



Atsumu Kubota^{1,2,*}, Mitsuhiko Kimoto^{1,3}, Takamasa Iio^{1,4}, Katsunori Shimohara² and Masahiro Shiomi¹

- ¹ Interaction Science Laboratories, ATR, Seika-cho 619-0237, Japan; kimoto@atr.jp (M.K.); tiio@mail.doshisha.ac.jp (T.I.); m-shiomi@atr.jp (M.S.)
- ² Department of Information Systems Design, Doshisha University, Kyotanabe 610-0321, Japan; kshimoha@mail.doshisha.ac.jp
- ³ Faculty of Science and Technology, Keio University, Yokohama 223-8522, Japan
- ⁴ Faculty of Culture and Information Science, Doshisha University, Kyotanabe 610-0321, Japan
- Correspondence: kubota2019@sil.doshisha.ac.jp

Abstract: This paper addresses the effects of visual reaction times of a turn around behavior toward touch stimulus in the context of perceived naturalness. People essentially prefer a quick and natural reaction time to interaction partners, but appropriate reaction times will change due to the kinds of partners, e.g., humans, computers, and robots. In this study, we investigate two visual reaction times in touch interaction: the time length from the touched timing to the start of a reaction behavior, and the time length of the reaction behavior. We also investigated appropriate reaction times for different beings: three robots (Sota, Nao and Pepper) and humans (male and female). We conducted a websurvey based experiment to investigate natural reaction times for robots and humans, and the results concluded that the best combinations of both reaction times are different between each robot (i.e., among Sota, Nao and Pepper) and the humans (i.e., between male and female). We also compared the effect of using the best combinations for each robot and human to prove the importance of using each appropriate reaction timing for each being. The results suggest that an appropriate reaction time combination investigated from the male model is not ideal for robots, and the combination investigated from the female model is a better choice for robots. Our study also suggests that calibrating parameters for individual robots' behavior design would enable better performances than using parameters of robot behaviors based on observing human-human interaction, although such an approach is a typical method of robot behavior design.

Keywords: reaction time; human-robot interaction; touch interaction

1. Introduction

Designing an appropriate reaction time is one of the essential topics in the humanrobot interaction research field. The sensing capability and recognition processes of social robots are different from those of a human being; therefore, the reaction time of social robots in interaction is also different. However, if such social robots have not considered human reaction time, the people who interact with them will have strange feelings during an interaction. For example, modern industrial robots can react and move much more quickly than humans due to their advanced sensing and actuation systems, but social robots do not necessarily need such features. Rather, social robots are required to operate at the appropriate time in response to people.

Therefore, robotics research has studied various reaction times such as those between interactions [1], emotional and affective expressions [2,3], between actions [4,5], and between conversations [6]. For example, Frederiksen et al. reported that the reaction times of robots in physical interactions influence the affective state of interaction between people, and human-like reaction times have affective impressions upon them [3]. Alhaddad et al. reported that a late response (more than 1 s) might lead to negative influences toward



Citation: Kubota, A.; Kimoto, M.; Iio, T.; Shimohara, K.; Shiomi, M. From When to When: Evaluating Naturalness of Reaction Time via Viewing Turn around Behaviors. *Appl. Sci.* 2021, *11*, 11424. https://doi.org/ 10.3390/app112311424

Academic Editors: Koen Hindriks, An Jacobs and Hoang-Long Cao

Received: 24 September 2021 Accepted: 29 November 2021 Published: 2 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). interacting children [2]. Yamamoto et al. investigated preferred reaction times during conversational greetings and concluded that robots should react within 300 ms [6]. Kanda et al. described the preferred reaction time in gestural reactions and concluded that a robot should delay its reaction for 890 ms to create a more natural feeling in a route-guidance interaction [1]. In conversational interaction, a past study reported that a 1-s reaction time is slightly better than a 0-s reaction time [7]. The knowledge has been applied for social robots to realize natural reaction times to interact with people in real environments [8].

In this research, we focus on the reaction time of a robot when it is touched. Although past studies report the natural reaction time when a person directly touches a robot [4], they have been limited to robots that closely resemble humans, like androids [5]. Moreover, these studies examined the appropriate reaction time using relatively large time resolutions. For example, a past study reported that only a 0-s reaction time is slightly better than a 1-s reaction time [4], but the study did not provide a detailed reaction timing behavior design. In other words, it is not yet clear whether they can be applied to robots with different appearances and sizes, and the resolutions of reaction timing are not precise. Therefore, we conducted a web-based questionnaire survey using several types of robots to find out how the appropriate reaction time changes depending on the robot's appearance and size considering shorter time resolution. This is one of the unique points of this study.

