

Warming up cool cooperators

Received: 5 October 2020

Accepted: 27 July 2023

Published online: 14 September 2023

 Check for updates

Eamonn Ferguson^{1,2}✉, **Claire Lawrence**³, **Sarah Bowen**⁴,
Carley N. Gemelli⁵, **Amy Rozsa**⁶, **Konrad Niekrasz**⁶, **Anne van Dongen**⁷,
Lisa A. Williams⁸, **Amanda Thijsen**⁵, **Nicola Guerin**⁵, **Barbara Masser**^{2,5,9}
& **Tanya E. Davison**^{5,10}

Explaining why someone repeats high-cost cooperation towards non-reciprocating strangers is difficult. Warm glow offers an explanation. We argue that warm glow, as a mechanism to sustain long-term cooperation, cools off over time but can be warmed up with a simple intervention message. We tested our predictions in the context of repeat voluntary blood donation (high-cost helping of a non-reciprocating stranger) across 6 studies: a field-based experiment ($n = 5,821$) comparing warm-glow and impure-altruism messages; an implementation study comparing a 3-yr pre-implementation period among all first-time donors in Australia ($N = 270,353$) with a 2-yr post-implementation period ($N = 170,317$); and 4 studies ($n = 716, 1,124, 932, 1,592$) exploring mechanisms. We show that there are relatively warm and cool cooperators, not cooling cooperators. Cooperation among cool cooperators is enhanced by a warm-glow-plus-identity message. Furthermore, the behavioural facilitation of future cooperation, by booking an appointment, is associated with being a warm cooperator. Societal implications are discussed.

Human cooperation presents the individual with the dilemma of deciding between what is good for them versus what is good for others¹. This dilemma has been defined in the literature in a variety of ways ranging from paying a personal cost to benefit others¹ to engaging in behaviours with mutual benefits^{2–4} that do not necessarily require coordinated action but may include reciprocal interaction^{2,3,5}. Indeed, there is a clear clustering of cooperative behaviours, ranging from altruism, in terms of the simple division of resources without direct reciprocity (for example, dictator game), to more complex reciprocal interactions (for example, public good and trust games), that define the human cooperative phenotype⁵. Human society is characterized by a remarkably high degree of cooperation¹ and for such high levels of cooperation to survive, definable benefits for the helper (and recipient) need to

outweigh costs^{1,4,6–8}. Indeed, several benefits (for example, reputation building) have been identified^{1,7–21}. Here we examined one aspect of the cooperative phenotype—altruism—whereby the donor bears a cost to benefit a non-reciprocating stranger in the absence of any other mechanisms to support altruism (for example, costly punishment)^{1,8,14,15,22}. Under such circumstances, we argue that warm glow sustains high-cost altruistic cooperation directly through reinforcement and indirectly by facilitating a behavioural commitment to cooperation. We tested these hypotheses using voluntary non-remunerated blood donation (VNRBD) as a real-world model of cooperation.

VNRBD follows the structure of the anonymous dictator and public goods games used to demonstrate warm glow in the lab^{20,21,23–26} and presents cooperation that incorporates both altruism^{27–29} and

¹School of Psychology, University of Nottingham, Nottingham, UK. ²National Institute for Health and Care Research Blood and Transplant Research Unit in Donor Health and Behaviour, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. ³Lawrence PsychAdvisory, Nottingham, UK. ⁴School of Economics, University of Nottingham, Nottingham, UK. ⁵Clinical Services and Research, Australian Red Cross Lifeblood, Melbourne, Victoria, Australia. ⁶Corporate Strategy and Transformation, Australian Red Cross Lifeblood, Melbourne, Victoria, Australia. ⁷Department of Psychology, Health, and Technology, University of Twente, Enschede, the Netherlands. ⁸School of Psychology, University of New South Wales, Sydney, New South Wales, Australia. ⁹School of Psychology, The University of Queensland, Brisbane, Queensland, Australia. ¹⁰Monash Art, Design, and Architecture, Monash University, Melbourne, Victoria, Australia. ✉e-mail: eamonn.ferguson@nottingham.ac.uk

contributing to the public good^{30,31}. Specifically, blood donors in a VNRBD system pay a cost to donate (for example, time, risk of fainting) to benefit a non-reciprocating stranger they will never meet (the recipient)²⁹. Donation in a VNRBD system also contributes to the public good³⁰, as donated blood is both non-excludable (freely available to anyone who needs it) and non-rivalry (there is sufficient blood, such that transfusing one person will not mean that another person cannot be transfused)³¹. This sufficiency of supply within a VNRBD system is maintained by transfusion services modelling and managing future supply and demand characteristics^{32,33}.

Andreoni^{11,20,21} proposed that altruism can be motivated by one of the following: (1) warm glow, (2) pure altruism or (3) impure altruism. Warm glow refers to helping that is motivated solely by the positive feelings experienced as a result of helping and not by achieving a particular charitable outcome. In contrast, pure altruism refers to helping, not to feel good about helping, but solely to achieve a successful charitable outcome. Finally, impure altruism is a combination of warm glow and pure altruism: the giver experiences warm glow and wants to donate sufficient resources to ensure a successful outcome. As such, warm glow can be assessed either as experienced positive emotions arising from giving or behaviourally, in terms of individual contributions to a cause when these contributions result in no material benefit to that cause²³ (Supplementary File 1 for details on assessing warm glow).

The general hypotheses of warm-glow theory are supported by a large body of studies^{23–26,34–48} which is summarized in Fig. 1a. Specifically, evidence shows that people experience warm glow following helping^{23,49–51} (ω), with the degree of warm glow being proportional to the amount donated to charity^{23,26}, with greater warm glow experienced when the donated resource was earned⁴⁵. Furthermore, both the degree of experienced warm glow^{37,49} and exogenous manipulations of warm glow (χ)^{52,53} have been shown to predict future helping (ρ). A mechanism to account for why experienced warm glow predicts future helping is likely to be reinforcement arising from the expectation of experiencing warm glow again (λ)²³. This reinforcing capacity of experienced warm glow is supported by evidence that warm glow is associated with: (1) activation of neural reward centres³⁶, (2) increased effort to help charity⁴³, (3) repeated acts of generosity⁴² and (4) the expectation of future reward²³. Further, altruism is itself rewarding⁵¹. However, while exogenous manipulations of warm glow enhance cooperation, the mechanisms by which they achieve this are unknown^{52,53}.

The current paper develops this understanding (Fig. 1b). We start from the premise that donors experience varying degrees of experienced warm glow (ω)^{26–28,50}. Some experience stronger warm glow (ω_{ii} ; ‘warm cooperators’) than others (ω_i ; ‘cool cooperators’), specifically $\omega_{ii} > \omega_i$. The function $\omega_{ii} > \omega_i$ represents a relative relationship whereby cool cooperators still experience some warm glow but to a lesser extent than warm cooperators. Warm cooperators (ω_{ii}) are predicted to have a greater probability of repeat cooperation via direct reinforcement effects of warm glow (λ). We also propose a second indirect behavioural mechanism, booking another appointment to donate during the donor’s current donation (β), through which stronger experienced warm glow increases the probability of repeat cooperation (ρ). Psychologically, booking will enhance a commitment to cooperate (\emptyset)⁵⁴ and reinforce the donor’s self-identity as a donor (π)^{54,55}, which in itself is a crucial predictor of prosociality^{56,57}. A behavioural commitment to cooperate will also facilitate future cooperation through planning and scheduling (putting the date in the diary), which reduces the likelihood that a further donation will be derailed by competing commitments or forgetting⁵⁸. Thus, booking an appointment will be a strong predictor of subsequent cooperation. Consistent with evidence that positive affect is associated with taking steps to facilitate future prosociality⁵⁹, we predict that making a booking (commitment to cooperation) will be more likely for warm cooperators (ω_{ii}) than for cool cooperators (ω_i).

For cool cooperators (ω_i), the reinforcing effects of warm glow (λ) and the likelihood of booking another appointment (β) are both

less likely. Therefore, we propose that an intervention based on warm glow (χ) will be more effective for these cool cooperators by eliciting feelings of warm glow. Further, as donor identity (π) strongly predicts repeat donation, a warm-glow message that also emphasizes donor identity will be more effective at encouraging cool cooperators to donate again. Therefore, a message that highlights both warm glow and donor identity should be especially effective for cool cooperators who can be identified behaviourally as those who do not book their next appointment to donate.

However, the reinforcing effect of warm glow (λ) may diminish over time, especially if the interval between cooperative acts is long, resulting in affective decay^{60,61} and unintentional free-riding^{62,63}. Thus, we may have cooling cooperators. Research based on affective chronometry (the study of mood change over time) suggests that affect decays^{60,61}. However, there is evidence that for positive affect augmenters, such as our warm cooperators, positive affect decays at a slower rate than for negative affect augmenters, such as cool cooperators⁶⁴. Therefore, for cool cooperators (ω_i), this decay may be more rapid than for warm cooperators (ω_{ii}). Understanding the temporal dynamics of warm glow should help explain why warm-glow messages are more effective for cool cooperators, for example, if their warm glow decays more quickly⁵¹.

