself).

In addition, we asked whether the participants already knew Garmi, how they liked the name, whether it matched the design, and with which gender they associated the robot. Finally, we asked participants to enter their demographic information: age, gender, nationality, children, and experience in care as well as in robotics.

2) Survey Robot design

We were able to conduct the second survey at the Robotics exhibition at the Deutsches Museum on several dates from February 2022 to March 2023.

Unlike the first survey, this data collection was supervised and participants could ask questions about the topic. They were also offered candy bars as a reward. The survey in German and English with a total of 500 robots consisted of 10 versions L01 – L10 (Fig. 3) with an average of 50 robots per version, with the exception of the first two variants. The versions were 4 to 5 pages long.

The participants were able to answer six questions per robot on a 5-Likert scale. The first two question "How do you like the robots?" and "Would you want to have these robots yourself?" were aimed at likeability and willingness to use. The other four questions were related to the four tasks of serve, need, mobility, and multimedia, which were also collected in the first survey.

In addition to the paper & pencil, an online version was created on the SoSci survey platform for the L01 and L02 variants with a total of 124 robots[1]. This variant offered Polish and Czech language in addition to German and English. The names of the robots also served as links and were linked to background information about the robots from project sites or developers.

The goal of this second survey was to examine larger contexts in liking the robot design. It was mainly about different structures like stationary, bipedal or wheeled robots, respectively different design styles. From this survey, seven robots, including GARMI were selected and compared in terms of likeability and tasks. Specifically, these were PR2, Tiago, Twendy-One, Toyota HSR, Riba I and Care-O-bot 4. The criteria for selection were one- or two-arm manipulation of objects and suitability for assisting with care tasks.

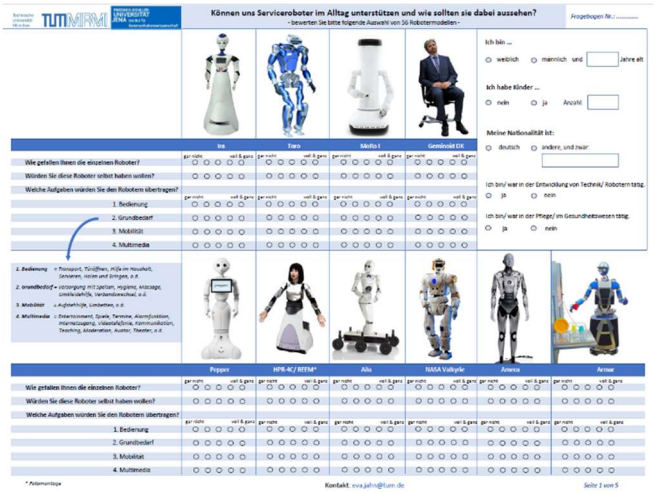

Fig. 3. Robot design questionnaire, version L01 page 1.

[1] https://www.soscisurvey.de/roboterdesign/

III. RESULTS

1) Survey GARMI

A total of 250 participants took part in the first survey, of which 66.7% were female, 33.3% male and 0.0 % diverse. Participants ranged in age from 8 to 90 years, with a mean of 36.65 years and a standard deviation of 20.17 years.

Responses were given on a 5-Likert scale ranging from not at all to fully, with the scale center at 3. For interpretation, the scale was divided into five domains:

- [1.0; 1.8] negative
-]1.8; 2.6] slightly negative
-]2.6; 3.4] average
-]3.4; 4.2] slightly positive
-]4.2; 5.0] positive

The following values apply to the significances:

- $p<0.001$ *** high
- $p<0.01$ ** middle
- $p<0.5$ * low
- $p\geq0.5$ n.s. not significant

We found significant differences by experience in care in the design evaluation of GARMI. Participants with experience rated the individual body parts as well as the overall impression higher than participants without this experience. The best results are achieved head, face, color and hand. For arm and base, the mean values of the group without experience are below the scale center of 3 (Fig. 4).

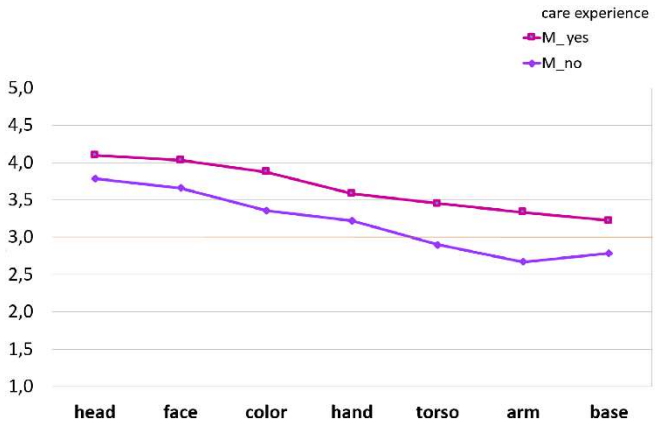

Fig. 4. Likeability to individual body parts according to experience in care.