For this purpose, we investigated the most natural reaction time using three robots of different appearances and sizes from humans. Due to the difficulties of using a large number of different robots, we selected typical and commercial robot platforms with different appearances and sizes in Japan: Sota (a desktop-sized robot), Nao (a mediumsized robot), and Pepper (a child-sized robot). Using different kinds of robots to investigate appropriate reaction timing individually is one unique point of this study.

Another unique point of this study is the resolution and type of reaction timing. We shortened the resolution of the reaction behavior and examined two different time lengths in this study. For the resolution, we used a 200-ms resolution of time slices between 0 and 1 s. For the two types of reaction time lengths, we focused on (1) the time length between touch and the start of a reaction behavior (before-reaction time) and (2) the time lengths of the reaction behavior (after-reaction time) (Figure 1). We also verify whether the reaction time model built from one entity can be applied to other entities as an advanced approach.



Figure 1. Before-reaction and after-reaction time concept.

Since the web-based questionnaire survey uses video, the impression of the reaction time may differ from the natural reaction time when the robot is actually touched. However, we note that a turning around behavior is not a tactile stimulus but rather a visual one. Therefore, we believe that our findings will be useful for actual robots, in particular, in a situation when people observe physical interaction between other persons and robots. In fact, recent studies about human-robot physical interaction adopt a video-based web survey to evaluate the impression changes by observing people and robots interacting with each other physically [9]. We believe that a web survey is a possible method for evaluating physical interactions between people and robots. In summary, we address the following research question via an experiment: Are the appropriate reaction times (before-and after-reactions) different for humans and robots?

In addition, we are interested in whether the robots' reaction timing is different due to beliefs, i.e., robot or human. In other words, we investigated the appropriate reaction timing when the appearance and size are the same, but the priming information is different. A past study reported that people who believed that they were interacting with an autonomous robot felt more enjoyment and interacted longer than with a teleoperated robot [10]. In summary, we also address the following research question via an additional experiment:

Are the appropriate reaction times (before-and after-reactions) different due to their beliefs? The organization of the paper is as follows: Section 1 presents the importance and current studies about reaction timing design in human-robot interaction. Section 2 presents materials and methods that are used in experiment I, including detailed explanations about experiment settings. Section 3 shows the analysis results and discussions based on the results of Experiment I, Section 4 presents the details of Experiment II, and Section 5 presents the general discussion. Section 6 provides the conclusion of this study.

2. Materials and Methods of Experiment I

2.1. Factors for Visual Stimulus

This study investigates the differences inappropriate reaction times for three robots as models for visual stimulus: Sota, Nao, and Pepper. We also prepared a male and female person as the visual stimulus of human models to investigate the appropriate reaction times of human beings.

In all videos, models are placed on the right side. For Pepper and human models, only their upper body parts are shown. The experimenter's right-hand touches the model's left shoulder, and the models react by looking to the left. We note that we designed the reaction behavior to be "turning around" as shown in Figure 1. We used a simple rotating behavior because we would like to apply this knowledge for various robots, even though they have fewer DOFs.

In all the videos, we unified the accelerations of the reaction behaviors among the videos to avoid any acceleration effects. The videos did not include any sounds. The length of each video is five seconds, and the file is compressed to decrease total size (less than 1 MB) to avoid delays in online surveys. The width and height of each video are 1280 and 720, respectively, and the fps of each video is 29.97. We confirmed that these settings gave enough values to be notified of the 200 ms differences before the experiment.

2.2. Model

In this study, we used three different robots and two human models. We note that the robot's behaviors are controlled by an operator and adjusted in the video editing because this study aimed to investigate the effects of timing for reaction time, and not a technical approach to realize such behaviors. In fact, it was quite difficult for human models to react with accurate timings.

Sota: Figure 2 shows Sota (VSTONE), which has eight degrees of freedom (DOFs): three DOFs for its head, two DOFs for both arms, and one DOF for its lower body. It is 28 cm tall.



Figure 2. Video stimulus with Sota.

Nao: Figure 3 shows Nao (Aldebaran Robotics), which has 25 degrees of freedom (DOFs): two DOFs for its head, six DOFs for both arms, and 11 DOFs for its lower body. It is 58 cm tall.



Figure 3. Video stimulus with Nao.

Pepper: Figure 4 shows Pepper (Softbank Robotics), which has 20 degrees of freedom (DOFs): two DOFs for its head, six DOFs for both arms, and six DOFs for its lower body. It is 121 cm tall.



Figure 4. Video stimulus with Pepper.