So far, two field-based experiments have exogenously manipulated warm-glow messages to test their impact on real-world cooperation^{52,53}. Both experiments involved donating financial resources. The first study compared the effects of warm-glow messages (for example, ‘Please donate your refund – think of the good feeling in helping others.’) against norm-based messages (for example, ‘Many of our customers from this store regularly donate their refunds...’) displayed at supermarkets on the likelihood of donating bottle return refunds to charity⁵². Significantly higher donation rates were observed in supermarkets displaying warm-glow messages than supermarkets displaying norm-based messages⁵². However, as messages were randomly allocated by supermarket rather than by consumer, it cannot be ruled out that these results were due to differences in consumer populations between supermarkets⁵².

In the second field-based experiment, participants were randomly allocated to one of three message conditions: warm glow (‘... Warm your heart’), pure altruism (‘... Make Alaska better for everyone’) and a no-treatment control (no message). The likelihood of participants donating to a chosen charity following a financial windfall and the amount donated were higher with exposure to the warm-glow than to the other messages⁵³.

We further developed this work on the effectiveness of exogenous warm-glow messages. While previous research compared warm glow with pure altruism³³, studies have consistently shown that blood donation is unlikely to be an example of pure altruism^{27–29,50,65–67}. However, impure-altruism may motivate blood donors^{27,65,66}. Therefore, we contrast warm-glow with impure-altruism messages; our warm-glow message focuses on the donor’s warm glow with no reference to benefits to recipients, whereas our impure-altruism message focuses both on the donor’s warm glow and benefits to recipients (Supplementary File 2 and Table 1). Thus, both warm-glow and impure-altruism messages feature warm glow as an active ingredient of the message. However, the impure-altruism message also includes pure altruism as an active ingredient (benefits to the recipient). Thus, we tested whether warm-glow or impure-altruism messages would enhance return donations more, especially for those who had not committed to a repeat donation (cool cooperators).

We further explored whether priming donor identity would enhance the effectiveness of warm-glow and impure-altruism messages. We predicted that (1) messages including a donor-identity prime would be more effective than those that did not prime identity and, (2) messages with a higher number of active elements would be more effective than those with fewer active elements. Therefore, we predicted that the rank

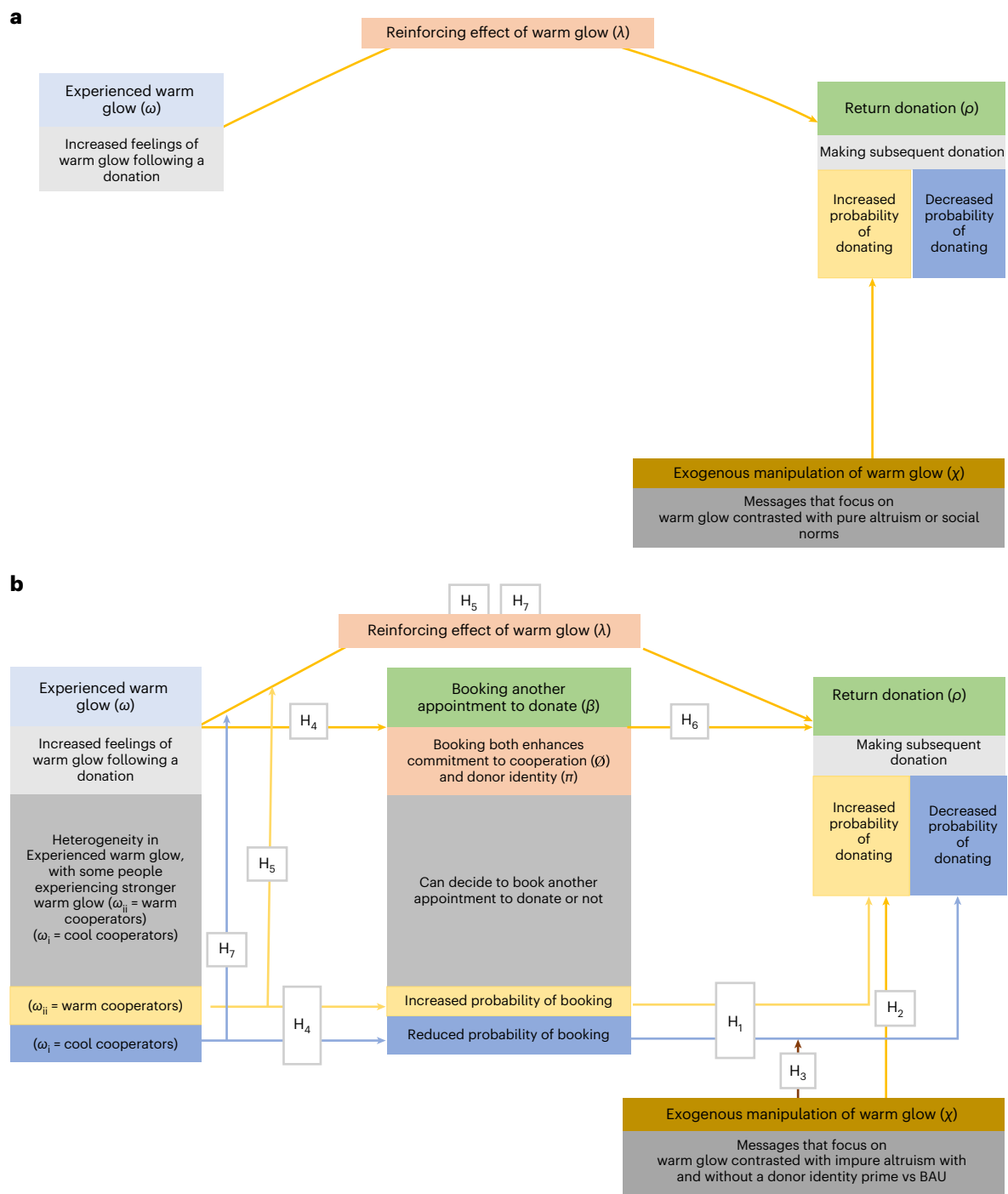


Fig. 1 | Theoretical models linking warm glow to cooperation. a, Current knowledge. Summary of the general findings supporting warm-glow theory. Specifically, people experience variation in warm glow from giving (ω) (see main text for factors that influence the extent of experienced warm glow) and this predicts future donations (ρ) based on the reinforcing effects of warm glow (λ) (see main text). Exogenous manipulations of warm glow (χ) predict repeat donations. **b**, Model to be tested. We extend **a** by suggesting an additional behavioural mechanism, booking another appointment to donate (β), which enhances commitment to cooperate (\emptyset) and identity as a donor (π). Overall increased feelings of experienced warm glow increase the probability of booking, with this being stronger for warm cooperators (those with higher

levels of experienced warm glow) than cool cooperators (those with lower levels of experienced warm glow). Booking (β) predicts return donation (ρ), which is again stronger for warm cooperators. An exogenous manipulation of warm glow/impure altruism will increase the probability of making another donation if a donor identity prime is present. As cool cooperators are less likely to book and have their donor identity strengthened, the warm-glow/impure-altruism-plus-identity messages should be maximally effective for cool cooperators. Green, behavioural mechanism; peach, psychological mechanism; brown, exogenous manipulation; light blue, overall experienced warm-glow; yellow, effects linked to being a warm-cooperator; dark blue, effects linked to being a cool cooperators; greys, indications of associated behaviours or action.

order of message effectiveness would be: ‘impure-altruism-plus-identity’, ‘warm-glow-plus-identity’, ‘impure-altruism-only’ and finally ‘warm-glow-only’ (Supplementary File 2).