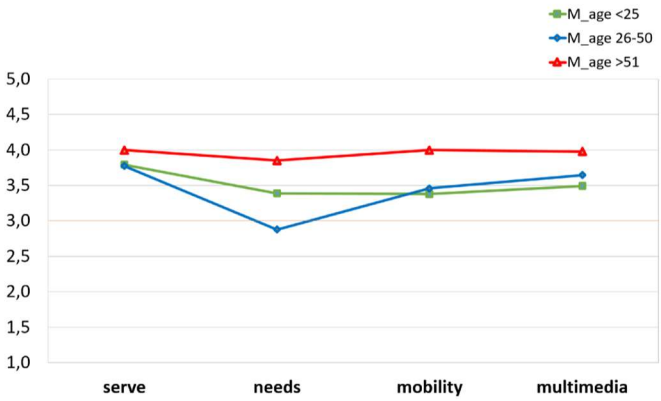

Fig. 5. GARMI acceptance of tasks according to age groups.

The highest acceptance for the tasks is among older people over 51 years, and the lowest among middle-aged people between 26 and 50 years. The tasks serve and multimedia, i.e. those without direct physical contact, have the highest acceptance in the mean value (Fig. 5).

The mean values for the areas of deployment are also highest for older persons over 51 years of age, with nursing home in the first place. The lowest acceptance is for the question of own use (Fig. 6).

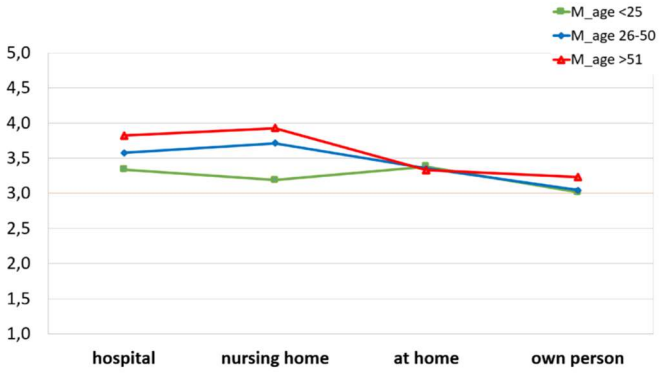

Fig. 6. GARMI acceptance of areas of use according to age groups.

2) Survey Robot design

A total of 2712 participants took part in the second survey, of which 51.9 % were female, 47.3 % male and 0.8 % diverse. Participants ranged in age from 3 to 96 years, with a mean of 28.51 years and a standard deviation of 18.51 years.

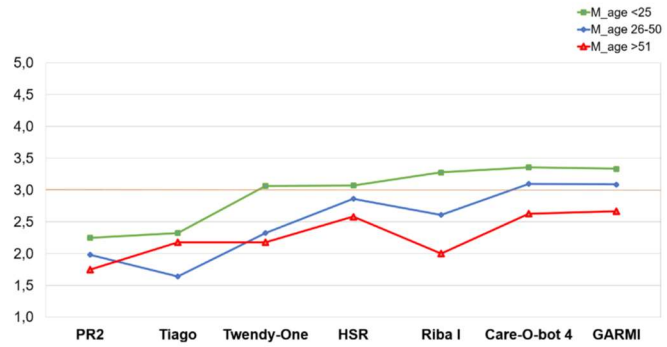

Fig. 7. Likeability to selected service robots according to age groups.

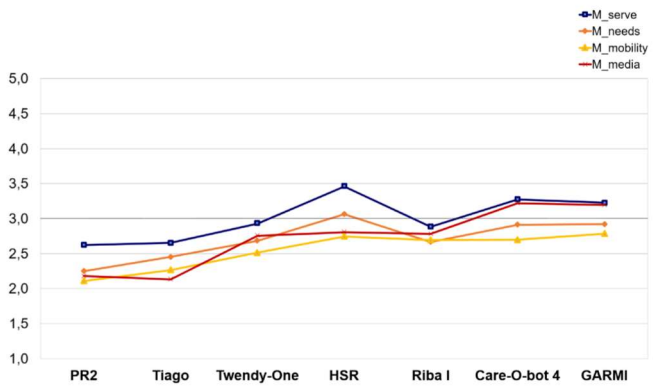

Fig. 8. Acceptance of robot tasks by selected robots.

The robot comparison shows a significant difference in the evaluation by age. The highest likeability values are found among young people up to 25 years of age. The robots are sorted in ascending order according to this age group, with the highest value for GARMI. Only Care-O-bot 4 and GARMI achieve mean values above the scale mean in the age groups up to 25 and between 26 and 50 years. Older people over 51 years of age gave all robots a reserved rating (Fig. 7).