Human: We prepared both a male (Figure 1) and a female (Figure 5). Their heights are about 172 cm and 153 cm.



Figure 5. Video stimulus with a female model.

2.3. Time Length between a Touch and a Reaction

In the video stimulus, after the experimenter's hand touched the left shoulders of the robot models (e.g., Figure 2—left), they reacted to the touch after a certain period (e.g., Figure 2—middle). We investigated the effects of this time length, i.e., the time length between the timing of touch and the start of a reaction as the *before-reaction* time.

Past studies investigated the reaction times to touch stimuli and identified a wide range from 150 to 400 ms [11–14]. If we use excessively short time resolutions (like 50 ms) that require a huge number of conditions, investigating the reaction time effects is not appropriate. When considering the implementation of a specific reaction time to real robots, it may be difficult to achieve a fairly precise frequency because of their servo characteristics.

Based on these considerations, we used time slice resolutions of 200 milliseconds and investigated this time length between 0 to 1 s in 200-ms intervals (0, 200, 400, 600, 800 and 1000 ms). We used 0 s because recent sensing systems such as vision, infrared-based, or laser-based distance sensors will enable robots to accurately estimate touch times in the near future by anticipating the trajectories of people's hands.

This study investigates the differences in appropriate reaction times for three robots as models for visual stimulus: Sota, Nao and Pepper. We also prepared a male and female person as the visual stimulus of human models to investigate the appropriate reaction times of human beings.

2.4. Time Length of a Reaction

In the video stimulus, after the models started their reaction (e.g., Figure 2—middle), a certain time is needed to finish it (e.g., Figure 2—right). We also investigated the effects of the time length, i.e., between the start and the end timings of the reaction as the *after-reaction* time. Similar to the *before-reaction* time, we used time slice resolutions of 200 milliseconds, but we did not use 0 s because it is difficult to finish the reaction behavior within 0 s. Therefore, we investigated time lengths between 0.2 to 1 s in 200-ms intervals: 200, 400, 600, 800 and 1000 ms.

2.5. Procedure

The survey was conducted online by a web page, and a survey company recruited all participants. First, they read an explanation of the data collection and how to evaluate each video via a document. We explained that we were investigating the natural reaction times in response to touch behavior for autonomous robots.

In this data collection, participants observed videos with different reaction times and models. We prepared 30 videos for each model based on the combinations of *before*-and *after*-reaction times. In addition, we prepared two additional videos for each model to check whether participants carefully watched the videos and their answers' qualities. For this purpose, the models of the two videos immediately looked to the left after touching, without delay; i.e., the before-and after-reaction times are zero.

In total, 32 videos were presented with one of the models. The participants watched each video and evaluated the naturalness of the reaction time of the model 32 times. Each video file is relatively small; therefore, participants could watch them after completely downloading them to avoid delays due to network connections.

2.6. Measurements

For evaluation, we prepared only one questionnaire item to decrease the load of the participants; the item is on a 1 to 10 scale, where one is very unnatural, and ten is very natural. Due to a large number of combinations, we decided the order of each video in advance. To avoid similar orders during each video, we sorted the videos by considering before-and after-reaction time to avoid bias evenly. However, the order of videos is the same between participants. Note that the two videos with zero before-and after-reaction times are placed 16th and 32nd. We used the evaluations of these two videos to filter out participants who did not carefully check videos because a past study reported the need for screening participants in a web survey [15]. For this purpose, we only used the data of participants who gave a very unnatural rating for the two videos because such behaviors are impossible actions for all models.

2.7. Participants

In total, 1212 people participated in this survey (Sota: 266, Nao: 242, Pepper: 266, Male: 219, Female: 219). After the filtering process, the valid participant number was 780 (64.4%, Sota: 143, Nao: 137, Pepper: 159, Male: 166, Female: 175).

3. Results and Discussion of Experiment I

3.1. Best Combinations between Before-and After-Reaction Time

Figure 6 and Tables 1–5 show the results of each model. In the tables for each graph, the bold texts mean the highest value in each combination. The figure shows that the best combinations between before-and after-reaction times are different due to the models. As shown in the results, the peak is only one kind in each model. In this case, we do not need to consider complex equations for implementation, except for a visualization purpose of their trends; therefore, we just use the best combinations in each model for comparison. Even though the naturalness scores of other combinations are close to the best combination's value, using the best combinations is a reasonable approach in the context of implementing robots' behaviors. The participants felt more naturalness with relatively slow before-reaction time for the robot models and the female model (400~600 ms) compared to the male models (200 ms). On the other hand, the participants identified more naturalness with a relatively fast after-reaction time for Sota, Nao, and the female model (400~600 ms) compared to Pepper and the male models (800 ms).