Overall, the aim of this paper is to extend the existing literature on warm glow and test key predictions outlined in Fig. 1b. Specifically: (1) the effectiveness of warm-glow/impure-altruism messages will be

Table 1 | Detailed hypotheses by study

Hypothesis	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6
Behavioural effects						
H ₁ : Commitment-to-cooperate effect: ‘Those who book another appointment will be more likely to return than those who do not book.’~	✓	✓		✓	✓	
H ₂ : Effect of the warm-glow message: ‘Compared to a BAU control, the impure-altruism-plus-identity and warm-glow-plus-identity messages will have the strongest effects on donor return.’~	✓	✓				
H ₃ : Message by donor commitment-to-cooperate effect: ‘A message that contains an identity component (warm-glow and/or impure-altruism) will have a significant effect on donor return only in those who did not book another appointment.’~	✓	✓				
Behavioural mechanisms						
H ₄ : Experienced warm-glow commitment-to-cooperation effect: “Experienced warm glow will be higher in those who have booked another appointment compared to those who have not.”			✓	✓	✓	
H ₅ : Direct effect of experienced warm glow: “Increased experienced warm glow predicts returning to make a repeat donation.”				✓	✓	
H ₆ : Indirect effect of experienced warm glow: “Experienced warm glow will be indirectly linked to return donor behaviour through making a commitment to cooperate (booking another appointment).”~				✓		
H ₇ : Temporal dynamics of warm glow: “Warm glow will decay over time, especially for cool cooperators (those who have not booked another appointment).”					✓	
Treatment messages: active ingredient						
H ₈ : Affordance of warm glow: “Both warm-glow and impure-altruism messages should afford more experienced warm glow, compared to BAU message.”						✓
H ₉ : Identity enhances warm glow: “Experienced warm glow should be greater in the presence of an identity prime.”~						✓
H ₁₀ : Warm-glow messages focus on the donor: “The warm-glow compared to the impure-altruism message should focus more on the donor and less on the recipient or both donor and recipient.”~						✓
H ₁₁ : Impure-altruism messages focus on maintaining blood supply: “The impure-altruism compared to the warm-glow message should focus more on maintaining blood supply (the public good).”~						✓
H ₁₂ : “Focus on the donor mediates the association between exposure to a warm-glow-plus-identity message and experienced warm glow.”~						✓

Studies and design. Study 1. Longitudinal field-based experiment: 5,821 first-time blood donors were randomly allocated to one of five warm-glow and impure-altruism conditions (with or without an identity prime) and a BAU control. Study 2. Pre-, post-implementation study comparing donor return rates in 270,353 first-time donors before the implementation of the warm-glow-plus-identity message with 170,317 donors post-implementation. Study 3. Survey of 716 first-time donors. Study 4. Survey of 1,124 donors, crossing donor status (first-time, novice) with donor type (plasma, blood). Study 5. Longitudinal study with 932 first-time donors by measuring warm glow at two timepoints (baseline). Study 6. Online between-subjects experiment: 1,592 members of the public were randomized to one of six conditions. ~Pre-registered prediction. All other behavioural effects and behavioural effects hypotheses were developed in the initial paper and tested formally in the subsequent revisions. Treatment-message hypotheses were developed in response to reviewers’ comments. Data from studies 3 to 5 were secondary data analyses to address components of the warm-glow model in the initial submission. Study 6 was designed to test hypotheses developed after the initial submission of the paper.

enhanced when presented in conjunction with a donor-identity prime, (2) warm-glow/impure-altruism messages with a donor-identity prime will be most effective for those who have not made a commitment to cooperation, (3) experienced warm glow will have a direct effect on future cooperation, (4) experienced warm glow will have an indirect effect on future cooperation through a commitment to cooperation and (5) experienced warm glow decays over time, with that decay being greater for cool vs warm cooperators. We addressed these aims by testing the hypotheses set out in Table 1 (hypotheses marked with a ~ were pre-registered) across six studies.

Studies 1 and 2 test behavioural hypotheses 1–3 and address the main question: “Compared to a business-as-usual (BAU) control, does a warm-glow/impure-altruism message plus an identity prime encourage repeat donation, specifically among those who have not booked an appointment for a future donation?” Study 1 is a field-based experiment that examined donor return behaviour as a function of booking status (booked, not booked) and four active message arms created by crossing message type (warm-glow, impure-altruism) with the donor identity prime (“... that’s when you became a blood donor ...”: yes, no) (see Fig. 2 and Supplementary File 2 for details of message

structure) compared to a BAU control. First-time blood donors (5,821) were randomly allocated to one of the five treatment conditions. Study 2 tested the most effective message from the field-based experiment (warm-glow-plus-identity) in an implementation trial. This trial compared return behaviour among all first-time Australian blood donors in three 1-yr time windows before the warm-glow-plus-identity message was implemented (total $n = 270,353$; 54.6% had not booked) and two 1-yr time windows after implementation (total $n = 170,317$; 65.6% had not booked).

In studies 3 to 5, we explored the behavioural mechanisms that explain the effectiveness of the ‘warm-glow-plus-identity’ message (hypotheses 4–7). These address the main question “Does experienced warm glow predict future cooperation directly and indirectly via a commitment-to-cooperation?”. In Study 3 (survey of 716 first-time donors), we evaluated whether booking status (booked, not booked) was associated with warm glow or pure altruism. Study 4 (survey of 1,124 donors) compared four groups of donors by crossing donor status (first-time, novice) with donor type (plasma, whole blood) to explore the association between booking status and experienced warm glow and whether experienced warm glow and booking status predicted

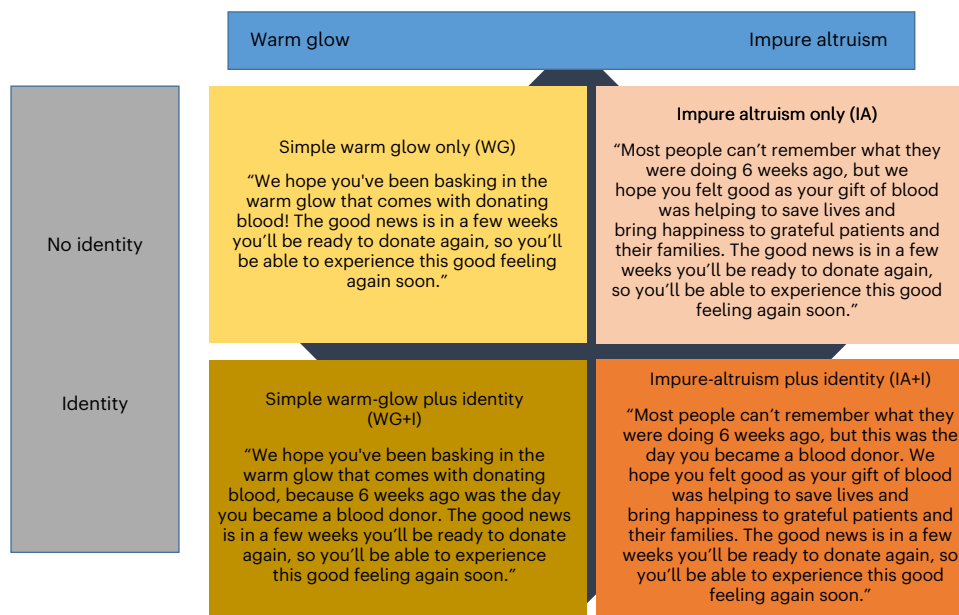


Fig. 2 | Treatment messages used in Studies 1 and 6. The messages formed when warm glow (warm glow, impure altruism) is crossed with identity (present, absent).

return behaviour. Study 5 investigated the temporal dynamics of experienced warm glow (does warm-glow decay and, if so, is the rate of decay higher among those who did not book another appointment) among $n = 932$ first-time donors by measuring experienced warm glow at two timepoints (baseline, 12-week follow-up). Study 5 also explored the association between experienced warm glow, booking status and return behaviour. Study 4 also explored the indirect effect of warm glow on return behaviour via a commitment to cooperate. Study 6, an online experiment ($n = 1,592$), explored the active ingredients of the ‘warm-glow-plus-identity’ message that make it the most effective (hypotheses 8–12).

Results

We first present the results of tests of the behavioural effects (hypotheses 1–3; Studies 1–2), followed by tests of the behavioural mechanisms (hypotheses 4–7; Studies 3–5) and the active elements of the treatment messages (hypotheses 8–12; Study 6).

Study 1 Field-based experiment

Table 2 shows the number of donors by treatment (warm-glow-only, warm-glow-plus-identity, impure-altruism-only, impure-altruism-plus-identity and no-message BAU control), their implied commitment to cooperate (booked or did not book a future appointment) and return rates. As expected, based on the randomization to treatments and the fact that donors were exposed to the treatment messages after their booking decisions, the likelihood of booking did not statistically significantly vary across treatments ($\chi^2_{(4)} = 4.198, P = 0.380, \varphi = .027$). Thus, any variability in return rates is due to the treatment messages and cannot be attributed to initial variation in booking status.

Effect of behavioural commitment (H_1). The final column of Table 2 shows that for each treatment group, making a booking resulted in higher return rates.

Behavioural effects of warm glow (H_2 and H_3). First, we compared whether exposure to an active message, regardless of the message content (combined across the four treatments) influenced return rates. The results comparing return rates for the no-message BAU control to those who were exposed to an active message, regardless of content, are presented in Table 2 (last row). The results show that return rates

were not significantly different following exposure to any active message, both for donors who had booked and those who had not. Thus, we found no evidence that simple exposure to any message, regardless of content, influenced return rates.

Next, we tested whether exposure to a specific treatment message (for example, warm-glow-plus-identity or impure-altruism-plus-identity) increased return rates more than exposure to the BAU control among those who booked or those who did not book. The ‘warm-glow-plus-identity’ message was the only significant predictor of return rates among those who had not booked another appointment, and there were no significant effects among those who had booked (Supplementary File 2 and Table 2).