In terms of tasks, serving clearly leads, followed by multimedia. Toyota HSR, Care-O-bot 4 and GARMI achieve the highest mean scores (Fig. 8).

We did not find significant differences for the independent variables of gender and experience in robotics. Data for the graphs Fig. 4 to Fig. 8 are shown as tables in the appendix.

IV. DISKUSSION

The design evaluation of GARMI as well as the 500 robots are based on images. The data on likeability is thus limited to visual perception. Other influencing factors such as size, speed, voice, acoustic, optical signals and, above all, direct interaction play an important role in overall acceptance and would have to be investigated separately.

ACKNOWLEDGMENT

We thank the Pinakothek der Moderne, exhibition “KI.ROBOTIK.DESIGN” as well as the Deutsches Museum, exhibition “Robotik” for their support and intensive cooperation.

REFERENCES

[1] World Health Organization, World report on ageing and health, 2015. https://apps.who.int/iris/handle/10665/186463 *(accessed 03.05.2023)*

[2] M. Lutze, G. Glock, J. Stubbe and D. Paulicke, Digitalisierung und Pflege-bedürftigkeit - Nutzen und Potenziale von Assistenz-technologien, GKV Spitzenverband Berlin, Modellprogramm zur Weiterentwicklung der Pflegeversicherung, Band 15, 2019.

[3] U. Gräske, Intelligente Technik in der Pflege. Von den Chance und Risiken einer Pflege 4.0, Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin 2015.

[4] L. Sullivan: The tall office building artistically considered, 1896.

[5] X., Jin and A. Howard, "The impact of first impressions on human-robot trust during problem-solving scenarios." *2018 27th IEEE international symposium on robot and human interactive communication (RO-MAN)*. IEEE, 2018.

[6] J. Goetz, S, Kiesler and A. Powers, Matching robot appearance and behavior to tasks to improve human-robot cooperation. RO-MAN 2003. The 12th IEEE Internationl Workshop on root and Human Interactive Communication: proceedings, October 31-November 2, Millbrae, California, USA, 2003.

[7] E. T. Jahn, L. F. C. Figueredo, A. Naceri, J. Vorndamme, Ch. Jähne and S. Haddadin, Scenarios to support nursing home residents: development & pilot tests with the humanoid service robot GARMI. Project KoBo34: intuitive interaction with a coooerative service robot for the 3rd and 4th age. Poster, 1th Symposium on Geriatronics, MIRMI Munich Institute of Robotics and Machine Intelligence, Technical University Munich, June 20-June 21, Garmisch-Partenkirchen, 2022.
https://www.researchgate.net/publication/363372962

[8] C. Horstmannshoff, T. Sollfrank, E. T. Jahn and M. Müller, Interaktion potenzieller Nutzer*innengruppen mit einem kooperativen Assistenzroboter für das 3. und 4. Lebensalter (KoBo34) im Rahmen von Testszenarien im Laborkontext. Poster, 23. Jahrestagung des EbM-Netzwerks 2022. Lübeck, 01.-03. September 2022.
https://www.researchgate.net/publication/363335439

[9] E. Jahn, Akzeptanz von Pflegerobotern – Untersuchung eines Konzeptes für die Kranken- und Altenpflege in Deutschland, AV Akademikerverlag, 2015.
https://www.db-thueringen.de/receive/dbt_mods_00025641

APPENDIX

Likeability GARMI	head	face	color	hand	torso	arm	base
Care experte	slightly positive	slightly positive	slightly positive	slightly positive	slightly positive	average	average
M (SD)	**4.10 n.s.** (0.94)	**4.03 n.s.** (0.80)	**3.87*** (1.18)	**3.58 n.s.** (1.09)	**3.45*** (1.09)	**3.33****(1.06)	**3.23 n.s.** (1.18)
participants	N = 31	N = 31	N = 31	N = 31	N = 31	N = 30	N = 31
Care layman	slightly positive	slightly positive	average	slightly positive	average	average	average
M (SD)	**3.79***** (1.07)	**3.66 n.s.** (1.21)	**3.35*** (1.4)	**3.22 n.s.** (1.20)	**2.90*** (1.15)	**2.67**** (1.20)	**2.78 n.s.** (1.21)
participants	N = 188	N = 184	N = 187	N = 188	N = 183	N = 186	N = 185

Tasks	serve	needs	mobility	multimedia
Age >25	slightly positive	average	average	slightly positive
M (SD)	**3.79 n.s.** (1.23)	**3.39*** (1.30)	**3.38*** (1.30)	**3.49 n.s.** (1.38)
participants	N = 116	N = 116	N = 116	N = 116
Age 26-50	slightly positive	average	slightly positive	slightly positive
M (SD)	**3.77 n.s.** (1.24)	**2.88**** (1.23)	**3.46 n.s.** (1.22)	**3.65 n.s.** (1.25)
participants	N = 48	N = 48	N = 48	N = 48
Age <51	slightly positive	slightly positive	slightly positive	slightly positive
M (SD)	**4.00 n.s.** (1.04)	**3.85**** (0.98)	**4.00*** (1.04)	**3.97 n.s.** (1.06)
participants	N = 40	N = 40	N = 40	N = 39