Figure 6. Questionnaire results of perceived naturalness. In the table, the bold text with a large font means the highest value in each model. (a) Sota, (b) Nao, (c) Pepper, (d) Male, (e) Female.

		Before					
		0	200	400	600	800	1000
	200	3.66	4.97	5.09	5.11	5.12	4.97
After	400	5.20	7.02	7.01	7.65	6.73	5.59
	600	6.37	7.17	7.52	7.17	6.64	6.02
	800	5.80	6.90	6.90	5.74	5.58	5.24
	1000	5.27	5.87	5.69	4.66	4.40	4.69

Table 1. The perceived naturalness of Sota. The bold text with a large font means the highest value.

Table 2. The perceived naturalness of Nao. The bold text with a large font means the highest value.

			Before					
		0	200	400	600	800	1000	
	200	3.42	3.96	4.04	4.18	3.88	3.89	
	400	4.64	6.39	6.41	7.28	6.15	5.36	
After	600	5.82	7.61	7.86	7.46	6.98	5.88	
	800	5.93	7.32	7.67	6.69	5.75	5.53	
	1000	5.47	6.47	5.96	5.04	4.68	4.79	

		Before						
		0	200	400	600	800	1000	
	200	3.20	4.04	4.02	3.96	3.99	3.84	
	400	4.86	6.23	6.15	6.62	5.82	4.62	
After	600	6.17	7.80	8.11	7.84	6.92	6.04	
	800	6.01	7.82	8.15	7.64	6.77	6.38	
	1000	7.21	6.97	6.82	5.91	5.89	5.31	

Table 3. The perceived naturalness of Pepper. The bold text with a large font means the highest value.

Table 4. The perceived naturalness of the male model. The bold text with a large font means the highest value.

			Before					
		0	200	400	600	800	1000	
After	200	2.37	2.69	2.71	2.29	3.34	2.66	
	400	5.23	6.22	6.86	6.45	5.66	4.74	
	600	7.88	8.44	8.65	8.06	7.36	6.23	
	800	7.80	9.06	8.72	8.08	7.13	7.03	
	1000	8.51	8.62	7.92	7.36	6.73	6.56	

Table 5. The perceived naturalness of the female model. The bold text with a large font means the highest value.

		Before					
		0	200	400	600	800	1000
	200	2.44	2.92	2.51	2.29	2.52	2.41
	400	5.21	6.11	5.78	6.12	4.93	4.33
After	600	7.40	8.21	8.66	8.06	7.49	6.19
	800	7.78	8.61	8.57	7.85	7.08	6.70
	1000	7.72	8.59	7.90	7.21	6.65	6.49

For reference, we conducted a 3-factor ANOVA with perceived naturalness as a dependent variable with before-reaction, after-reaction times, and the model. The results showed that there is a significant effect on all factors. There were significant differences in the before-reaction time (*F* (5, 3870) = 231.735, *p* < 0.001, *partial* η^2 = 0.230), in the after-reaction time (*F* (4, 3096) = 1112.153, *p* < 0.001, *partial* η^2 = 0.590), in the model (*F* (4, 774) = 5.120, *p* < 0.001, *partial* η^2 = 0.026), in the simple interaction effects between before-reaction and after-reaction times (*F* (20, 15,480) = 53.198, *p* < 0.001, *partial* η^2 = 0.064), in the simple interaction effects between before-reaction and model (*F* (16, 15,480) = 74.044, *p* < 0.001, *partial* η^2 = 0.277), and in the two-way interaction effect (*F* (80, 15,480) = 3.127, *p* < 0.001, *partial* η^2 = 0.016). Due to a large number of combinations, we did not describe the details of interaction effects.

3.2. Is the Human Model Applicable to Robots?

To prove the importance of appropriate reaction timing for different models, we compared the best combinations of before-and after-reaction times between the models. If an appropriate combination of reaction times is not compatible between different existences, robotics researchers should use different combinations. In other words, it would not

be better to directly use the extracted parameters based on the observations of human behaviors to robot behavior designs.

Figure 7 shows the questionnaire results of the participant's feelings about the naturalness of each model with each best combination. For Sota, we conducted a repeated measure ANOVA, and the results showed a significant difference (F (3, 426) = 7.437, p < 0.001, *partial* η^2 = 0.050). Multiple comparisons with the Bonferroni method showed significant differences: Sota's timing > Male's timing (p = 0.003), and Pepper's timing (p = 0.02). We found no significant differences between Sota's timing and Nao's and Female's timing (p = 1.000).