To consolidate the above univariate analyses, we conducted an intention-to-treat analysis with treatment, booking status and the treatment by booking-status interaction predicting returning to donate. These analyses controlled for other salient predictors of return to donate (age, gender and blood type). A logistic regression model (Table 3) indicated that the likelihood of returning increased with age and was higher among O– blood donors than O+ donors. Donors exposed to the ‘warm-glow-plus-identity’ message were 1.279 times more likely to return than the BAU control. No other treatment messages impacted return significantly relative to the BAU control. A significant main effect of booking status showed that, overall, donors who booked another appointment were 3.242 times more likely to return than those who had not booked.

The effect of booking status interacted significantly with the ‘warm-glow-plus-identity’ message ($B = -0.506, s.e. = 0.180; P = 0.005$; odds ratio (OR) = 0.603; 95% confidence interval (CI) = 0.424, 0.859). Margins analysis was used to explore this interaction. Compared with the BAU control, those in the ‘warm-glow-plus-identity’ treatment who had not booked were significantly more likely to return ($dx/dy = 0.058, s.e. = 0.025; P = 0.023; 95\% CI = 0.008, 0.108$; Cohen’s $d = 0.133$), but for those who had booked, there was no statistically significant variation ($dx/dy = -0.060, s.e. = 0.033; P = 0.070; 95\% CI = -0.124, 0.005$; Cohen’s $d = -0.143$) (Supplementary File 2, and Tables 3 and 4).

Study 2 Implementation analysis

The ‘warm-glow-plus-identity’ message was implemented and sent to all first-time donors in Australia, 6 weeks after their first donation (Supplementary File 3). We compared return rates among those who

Table 2 | Descriptive breakdown of the number (percentage) of first-time donors by treatment and whether they booked or did not book an appointment to donate again and then attended to make a subsequent donation for Study 1

Treatment (message received)	All first-time donors		First-time donors who booked		First-time donors who did not book		Comparison of return rate of those who booked vs those who did not ^a
	Number of first-time donors assigned to each treatment	Number (%) of first-time donors	Number (%) who donated	Number (%) of first-time donors	Number (%) who donated		
BAU	1,161	416 (35.8)	277 (66.6)	745 (64.2)	278 (37.3)	$\chi^2_{(1)}=91.660, P<0.001, \phi=0.281$	
WG	1,156	428 (37.0)	292 (68.2)	728 (63.0)	269 (37.0)	$\chi^2_{(1)}=105.539, P<0.001, \phi=0.302$	
WG+I	1,158	439 (37.9)	266 (60.6)	719 (60.1)	307 (42.7)	$\chi^2_{(1)}=34.915, P<0.001, \phi=0.174$	
IA	1,166	464 (39.8)	286 (61.6)	702 (60.2)	279 (39.7)	$\chi^2_{(1)}=53.616, p P<0.001, \phi=0.214$	
IA+I	1,160	441 (38.0)	269 (61.0)	719 (62.0)	285 (39.6)	$\chi^2_{(1)}=49.983, p P<0.001, \phi=0.208$	
Comparison BAU vs Combined treatments on return rate			$\chi^2_{(1)}=2.073 P=0.150, \phi=0.031$		$\chi^2_{(1)}=1.496 P=0.226, \phi=0.020$		

All tests are two-tailed and no adjustment for multiple comparisons was applied. WG, warm-glow-only; WG+I, warm-glow-plus-identity; IA, impure-altruism-only; IA+I, impure-altruism-plus-identity. An overall comparison of the BAU treatment to the combined message treatments shows that there was no overall effect of simply being exposed to a message ($\chi^2_{(1)}=0.211, P=0.646, \phi=0.006$); results in the last row also show no effect of simply being exposed to a message for those who booked and those who did not book. Results in the last column show that booking enhances return rates across all treatments. ^aExact P values for the comparison of return rates of those who booked vs those who did not for (1) BAU, $P=1.0277 \times 10^{-21}$; (2) WG, $P=9.3047 \times 10^{-25}$; (3) WG+I, $P=3.4445 \times 10^{-9}$; (4) IA, $P=2.4376 \times 10^{-19}$, and (5) IA+I, $P=1.5511 \times 10^{-12}$.

Table 3 | Logistic regression predicting return behaviour at 3 months in first-time donors

	B (s.e.)	P	OR	95% CI		Cohen's d
				Lower	Upper	
Gender	-0.03 (0.06)	0.598	0.971	0.896	1.084	-0.016
Age	0.023 (0.002)	<0.001	1.023	1.019	1.028	0.012
Blood group		0.022				
A-	0.098 (0.115)	0.394	1.103	0.880	1.382	0.054
A+	-0.009(0.060)	0.838	0.991	0.880	1.116	-0.005
O-	0.266 (0.093)	0.004	1.304	1.087	1.565	0.146
Message treatment		0.120				
WG+I	0.246 (0.108)	0.023	1.279	1.035	1.580	0.136
WG	-0.010 (0.109)	0.928	0.990	0.800	1.226	-0.005
IA+I	0.105 (0.109)	0.334	1.111	0.898	1.374	0.058
IA	0.111 (0.109)	0.308	1.118	0.903	1.384	0.061
Booking status	1.176 (0.130)	<0.001	3.242	2.513	4.182	0.648
Interactions		0.004				
Booking status×WG+I	-0.506 (0.180)	0.005	0.603	0.424	0.859	-0.279
Booking status×WG	0.097 (0.184)	0.596	1.102	0.768	1.581	0.053
Booking status×IA+I	-0.335 (0.180)	0.064	0.715	0.502	1.019	-0.185
Booking status×IA	-0.297 (0.180)	0.098	0.743	0.523	1.057	-0.164
Intercept	-1.263 (0.119)	<0.001	0.283			
R ²	0.10					

Gender (0 = men; 1 = women); blood group O+, reference category; booking status (0 = did not book another appointment to donate; 1 = did book another appointment to donate). The reference group for the message treatments is the BAU no message control. All tests are two-tailed and no adjustment for multiple comparisons was applied. B coefficients are unstandardized coefficients. Exact P values for (1) age, $P=9.5071 \times 10^{-23}$; (2) booking status, $P=1.4149 \times 10^{-19}$; and (3) constant, $P=9.6695 \times 10^{-27}$.

had booked another appointment with those who had not booked, across three time windows preceding the implementation and two time windows following the implementation. The number of first-time donors in each time window, the number who booked or not and the

number who returned to donate can be found in Supplementary File 3, Fig. 1 and Table 5.

We aggregated these data over the three pre-implementation time windows ($n = 270,353$, of which 122,681 booked and 147,672 did

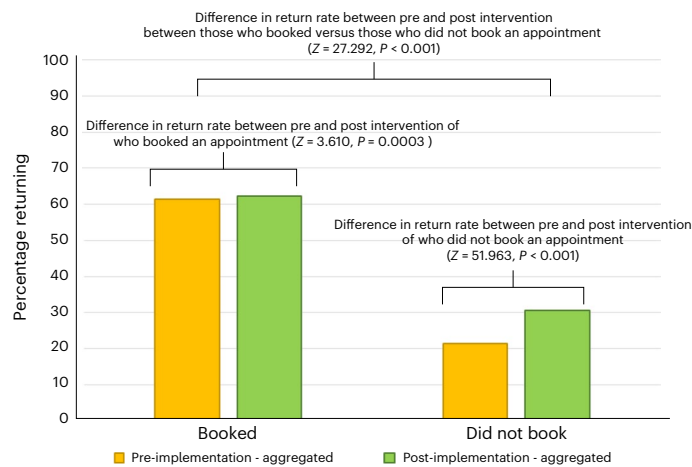


Fig. 3 | Aggregate return behaviour pre- and post-implementation of the ‘warm-glow-plus-identity’ message in new donors as a function of booking status (Study 2). For the pre-implementation, the aggregated point estimate for the percentage return rate for those who booked was 61.47 (95% CI = 61.20, 61.74; $n = 75,414/122,681$) and for those who did not book, 21.49 (95% CI = 21.28, 21.70; $n = 31,738/147,672$). For the post-implementation, the aggregated point estimate for the percentage return rate for those who booked was 62.35 (95% CI = 61.96, 62.74; $n = 36,472/58,500$) and for those who did not book, 30.55 (95% CI = 30.28, 30.82; $n = 34,156/111,817$). For those who had booked, there was a 0.88% (95% CI 0.402, 1.358) significant ($Z = 3.610$, $P = 0.0003$) increase in returning donors. For those who had not booked there was a 9.06% (95% CI = 8.718, 9.402) significant ($Z = 51.963$, $P < 0.001$) increase in returning donors. There was a significant interaction with the percentage increase in return rates for those who had not booked, significantly greater than those who had booked ($Z = 27.292$, $P < 0.001$), with a percentage difference of 8.18% (95% CI = 7.593, 8.767). Analyses were conducted using procedures for Z tests for proportions detailed in refs. 98,99 and implemented in ZumaStat 4.0. All analyses were two-tailed and no adjustments were made for multiple comparisons. Error bars are 95% CIs.