Areas of use	hospital	nursing home	at home	own person
Age >25	average	average	average	average
M (SD)	**3.34 n.s.** (1.30)	**3.19*** (1.36)	**3.38 n.s.** (1.32)	**3.02 n.s.** (1.54)
participants	N = 114	N = 114	N = 114	N = 113
Age 26-50	slightly positive	slightly positive	average	average
M (SD)	**3.58 n.s.** (1.03)	**3.71 n.s.** (1.04)	**3.36 n.s.** (1.30)	**3.04 n.s.** (1.36)
participants	N = 45	N = 45	N = 45	N = 45
Age <51	slightly positive	slightly positive	average	average
M (SD)	**3.83 n.s.** (0.93)	**3.93*** (0.80)	**3.33 n.s.** (1.13)	**3.23 n.s.** (1.35)
participants	N = 40	N = 40	N = 39	N = 39

							
System	**PR2**	**Tiago**	**Twendy-One**	**HSR**	**RIBA I**	**Care-O-Bot 4**	**GARMI**
Producer	Willow Garage	PAL Robotics	Sugano Lab.	Toyota	RIKEN	Fraunhofer IPA	FRANKA EMIKA
Year	2008	2015	2009	2012	2011	2015	2017
Country	US	ES	JP	JP	JP	DE	DE

Likeability	PR2	Tiago	Twendy-One	HSR	RIBA I	Care-O-Bot 4	GARMI
Age >25	slightly negative	slightly negative	average	average	average	average	average
M (SD)	**2.25*** (1.17)	**2.32**** (1.35)	**3.06***** (1.27)	**3.07*** (1.37)	**3.28***** (1.54)	**3.35**** (1.26)	**3.33**** (1.30)
participants	N = 85	N = 142	N = 79	N = 126	N = 105	N = 79	N = 90
Age 26-50	slightly negative	negative	slightly negative	average	average	average	average
M (SD)	**1.98 n.s.** (1,13)	**1.64**** (0.95)	**2.32***** (1.21)	**2.86 n.s.** (1.23)	**2.61*** (1.40)	**3.09 n.s.** (1.22)	**3.09 n.s.** (1.22)
participants	N = 101	N = 55	N = 109	N = 142	N = 51	N = 106	N = 104
Age <51	negative	slightly negative	slightly negative	slightly negative	slightly negative	average	average
M (SD)	**1.75*** (1,01)	**2.18 n.s.** (1.13)	**2.18***** (1.17)	**2.58*** (1.37)	**2.00***** (1.19)	**2.63**** (1.31)	**2.67**** (1.39)
participants	N = 59	N = 17	N = 68	N = 76	N = 28	N = 67	N = 60

Tasks	PR2	Tiago	Twendy-One	HSR	RIBA I	Care-O-Bot 4	GARMI
Serve	average	average	average	slightly positive	average	average	average
M (SD)	**2.63***** (1.46)	**2.65**** (1.44)	**2.93 n.s.** (1.40)	**3.46***** (1.46)	**2.88 n.s.** (1.51)	**3.28**** (1.39)	**3.23*** (1.35)
participants	N = 216	N = 210	N = 229	N = 313	N = 191	N = 229	N = 228
Basic needs	slightly negative	slightly negative	average	average	average	average	average
M (SD)	**2.25***** (1.31)	**2.45***** (1.41)	**2.68***** (1.33)	**3.06 n.s.** (1.46)	**2.67**** (1.48)	**2.91 n.s.** (1.36)	**2.92 n.s.** (1.37)
participants	N = 215	N = 208	N = 229	N = 308	N = 193	N = 230	N = 226
Mobility	slightly negative	slightly negative	slightly negative	average	average	average	average
M (SD)	**2.11***** (1.32)	**2.26***** (1.33)	**2.51***** (1.35)	**2.74**** (1.41)	**2.70**** (1.51)	**2.70**** (1.40)	**2.78*** (1.38)
participants	N = 215	N = 208	N = 230	N = 298	N = 194	N = 233	N = 229
Multimedia	slightly negative	slightly negative	average	average	average	average	average
M (SD)	**2.18***** (1.36)	**2.13***** (1.32)	**2.76**** (1.40)	**2.81*** (1.45)	**2.78*** (1.54)	**3.22*** (1.49)	**3.19*** (1.45)
participants	N = 211	N = 205	N = 231	N = 306	N = 192	N = 231	N = 228