Figure 7. Comparison with other best combination timing for Sota, Nao and Pepper. (**a**) naturalness of Sota, (**b**) naturalness of Nao, (**c**) naturalness of Pepper. * p < 0.1.

For Nao, we conducted a repeated measure ANOVA, and the results showed a significant difference (F (3, 408) = 4.385, p = 0.005, *partial* η^2 = 0.031). Multiple comparisons with the Bonferroni method showed significant differences: Nao's and Female's timing > Sota's timing (p = 0.002) and Male's timing (p = 0.037). We found no significant differences with Pepper's timing (p = 1.000).

For Pepper, we conducted a repeated measure ANOVA, and the results showed a significant difference (*F* (3, 474) = 27.935, *p* < 0.001, *partial* η^2 = 0.150). Multiple comparisons with the Bonferroni method showed significant differences: Pepper's > Sota's timing (*p* < 0.001). We found no significant differences between Peeper's timing and Male's timing (*p* = 0.346) as well as Nao's and Female's timing (*p* = 1.000).

These results showed that each robot's best timing combinations are different between robot models and human models. On the other hand, the results showed that the best combinations of the female model did not show significant differences compared to all robot models. We note that there is no evidence of absence; therefore, the results could not claim to show clear advantages of the female model's best combination; they only showed the disadvantages of using the best combination of the male model. Therefore, we conducted a non-inferiority analysis for Sota's data and Pepper's data to test whether applying the female's best combination for the reaction time design is reasonable or not as an additional analysis. The non-inferiority margin was set at 10%, 1 sided test, $\alpha = 0.025$.

The results showed that the female's best combination (mean: 7.65, Confidence Interval (CI): 7.28 to 8.29) was non-inferior to the Sota's best combination (mean: 7.52, CI: 7.15 to 8.16) within the Sota's data (p < 0.001). In addition, the female's best combination (mean: 8.15, CI: 7.78 to 8.61) was non-inferior to Pepper's best combination (mean: 8.11, CI: 7.74 to 8.57) within Pepper's data (p < 0.001). Therefore, using the female's best combination would be not the best choice, but a better choice for reaction time design.

3.3. Summary and Implication for Designing the Reaction Time

The experiment results showed that the best combination of the male model showed significantly lower values for the naturalness of Sota and Nao. In addition, using different robot's best combinations also showed lower values for Sota. On the other hand, the best combination of the female model, which is the same as Nao, did not show significant

differences compared to the other models; additional analysis showed non-inferior results between the female's best combination and the sota as well as pepper's best combination.

Based on these results, we suggest two design guidelines for designing the reaction time for social robots when touching: (1) Measure the best combination of before-and after-reaction time for a newly designed robot because the best values are different between robots. (2) If it is difficult, applying the best combination from the female's model (i.e., 400 ms for before-reaction and 600 ms for after-reaction) would make better impressions because using the best combination from the male model showed clear disadvantages for some of the robots in this study.

3.4. Mathematical Modeling of the Reaction Time

We analyzed the collected data by surface fitting with a polynomial spline (Equation (1), *Before* is the before-reaction time value, and *After* is the after-reaction time value). We used the Akaike Information Criterion (AIC) to adjust each model's likelihood. Figure 8 and Tables 6–10 shows surface fitting results with the averaged data points of the naturalness of each combination and the parameters. As shown in the figures, the slope changes depending on the model. The figures also showed that the peak values shift toward slower after-reaction time by comparing Sota, Nao, Pepper and Human.

$$Naturalness = a + b * time (Before) + c * time (After) + d * time (Before)^{2} + f * time (After)^{2} + g * time (Before) * time (After)$$
(1)



Figure 8. Fitting results of reaction times. (a) Sota, (b) Nao, (c) Pepper, (d) Male, (e) Female.

Table 6. Coefficients for the fitting fuction of Sota.

a	1.408
b	5.564
С	15.08
d	-4.152
f	-2.883
g	-11.36

a	-0.413
b	6.116
с	19.7
d	-4.988
f	-2.52
g	-13.93

Table 7. Coefficients for the fitting fuction of Nao.

Table 8. Coefficients for the fitting fuction of Pepper.

a	-0.371
b	5.8
С	18.86
d	-4.894
f	-2.37
g	-12.05

Table 9. Coefficients for the fitting fuction of the male model.

a	-2.432
b	4.165
С	26.38
d	-3.439
f	-3.113
g	-15.69

Table 10. Coefficients for the fitting fuction of the female model.

a	-2.233
b	3.435
с	25.49
d	-3.532
f	-1.807
g	-15.41

4. Materials and Methods of the Experiment II

In this study, we compared three robots and two human models, and the results showed that their appropriate reaction timings are different; they have different appearances and sizes. When the appearance and size are the same, we wanted to establish whether their reaction timing was different due to beliefs, i.e., robot or human. Based on these considerations, we conducted an additional web survey to compare the effects of beliefs.