not book) and two slightly overlapping post-implementation windows ($n = 170,317$, of which 58,500 booked and 111,817 did not book). Figure 3 plots the aggregate return rates pre- and post-implementation for those who booked and those who did not. For those who had booked, there was a significant increase (0.88%; 95% CI = 0.402, 1.358; $Z = 3.610$; $P < 0.001$) in returning donors, whereas, for those who had not booked, the percentage increase was 9.06% (95% CI = 8.718, 9.402), which was significant ($Z = 51.963$, $P < 0.001$). There was a significant interaction with the percentage increase in return rates for those who had not booked, significantly greater than those who had booked ($Z = 27.292$; $P < 0.001$; % increase = 8.18%; 95% CI = 7.593, 8.768). Thus, the ‘warm-glow-plus-identity’ message positively affected return behaviour for both those who had booked and those who had not, but the effect was significantly larger for those who had not booked. Due to the slight overlap in the two post-implementation time windows, we also compared the two post-implementation windows separately with the pre-implementation window immediately before the post-implementation period. The pattern of results is identical to the aggregated analyses (see Supplementary File 3 and Fig. 2). We also ran a sensitivity analysis to adjust these aggregate return rates for aggregate age and percentage of women (see Supplementary File 3). The hypothesized positive effect of the warm-glow-plus-identity message for those who had not booked remained.

Studies 3 to 5 Behavioural mechanisms

Warm glow higher in those who book an appointment (H_4). Secondary analysis of data from Studies 3 to 5 allowed us to test the hypothesis that those who book will report higher experienced warm glow (Table 4, panels A–C, respectively).

As Andreoni’s model^{11,20,21} specifically contrasts warm glow with pure altruism, Study 3 was used to explore the relative effects of warm glow and pure altruism with respect to booking status (see Supplementary File 4 and Table 6 for the derivation of the measure of ‘anticipated warm glow’). We regressed (logistic model) booking status (0, did not book, 1, booked) on ‘anticipated warm glow’ and ‘pure altruism’. ‘Anticipated warm glow’ was significantly associated with booking status ($B = 0.0847$, s.e. = 0.0423; $P = 0.045$; 95% CI = 0.0018, 0.1677), such that higher levels of ‘anticipated warm glow’ were associated with booking another appointment, whereas for pure altruism there was no significant variation ($B = -0.0648$, s.e. = 0.07503; $P = 0.338$; 95% CI = -0.2118, 0.0822). We explored whether this association remains when other affective states associated with blood donation—‘anticipated feelings of calmness’ and ‘anticipated negative emotions’ (see Supplementary File 4 and Table 6 for the derivation of these measures)—are included in the model, along with controlling for donors’ age and gender. The resulting logistic regression is reported in Table 4a. The positive effect of anticipated warm glow on booking status remained but was not significant ($P = 0.097$). Anticipated negative emotions were not significantly associated with booking, as was pure altruism or anticipated calmness (a Heckman selection model (Supplementary File 4 and Table 7) showed no evidence of selection bias).

Study 4 provided a replication of Study 3 with a different measure of warm glow while controlling for donor status (first-time or novice) and donor type (whole blood or plasma). A logistic regression model was specified to test the effects of reported warm glow on booking another appointment, controlling for age, gender, donor status (first-time donors or novice donors) and donation type (whole blood or plasma). Warm glow was significantly and positively associated with booking status, but neither donor status nor donation type was significantly associated with booking (Table 4, panel B; for additional analyses predicting warm glow and robustness checks for sample clustering, see Supplementary File 4 and Tables 8–10).

Study 5 also shows that booking status predicts reported warm glow, with those who book reporting higher warm glow.

Across these analyses, there is clear evidence that making a commitment to cooperate (booking another appointment to donate) is associated with higher levels of warm glow. However, the effect size varies. Thus, using the mean scores for warm glow for those who had and had not booked across studies 3, 4 and 5, we report a simple random effects meta-analysis using Comprehensive Meta-Analysis v.2. The results show that there was an overall significant positive effect across the three studies (standard difference in means = 0.209, s.e. = 0.052; 95% CI = 0.106, 0.312; $P < 0.001$; Cohen’s $d = 0.201$; $r = 0.100$). Thus, there is a consistent positive association between reported warm glow and booking status (see Supplementary File 4 for analysis details and Supplementary Fig. 3 for Forest plot).

Warm glow (H_5) and booking (H_6) predict future donation. Results from Studies 4 and 5 (Supplementary File 5, and Tables 11 and 12) showed that both experienced warm glow (Study 4: B (s.e.) = 0.041 (0.016); $P = 0.008$; Cohen’s $d = 0.024$; OR = 1.042; 95% CI = 1.011, 1.074; Study 5: B (s.e.) = 0.048 (0.013); $P < 0.001$; Cohen’s $d = 0.026$; OR = 1.049; 95% CI = 1.022, 1.076) and booking status (Study 4: B (s.e.) = 1.408 (0.147); $P < 0.001$; Cohen’s $d = 0.776$; OR = 4.088; 95% CI = 3.068, 5.448; Study 5: B (s.e.) = 0.803 (0.111); $P < 0.001$; Cohen’s $d = 0.443$; OR = 2.232; 95% CI = 1.795, 2.775) positively predicted return behaviour.

Experienced warm glow predicts donations via booking (H_6). We proposed a path model (detailed in Supplementary File 5 and Fig. 4) that specified that experienced warm glow positively affected return behaviour via booking. The indirect effect of experienced warm glow on return via booking another appointment was significant ($B = 0.022$, s.e. = 0.006; $P < 0.001$; 95% CI = 0.013, 0.031).

Table 4 | Regression models from Studies 3 to 5

Panel A						
	<i>B</i> (s.e.)	<i>P</i>	OR	95% CI for OR		Cohen's <i>d</i>
				Lower	Upper	
Age	0.024 (0.005)	<0.001	1.024	1.014	1.035	0.013
Gender	0.229 (0.169)	0.171	1.257	0.906	1.743	0.126
Anticipated negative affect	0.066 (0.072)	0.356	1.068	0.928	1.230	0.036
Anticipated calmness	0.014 (0.067)	0.839	1.014	0.889	1.155	0.077
Anticipated warm glow	0.079 (0.048)	0.097	1.082	0.986	1.188	0.044
Pure altruism	-0.091 (0.077)	0.236	0.913	0.785	1.061	-0.050
Constant	-2.590 (1.127)	0.022	0.075	0.008	0.685	-1.428
<i>R</i> ² (<i>n</i>)	0.049 (716)					
Panel B						
Warm glow	0.066 (0.015)	<0.001	1.068	1.037	1.100	0.036
Gender	0.108 (0.134)	0.420	1.114	0.857	1.449	-0.060
Age	0.001(0.006)	0.809	1.001	0.990	1.013	0.001
Donor status	0.087 (0.245)	0.723	1.090	0.675	1.761	0.048
Donation type	-0.216 (0.186)	0.245	0.806	0.560	1.160	-0.112
Donor status×Donor type	0.273 (0.291)	0.349	1.314	0.742	2.325	0.151
Constant	-0.206 (0.309)	0.505	0.813	0.443	1.491	-0.114
<i>R</i> ² (<i>n</i>)	0.020 (1124)					
Panel C						
Gender	0.679 (0.208)	0.001		0.270	1.088	0.375
Age	-0.005 (0.007)	0.435		-0.019	0.008	-0.003
Booking status	0.597 (0.284)	0.036		0.040	1.152	0.329
Wave	0.078 (0.332)	0.809		-0.554	0.710	0.043
Booking status×Wave	0.216 (0.398)	0.587		-0.564	0.996	0.119
Constant	14.316 (0.365)	<0.001		13.596	15.035	0.375
<i>n</i> (observations)	932 (1,864)					

Panel A (Study 3): association between 'anticipated warm glow' and pure altruism with booking status (0 = did not book another appointment to donate; 1 = did book another appointment to donate) and gender (0 = men; 1 = women). Panel B (Study 4): donor status, donor type and warm glow to predict booking status. Gender (0 = men; 1 = women); donor status (0 = first time; 1 = novice); donor type (0 = plasma; 1 = whole blood); booking status (0 = did not book another appointment to donate; 1 = did book another appointment to donate). Panel C (Study 5): predicting the stability of warm glow over 12 weeks. Booking status (0 or 1), wave (0 = wave 1; 1 = wave 4), gender (0 = men; 1 = women). All tests are two-tailed and no adjustment for multiple comparisons was applied. *B* coefficients are unstandardized coefficients.