4.1. Factors for Visual Stimulus, Models, Procedure, Measurements, and Participants

In the additional survey, we used the same video stimuli with human models only. The difference is the videos' instructions. Here, the explanation of the data collection in the android setting included the priming information to persuade the participants: "We are investigating the appropriate reaction times of androids." The survey was also conducted online by a web page, and the same survey company recruited all participants. We excluded the participants who joined Experiment I.

Finally, they filled out questionnaires about the naturalness of each video on a 1 to 10 scale, where 1 is very unnatural, and 10 is very natural. Thus, we used the same measurement from experiment I.

In total, 404 people participated this survey (Male: 202, Female: 202). After a filtering process, the number of valid participants was 316 (78.2%, Male: 153, Female: 163). Figure 8 shows the results of each model.

4.2. Results

As shown in the Figure 9 and the Tables 11 and 12, the best combinations for the male model are different between humans and androids, but the female models are the same. Therefore, we conducted statistical analysis with ANOVA (two factors: belief (human or android) and time combination (200-before/800-after and 400-before/800-after)), but it did not show any significant differences (*F* (1, 471) = 0.492, *p* = 0.484, *partial* η^2 = 0.003). Therefore, we thought that beliefs about its existence did not significantly influence appropriate natural reaction time.



Figure 9. Questionnaire results of perceived naturalness in the expriment II. (**a**) the male model (android), (**b**) the female model (android).

			Before				
		0	200	400	600	800	1000
After	200	2.11	2.60	2.59	2.13	3.20	2.46
	400	4.71	5.88	6.80	5.99	5.24	4.66
	600	7.88	8.61	8.44	8.31	7.18	6.27
	800	8.01	8.72	8.80	8.01	7.44	6.83
	1000	8.54	8.39	7.90	7.12	6.82	6.66

Table 11. The perceived naturalness of the male model (android).

Table 12. The perceived naturalness of the female model (android).

		Before					
		0	200	400	600	800	1000
	200	2.31	2.54	2.28	2.34	2.73	2.42
After	400	4.99	5.95	6.14	5.83	4.99	4.47
	600	7.51	8.27	8.58	8.15	7.04	6.44
·	800	7.71	8.50	8.44	7.87	7.14	6.88
	1000	8.15	8.37	7.91	7.36	6.75	6.74

5. General Discussion

5.1. Factors That Possibly Change Impressions

Although belief was not shown to significantly influence reaction timing, we thought several possible factors existed that could change impressions. In this study, we asked participants to imagine various situations to evaluate the naturalness of the reaction behaviors. If we define such specific situations as surprised or absorbed, we may identify different data about before- and after-reaction times. We believe that our experimental results provide general knowledge about reaction behavior design in various contexts. Investigating appropriate reaction times for specific situations constitutes interesting future work.

One aspect of future work is to consider appropriate reaction timing and behavior for affective touch because past studies have reported the effects of affective context upon reaction time [16,17]. The reason why we did not deal with such affective touch in this study is because we wanted to find a basic reaction time as a baseline. Using different reaction timing to the baseline timing would be a useful way to judge appropriate reactions to affective touch. Factors such as affective touch interaction, touch parts (e.g., head or hand) [18], strength (perceived pain) [19], and frequency would influence appropriate reaction timing. Past studies have shown that these touch characteristics would change based on conveyed emotions via touches [20–22]. In this study, we used a gentle touch on the left shoulder. If the touch part is different (head or hand), the appropriate time will obviously also be different due to conveyed emotions. In addition, if a person violently touches the models, they should react more quickly and exhibit a different attitude.

From another perspective, investigation of the appropriate reaction times for children [17] and seniors [23] and comparing the results with the reaction times of adults would provide rich knowledge to understand human reaction times. People have different assumptions due to appearance and age; we assume that quicker/slower reaction times are appropriate for children/seniors respectively.

5.2. Implementation of Autonomous Turnaround Behaviors

Next, we discuss the possibility of making robots autonomous. Pepper and Nao have several touch sensors in their heads and hands. Such sensors do not cover their shoulders. Sota has no touch sensors. Therefore, attaching touch sensors to these robots would be necessary to enable them to react to touch behaviors autonomously. We propose that a capacitive sensor would be useful for this purpose. Several researchers have developed wearable touch sensors by using multiple capacitive sensors for robots [24–26]. Such capacitive sensors can roughly estimate the distance between the sensors and human body parts like a hand by observing the differences of time-series data. Therefore, attaching this type of sensor to a robot's body enables them to react toward touch behaviors.

5.3. Limitation

In this study, we used specific robots and human experimenters. The relationships between appearance/size and reaction time design will need more investigation. In addition, as described in the above sections, we fixed the accelerations of the reaction behaviors and did not focus on specific situations, such as being surprised. To apply our knowledge to different robots or children/seniors, we might need to investigate whether our collected data showed similar trends in these cases.

We only used the best combinations to compare the naturalness between humans and robots. As described above, many combinations are very close to the best combination, but we did not compare them by considering the context of implementing robots' behaviors and the number of peaks of these results. If we used a shorter time resolution (e.g., 50 ms), we might have observed a different number of peaks. In such a situation, a more detailed analysis would provide interesting knowledge about behavior designs. However, we think that the current time resolutions (i.e., 200 ms) would be reasonable to implement existing robots' behavior due to their hardware limitations.

Another limitation is the limited reaction behavior. The robots and human models only simply rotated toward touches. As described above, we employed this simple behavior because it can be applied to various kinds of robots. In future work we aim to investigate the effects of different kinds of reaction behaviors, such as more dynamic reactions, only voice reactions, etc.

However, even if several limitations exist, we believe that our study contributes sufficient information for designing appropriate reaction times for social robots and understanding how humans react to being touched. In fact, Nao and Pepper were often used for research platforms and robotics interfaces for actual services, and Sota is also used for commercial services in Japan. Therefore, even though we only investigate three different robots, the knowledge will influence a wide range of situations.

6. Conclusions

This study has investigated, via a web survey, the appropriate reaction time of robots when they are touched. Although robotics researchers investigated reaction time in conversational settings, they were less focused on touch interaction settings. For this purpose, we prepared different types of robot models and human models and investigated reaction timing effects by focusing on two types of reaction time lengths: before- and after-reaction times. The experiment results showed that appropriate reaction times are different for humans and robots, and found the best combinations for each model. The results provide useful knowledge to design robots' reaction timings and understand how people react toward touches.

Moreover, the results showed that using parameters made from observing human behaviors is sometimes not effective for a part of robots. Past human-robot interaction studies often conducted data collection to define parameters of robot's behaviors based on observing human-human interaction. Our study suggests that such parameters achieved better performance for robots, but calibrating parameters for individual robots' behavior design would achieve the best performance.

Author Contributions: Conceptualization, A.K., M.K., T.I., K.S., M.S.; methodology, A.K., M.K., T.I., K.S., M.S.; validation, A.K., M.K., T.I., M.S.; formal analysis, A.K., M.S.; investigation, A.K.; resources, A.K.; data curation, A.K.; writing—original draft preparation, A.K., M.S.; writing—review and editing, A.K., M.K., T.I., K.S., M.S.; visualization, A.K., M.S.; supervision, K.S., M.S.; project administration, M.S.; funding acquisition, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was supported in part by JST CREST Grant Number JPMJCR18A1, Japan.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of ATR (protocol code 20-501-4).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available from authors.

Acknowledgments: We thank Sayuri Yamauchi for her help during the execution of our experiments.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Kanda, T.; Kamasima, M.; Imai, M.; Ono, T.; Sakamoto, D.; Ishiguro, H.; Anzai, Y. A humanoid robot that pretends to listen to route guidance from a human. *Auton. Robot.* **2006**, *22*, 87–100. [CrossRef]
- Alhaddad, A.Y.; Cabibihan, J.-J.; Bonarini, A. Influence of Reaction Time in the Emotional Response of a Companion Robot to a Child's Aggressive Interaction. *Int. J. Soc. Robot.* 2020, *12*, 1279–1291. [CrossRef]
- Frederiksen, M.R.; Stoy, K. On the causality between affective impact and coordinated human-robot reactions. In Proceedings
 of the 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), Naples, Italy,
 31 August–4 September 2020; pp. 488–494.
- 4. Shiomi, M.; Minato, T.; Ishiguro, H. Subtle reaction and response time effects in human-robot touch interaction. In *International Conference on Social Robotics*; Springer: Cham, Switzerland, 2017; pp. 242–251.
- Shiomi, M.; Shatani, K.; Minato, T.; Ishiguro, H. Does a robot's subtle pause in reaction time to people's touch contribute to positive influences? In Proceedings of the 2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Nanjing, China, 27 August–1 September 2018; pp. 364–369.
- 6. Yamamoto, M.; Watanabe, T. Time delay effects of utterance to communicative actions on greeting interaction by using a voicedriven embodied interaction system. In Proceedings of the 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation, Kobe, Japan, 16–20 July 2003; pp. 217–222.
- 7. Shiwa, T.; Kanda, T.; Imai, M.; Ishiguro, H.; Hagita, N. How quickly should a communication robot respond? Delaying strategies and habituation effects. *Int. J. Soc. Robot.* **2009**, *1*, 141–155. [CrossRef]