Warm glow decays quicker in those not booking (H₁). We hypothesized that experienced warm glow cools over time and that this cooling would be quicker for those who did not book. We test this prediction using longitudinal data from Study 5, in which experienced warm glow was assessed 1 day (wave 1) and 12 weeks (wave 4) post-donation in those who had booked and those who had not. Table 4c shows the results of a generalized estimating equations model evaluating whether booking status predicted experienced warm glow across time. Results showed that those who booked reported higher average levels of experienced warm glow overall ($M = 15.32$, $s.e. = 0.155$) than those who did not ($M = 14.68$, $s.e. = 0.215$). There was no significant effect of time on experienced warm-glow levels or significant moderation effect of time by booking status. This indicates that the association of booking status on experienced warm glow was stable over time. Thus, rather than having cooling cooperators, we have cool cooperators (who do not book another appointment) and warm cooperators (who booked another appointment).

Study 6 The active ingredients of warm-glow-plus-identity

It is important to identify the active ingredients of the 'warm-glow-plus-identity' message that prompts its effectiveness. To reflect

theory, the warm-glow messages used across the studies focused solely on the donor, with no reference to recipient benefits, while the impure-altruism messages focused on both the donor and the recipient¹¹. Thus, both warm-glow and impure-altruism messages should afford experienced warm glow more than a BAU control, as both contain active warm-glow ingredients (H₈), with afforded warm glow greater when a person's identity as a donor is primed (H₉). Furthermore, warm-glow messages, compared with impure-altruism messages, should focus more on the donor than the recipient (H₁₀) and less on donating to maintain the public good (blood supply) (H₁₁). We predict that the warm-glow-plus-identity message is effective because the feelings of warm glow it generates are about the donor. Thus, exposure to a 'warm-glow-plus-identity message' should be indirectly linked to afforded warm glow through an enhanced focus on the donor relative to the recipient or both the donor and the recipient (H₁₂).

We tested these predictions using data from Study 6 in which participants were exposed to one of the four messages used in the field-based experiment. We assessed each message's ability to elicit: (1) 'afforded experienced warm glow' (possible range 2 to 14, where 14 is high warm glow), (2) a focus on the 'donor, the recipient or both' (-50 is focus on the blood donor, 0 is focus equally on both the donor

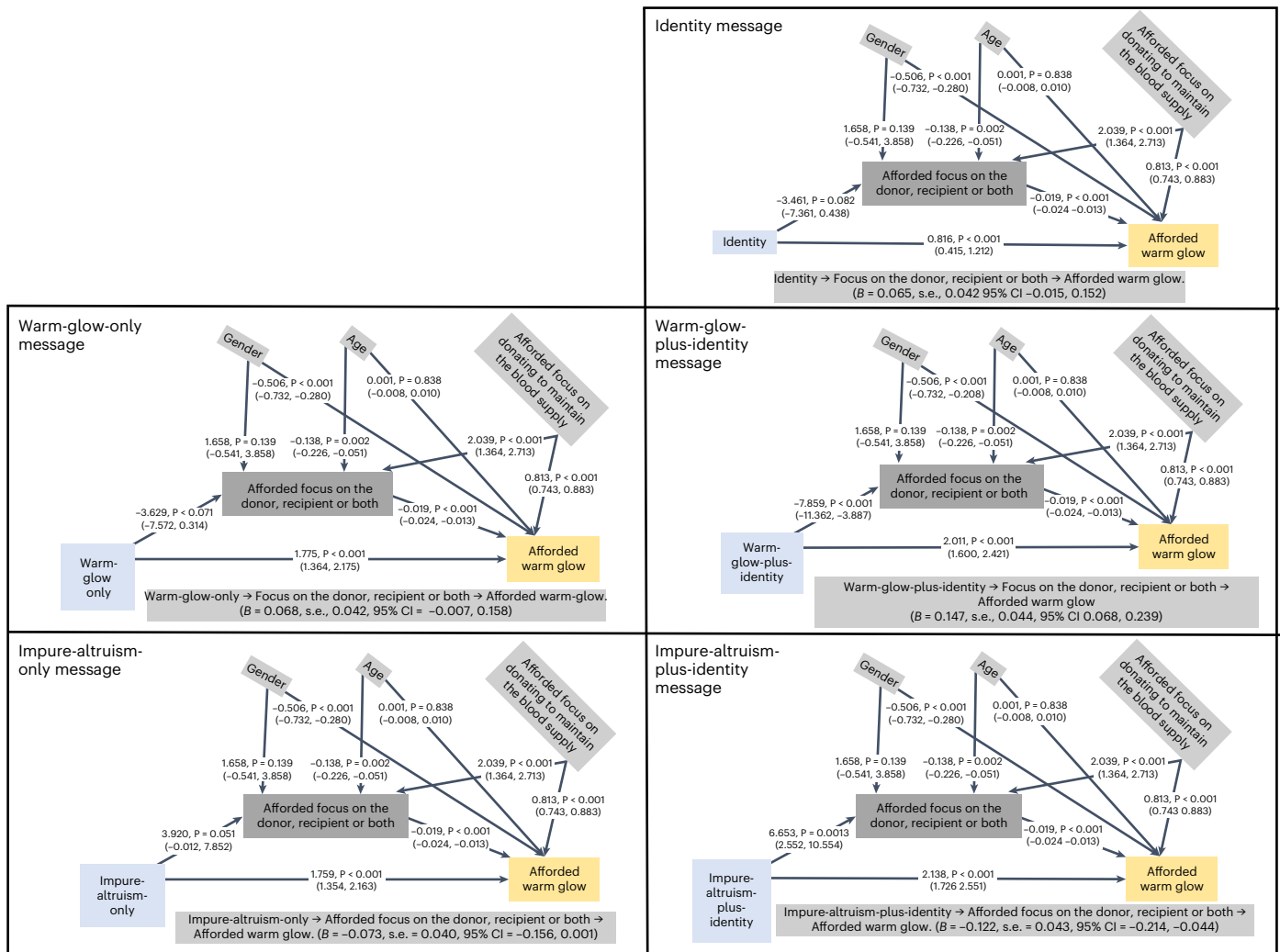


Fig. 4 | Mediation models specifying the effects of messages to predict ‘afforded warm glow’ directly and indirectly via a focus on the donor, recipient or both (Study 6). These models specify the effects of messages (X) to predict the outcome (Y) ‘afforded warm glow’ directly and indirectly via focus (‘afforded focus on the donor, recipient or both’ (the mediator, M)). These models were specified in PROCESS 4.0, with estimates based on 5,000 bootstraps

with a multinomial X (messages) and BAU as the reference ($n = 1,521$). Age, gender and supply (‘afforded focus on donating to maintain blood supply’) were specified as founders of M and Y . Predicting M from X has an R^2 of 0.086, $P < 0.001$. Predicting Y from both X and M has an R^2 of 0.333, $P < 0.001$. All analyses were two-tailed and no adjustments were made for multiple comparisons.

and the recipient, and +50 is focus on the recipient) and (3) perceptions about donating to maintain blood supply (1, not at all; to 7, completely). We included two control conditions: a simple reminder BAU control condition and an identity-prime-only condition.

Supporting H_8 , the warm-glow and impure-altruism messages were judged to afford more warm glow than BAU (Supplementary File 6 and Fig. 5, and the planned contrasts in File 6). Warm-glow messages were judged to afford slightly lower warm glow than the impure-altruism messages (Supplementary Table 13). Supporting H_9 , exposure to an identity prime resulted in greater afforded warm glow ($F_{(1,1,047)} = 24.126; P < 0.001; \eta_p^2 = 0.023; Mean_{identity} = 11.714; 95\% CI = 11.510, 11.917$ vs $Mean_{no-identity} = 11.418; 95\% CI = 11.220, 11.615$). Supporting H_{10} , impure-altruism messages compared with warm-glow messages afforded higher ratings for donating to maintain blood supply ($F_{(1,1,048)} = 106.842; P < 0.001; \eta_p^2 = 0.093; Mean_{impure-altruism} = 4.988; 95\% CI = 4.846, 5.131$ vs $Mean_{warm glow} = 3.917; 95\% CI = 3.772, 4.062$) (Supplementary Table 14). Supporting H_{11} , warm-glow messages afforded greater relative donor focus than the impure-altruism messages ($F_{(1,1,008)} = 92.500; P < 0.001; \eta_p^2 = 0.084; Mean_{warm glow} = -29.299; 95\%$

$CI = -31.277, -27.372$ vs $Mean_{impure-altruism} = -15.959; 95\% CI = -17.881, -14.038$) (Supplementary Table 15). Thus, warm-glow messages are more donor-focused and impure-altruism messages focus on both the donor and the recipient as well as on maintaining blood supply.

A series of mediation models with treatment message (X) predicting afforded warm glow (Y) directly and via an afforded focus on the donor, recipient or both (M) (Fig. 4) explored the hypothesis that the warm-glow-plus-identity message affords feelings of warm glow through a focus on the donor more than the recipient. All the messages, compared with the BAU control, predicted increased warm glow, as did a relative focus on the donor. There was a significant positive indirect effect of exposure to the ‘warm-glow-plus-identity’ message on levels of afforded warm glow via an afforded relative focus on the donor ($B = 0.147, s.e. = 0.044; 95\% CI = 0.068, 0.239$). In contrast, the significant indirect effect of exposure to an ‘impure-altruism-plus-identity’ message was negative ($B = -0.122, s.e. = 0.043; 95\% CI = -0.214, -0.044$). That is, exposure to an impure-altruism-plus-identity message is associated with an afforded greater recipient focus, linked to reduced afforded warm glow. There were no other significant indirect effects.

Therefore, the effectiveness of the ‘warm-glow-plus-identity message’ is attributable to experienced warm glow being associated with an increased relative focus on donor identity.

Discussion

We showed that warm glow is an important mechanism for sustaining high-cost cooperation directed towards a non-reciprocating stranger. We predicted that warm glow would decay and that exogenous warm-glow messages would act to remotivate cooling cooperators, especially among those who had not behaviourally committed to future cooperation. However, we found evidence that warm glow is stable over time. Therefore, we propose a revised model, whereby there are warm cooperators and cool (rather than cooling) cooperators, with a behavioural commitment to cooperate (booking another appointment) being associated with a warm cooperator, resulting in an increased likelihood of future cooperation.

Our observation that warm glow is stable is consistent with recent reports that warm glow, following blood donation, remains stable over several days^{42,68}. However, further studies are needed to measure warm glow over time, both in the lab and in the field, to assess patterns of stability and decay for different types of helping and what predicts any variation in these patterns. Indeed, levels of experienced warm glow are proportional to the cost of cooperation^{23,26} and blood donation is a costly act of cooperation that is therefore likely to generate high levels of warm glow⁶⁹. This may account for why experienced warm glow remains stable, even for cool cooperators. Less costly acts of cooperation may generate less warm glow that may decay⁶³. Thus, we identified stable warm and cool cooperators of a high-cost cooperative act and reported evidence that warm glow among warm cooperators facilitated a smooth transition to repeat cooperation^{51,59,70}. We show that this facilitation happens through two complementary mechanisms. First, a reinforcing route (λ)^{36,49,51} that signals a future reward²³. Second, warm cooperators take behavioural steps through commitment to cooperate that facilitates the likelihood of returning to donate again^{27,57}. In the context of blood donation, this is exemplified by booking the next donation, which minimizes derailment of future cooperation⁵⁸ and potentially reinforces the donor’s self-image as a committed donor⁵⁵. Thus, in the absence of an exogenous intervention, warm cooperators are more likely to return and cool cooperators drop out.

For any public health intervention to be effective, they have to contain active ingredients⁶⁷. We tested predictions for the active ingredient of ‘warm-glow-plus-identity’ messaging and showed that all messaging afforded warm glow, which was increased most when this messaging was donor-focused rather than focused on benefit to the recipient. Thus, encouraging other repeat acts of cooperation at a societal level (for example, vaccination, social distancing, mask-wearing)^{71–78} may be achieved most effectively via warm-glow messages that focus not only on the feeling of warm glow but also link the warm glow directly to the person’s identity as a cooperator. For example, the altruism vaccination hypothesis suggests that people are motivated to get vaccinated to protect others^{71–74} (but see ref. 75). Emphasizing warm glow linked to the person’s identity as a good citizen may enhance the effectiveness of such messages (for example, “Be the person who makes a positive contribution to helping our society, get a vaccination to avoid spreading the flu and feel the warm glow from helping.”). Volunteer blood donation is sustained by 3–4% of the eligible population donating blood at any one time^{29,67}. Many other cooperative acts aimed to help the wider society (for example, helping in a natural disaster or war zone) also rely on a few volunteers. Thus, our findings suggest that warm glow may be critical to ‘sustaining’ high-cost humanitarian aid in a small group of volunteers.

Improving conversion rates from first to repeat donations is a crucial objective for transfusion services worldwide, as repeat donors lead to savings on recruitment costs and improved donor safety^{79–82}. Warm-glow messages offer a simple, low-cost and effective means to help achieve this.

Furthermore, behavioural commitment to cooperation by booking another appointment to donate was a very powerful predictor of future prosocial behaviour, therefore making it easier for people to book another appointment will be beneficial. Taking this further, facilitating commitments to cooperate in other prosocial areas (for example, signing up for a flu vaccination appointment, making a diary commitment for volunteering) would also be likely beneficial.

Although the effect sizes reported in this paper are small for experienced warm glow and exogenous warm glow, a small effect size should be interpreted as important when the behaviour is hard to change (for example, making a repeated donation) but importantly, the intervention is simple and low cost, and the population effect is large⁸³. Indeed, we showed that the ‘warm-glow-plus-identity’ message resulted in an extra 3,049 donations in the year following implementation (Supplementary File 3).

Our findings show that the effect of warm glow extends to cooperative behaviours beyond donating money to the donation of bodily tissue (blood). This is interesting as these acts differ in substantive ways. Many people donate money and this is often low cost in terms of time, effort and discomfort but can range from low to high cost in terms of finances^{84–87}. However, few donate blood, which is generally high cost in terms of time, effort and potential discomfort⁸⁸. Notably, those who donate only blood or only money differ on several demographic (blood donors are younger, more likely to be childless) and psychological characteristics (blood donors are less interested in politics and are more risk-taking)⁸⁷. Furthermore, as levels of experienced warm glow are proportional to the cost of cooperation^{23,26}, blood donation is likely to generate greater warm glow than less costly forms of cooperation⁸⁸. It may be the case that a threshold exists for experienced warm glow, which once passed triggers sustained helping. Indeed, such a threshold model that becomes self-reinforcing has been suggested to sustain blood donor behaviour⁶⁷.

Finally, we acknowledge a number of limitations in the reported studies. First, we used secondary data analyses in 3 studies (studies 3 to 5) to explore the association between experienced warm glow, booking status and behaviour. These studies were not designed to assess these associations specifically and, as such, replication of these findings are needed, especially using more widely ranging assessments of experienced warm glow and emotional reactions to blood donation²³. Second, we assess the temporal dynamics of experienced warm glow at two timepoints, and it would be informative to assess warm glow across a greater number of timepoints post-donation and into subsequent donations to gain a more detailed understanding of the temporal dynamics of experienced warm glow as a function of booking status and time. Third, we did not assess exogenously manipulated warm glow and experienced warm glow together, hence we could not test directly whether the exogenous manipulation of warm glow enhances experienced warm glow in blood donors. However, the pattern of results across the studies indicates that the manipulation of warm glow, especially when donor identity is primed, does afford experienced warm glow and that this is associated with booking and return behaviour.

Methods

Ethics approvals and consent

Ethics approval for studies 1, 3, 4 and 5 was granted by the Australian Red Cross Lifeblood (formerly known as the Australian Red Cross Blood Service) Human Research Ethics Committee. Approvals were granted on 7 May 2018 for study 1 (Davison 04052018), 31 July 2020 for study 3 (Guerin 31072020), 18 September 2015 (2015#07) for study 4 and 15 November 2016 (2016#24) for study 5. Study 6 was approved by the University of Nottingham School of Psychology Ethics Committee (S1309). Informed consent was provided for studies 3 to 6. All donors attending to donate at the Australian Red Cross Lifeblood signed an individual general declaration consenting to assist blood donor research and as

such, the Australian Red Cross Lifeblood Human Research Ethics Committee did not require additional consent for study 1 as the messages were sent out as part of the normal donor recruitment protocol and posed minimal risk. Study 2 drew on data aggregated at the population level, collected by the Australian Red Cross Lifeblood. As all donors signed a general declaration consenting to assist blood donor research when they attend to donate, no additional ethics approval was required for study 2.

Study 1 Field-based experiment

Participant compensation. Participants did not receive compensation for this study.

Sampling and sample. A sample of 5,821 (Mean age = 31.94, s.d. = 11.96; % women = 60.7%) new whole-blood donors from across Australia with A-, A+, O- and O+ blood types (based on the Australian Red Cross Lifeblood donor segmentation policy) were recruited. However, A and O positive and negative groups are the most prevalent, accounting for 86% of Australian blood donors. Donors who had made their first whole-blood donation 6 weeks earlier and who had not donated previously in Australia were recruited. For each week of the field-based experiment, the entire eligible population was sampled, with all donors who met the eligibility criteria being selected and randomly allocated to conditions using simple randomization. This process continued until a priori target numbers per condition were achieved. Twenty donors were excluded: 12 due to email bounces and 8 due to being permanently deferred from donating blood after their initial donation, leaving a final sample of $n = 5,801$ (Mean age = 31.92, s.d. = 11.95; % women = 60.7%; see Supplementary File 7 for detail on sample bias and randomization checks, and Supplementary Fig. 9). This sample is generally representative when compared to the average age (32.5 yr) and percentage of women (56%) for all first-time whole-blood donors in the three pre-implementation windows in Study 2.