- Kanda, T.; Shiomi, M.; Miyashita, Z.; Ishiguro, H.; Hagita, N. A communication robot in a shopping mall. *IEEE Trans. Robot.* 2010, 26, 897–913. [CrossRef]
- Arnold, T.; Scheutz, M. Observing robot touch in context: How does touch and attitude affect perceptions of a robot's social qualities? In Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, Chicago, IL, USA, 5–8 March 2018; pp. 352–360.
- Yamaoka, F.; Kanda, T.; Ishiguro, H.; Hagita, N. Interacting with a human or a humanoid robot? In Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, CA, USA, 29 October–2 November 2007; pp. 2685–2691.
- 11. Galton, F. Exhibition of instruments (1) for testing perception of differences of tint, and (2) for determining reaction-time. *J. Anthropol. Inst. Great Br. Irel.* **1890**, *19*, 27–29. [CrossRef]
- 12. Welford, W.T.; Brebner, J.M.; Kirby, N. Reaction Times; Stanford University: Stanford, CA, USA, 1980.
- 13. Robinson, E.S. Work of the integrated organism. In *A Handbook of General Experimental Psychology;* American Psychological Association: Washington, DC, USA, 2007; pp. 571–650.
- 14. Lele, P.P.; Sinclair, D.C.; Weddell, G. The reaction time to touch. J. Physiol. 1954, 123, 187–203. [CrossRef] [PubMed]
- Downs, J.S.; Holbrook, M.B.; Sheng, S.; Cranor, L.F. Are your participants gaming the system? Screening mechanical turk workers. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; pp. 2399–2402.
- 16. Gélat, T.; Chapus, C.F. Reaction time in gait initiation depends on the time available for affective processing. *Neurosci. Lett.* **2015**, 609, 69–73. [CrossRef] [PubMed]
- 17. Bucsuházy, K.; Semela, M. Case study: Reaction time of children according to age. Procedia Eng. 2017, 187, 408–413. [CrossRef]
- Schicke, T.; Röder, B. Spatial remapping of touch: Confusion of perceived stimulus order across hand and foot. *Proc. Natl. Acad. Sci. USA* 2006, 103, 11808–11813. [CrossRef] [PubMed]
- Arendt-Nielsen, L.; Bjerring, P. Reaction times to painless and painful CO₂ and argon laser stimulation. *Eur. J. Appl. Physiol.* Occup. Physiol. 1988, 58, 266–273. [CrossRef] [PubMed]
- 20. Alenljung, B.; Andreasson, R.; Lowe, R.; Billing, E.; Lindblom, J. Conveying emotions by touch to the nao robot: A user experience perspective. *Multimod. Technol. Interact.* **2018**, *2*, 82. [CrossRef]
- Zheng, X.; Shiomi, M.; Minato, T.; Ishiguro, H. What kinds of robot's touch will match expressed emotions? *IEEE Robot. Autom.* Lett. 2019, 5, 127–134. [CrossRef]
- 22. Teyssier, M.; Bailly, G.; Pelachaud, C.; Lecolinet, E. Conveying emotions through device-initiated touch. *IEEE Trans. Affect. Comput.* **2020**, 1. [CrossRef]
- Lewis, R.D.; Brown, J.M.M. Influence of muscle activation dynamics on reaction time in the elderly. *Eur. J. Appl. Physiol. Occup. Physiol.* 1994, 69, 344–349. [CrossRef] [PubMed]
- Kang, M.; Kim, J.; Jang, B.; Chae, Y.; Kim, J.-H.; Ahn, J.-H. Graphene-Based three-dimensional capacitive touch Sensor for wearable electronics. ACS Nano 2017, 11, 7950–7957. [CrossRef] [PubMed]
- 25. Tairych, A.; Anderson, I.A. Capacitive stretch sensing for robotic skins. Soft Robot. 2019, 6, 389–398. [CrossRef] [PubMed]
- Shiomi, M.; Sumioka, H.; Sakai, K.; Funayama, T.; Minato, T. SOTO: An android platform with a masculine appearance for social touch interaction. In Proceedings of the Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, Cambridge, UK, 23–29 March 2020; pp. 447–449.