Procedure. New A-, A+, O- and O+ donors were recruited to the trial when they made their first donation between 16 April 2018 and 8 July 2018, with the first message issued on 22 May 2018. Donors were randomly allocated (using simple randomization) to one of four active arms and a BAU control arm. The control arm contained no message. All other communications from Lifeblood, including an SMS reminder that the donor could donate whole blood again sent 12 weeks post-donation, were identical across trial arms. Thus, the only difference between conditions was message type (warm-glow, impure-altruism) crossed with the donor identity prime (“... that’s when you became a blood donor ...”: yes, no) (see Fig. 2 and Supplementary File 2 for details of message structure).

Measures. We measured donor characteristics including (1) age, (2) gender (0 = men; 1 = women), (3) blood group (ABO: A-, A+, O-, O+) and (4) whether they had booked to make their next donation immediately after their first donation (scored 1) or had not initially booked (scored 0).

Donor return behaviour. Donor return was measured as verified attendance at the donor centre to make a high-cost donation (whole blood or plasma) within 3 months after becoming eligible. These data were collected using eProgesa 5.03. The data analysts who extracted the attendance data were blind to experimental treatments.

Power analysis. Warm glow has a small effect on predicted blood donor attendance^{53,55}. To achieve 80% power to detect a small effect with an alpha of 0.05 in a simple regression model with 8 predictors (for example, message type and identity prime conditions, age, gender, blood group, booking status and terms for the interaction of booking status with experimental conditions), 757 donors per arm are needed. Thus,

we aimed for 1,000 new donors per arm to allow for ad hoc exploratory analyses.

Pre-registration. The trial was pre-registered on the Open Science Framework (OSF reference: <https://osf.io/5m69k>). Pre-registered predictions are detailed in Table 1.

Study 2 Implementation study

Participant compensation. Participants did not receive compensation for this study.

Design. The most effective message of those tested in the field-based experiment was selected by Lifeblood to be rolled out nationally on 9 July 2018, with effectiveness measured by marketing click-through data (the highest proportion of click-to-open) from the original recruitment sample. Analysis of attendance data collected in the field-based experiment confirmed the click-rate message choice (see Supplementary File 8 and Fig. 10 for the timeline of the field-based experiment and implementation analysis). The implementation sample included all new donors across Australia who had donated whole blood, excluding donating for autologous and therapeutic reasons. We collated aggregate data on whether they had attended to make a second high-cost donation (whole blood or plasma) within 3 months and whether they had initially booked the next appointment in-centre. We compared the frequency of return in the 3 yr immediately before message roll-out with the frequency of return in two time windows after implementation. The 3-yr pre-implementation period covered three 12-month time windows: beginning on 16 April to 15 April for years (1) 2015–2016 ($n = 90,317$; Mean age = 32.0, s.d. = 13.6; % women = 55%, $n = 49,778$), (2) 2016–2017 ($n = 93,430$; Mean age = 32.5, s.d. = 13.4; % women = 56%, $n = 52,539$) and (3) 2017–2018 ($n = 86,606$; Mean age = 33.1, s.d. = 13.0; % women = 57%, $n = 49,531$). The post-implementation period covered two slightly overlapping time windows: (1) 9 July 2018 to 8 July 2019 ($n = 81,766$; Mean age = 34.9, s.d. = 13.5; % women = 57%, $n = 46,689$) and (2) 16 April 2019 to 15 April 2020 ($n = 85,551$; Mean age = 34.3, s.d. = 13.2; % women = 57%, $n = 50,480$). The first post-implementation window covers the 12 months from the moment the selected warm-glow message went live across Australia and the second window covers the April-to-April window consistent with the timing of the pre-implementation time windows.

Pre-registration. This study was not pre-registered. Rather, data were collated from the Australian Red Cross Lifeblood (via eProgesa 5.03) on a yearly aggregate level to test the prediction from the main finding of the field-based experiment that a warm-glow-plus-identity message would encourage more second-time donations in those who have not booked.

Study 3 Booking, warm glow and pure altruism

Participant compensation. Five AU\$100 gift vouchers were awarded after a random draw of participants.

Rationale. This study was designed primarily to examine blood donors’ concerns and attitudes towards blood donation during the coronavirus disease 2019 (COVID-19) pandemic. We used secondary data analysis to explore the association between anticipated warm glow (and other anticipated emotional states), pure altruism and booking status in first-time donors, as these were the focus of the field-based experiment.

Sample. Seven-hundred and sixteen first-time donors took part (Mean age = 41.20, s.d. = 14.85; % women = 65.5%, $n = 469$ (see Supplementary File 7 for sampling strategy)); the current sample was older and had a higher percentage of women.

Procedure. Donors were sent a de-identified online survey invitation between 12 August and 24 September 2020. A reminder was sent on 27

August 2020. Data were extracted on 24 September 2020. The survey was hosted on Qualtrics (<https://www.qualtrics.com/>). The survey was sent out after donors had made their donations and had the opportunity to book their next donation, but the time between booking and survey completion varied.

Measures. The following variables were used.

'Anticipated warm glow' was indexed using adjectives of the type that have been used to assess warm glow in previous research (for example, refs. 24,44). Specifically, donors were asked "What do you think donating blood in the future during the coronavirus pandemic would be like?", to which they responded using 12 adjectives (Unrewarding, Pointless, Displeasing, Negative, Stressful, Unsatisfying, Rewarding, Worthwhile, Pleasing, Positive, Relaxing, Satisfying), each measured on a 4-point scale (1 = Not at all, 2 = Slightly, 3 = Moderately and 4 = Extremely). Of these adjectives, Rewarding, Worthwhile, Pleasing, Positive and Satisfying represent the type of adjective that have been used to assess warm glow in previous research (for example, ref. 24). A principle-axis-factor analysis with oblique rotation (Supplementary File 4 and Table 6) applied to these 12 adjectives indicated a 3-factor solution, with the 'anticipated warm glow' adjectives Rewarding, Worthwhile, Pleasing, Positive and Satisfying ($\alpha = 0.77$) forming one factor²⁴, the adjectives Relaxing and Stressful forming a factor whereby high scores equated to relaxing (termed 'Anticipated calmness'; mean inter-item correlation = 0.34)⁸⁹ and the adjectives Pointless, Negative, Unrewarding, Unsatisfying and Displeasing forming the final factor, which was termed 'anticipated negative affect' ($\alpha = 0.68$)⁸⁹.

'Pure altruism' was assessed with two items concerning the donor's motivation for their last donation ("To help someone in need" and "It's important for society"), each measured on a 5-point scale (1 = Not important at all, 2 = Slightly important, 3 = Moderately important, 4 = Very important, 5 = Extremely important).

Booking status. We extracted data about whether donors had booked another appointment after their last donation from participants' linked but de-identified donor records ($n = 294$, 41.1% booked).

Additional selection model predictor. As the study took place during the COVID-19 pandemic, we included a question that assessed whether the donor had had a COVID-19 test ("Have you ever been tested for coronavirus?": Yes, No).

Power analysis. There was no data on which to base a power calculation. However, if we assume a small effect size, then we would need 363 donors who booked and 363 who did not ($n = 726$) with a power of 0.80 and $\alpha = .05$. In our sample, 294 booked and 422 did not. As such, the overall sample size was sufficient to detect a small effect.

Pre-registration. This study was not pre-registered.

Study 4 Warm glow, booking status, donor status and type

Participant compensation. Participants did not receive compensation for this study.

Rationale. The study was designed primarily to examine how donors' emotions change across a single donation. The study also included a measure of donors' intrinsic motivation to donate blood⁹⁰. Within the literature, warm glow has been conceptualized as intrinsic motivation^{24,44,91-93}, and intrinsic regulation from self-determination theory has been linked to assessing warm glow⁹³. Thus, warm glow was assessed using a subscale designed to measure intrinsic regulation for blood donation. In this study, we explored the association between booking and warm glow, controlling for donor status (first-time vs novice) and type (plasma vs whole blood).

Sample. In total, 1,124 donors (Mean age = 30.27, s.d. = 11.51; % women = 55.2%, $n = 621$ women) participated. Of these, 401 were first-time whole-blood donors (Mean age = 30.38, s.d. = 11.79; % women = 54.6%, $n = 219$), 208 were first-time plasma donors (Mean age = 31.07, s.d. = 11.50; % women = 51.0%, $n = 106$), 379 were novice whole-blood donors (Mean age = 29.61, s.d. = 11.49; % women = 60.4%, $n = 229$) and 136 were novice plasma donors (Mean age = 30.53, s.d. = 10.75; % women = 49.3%, $n = 67$) (see Supplementary File 7 for sampling strategy and details on sample bias). The sample was representative in terms of the percentage of women but slightly younger than the average age (32.5 yr) for first-time whole-blood donors in the three pre-implementation windows in Study 2.

Procedure. Participating donors were approached by a different researcher in each of three donation centres and prompted to complete questionnaires at several timepoints across their donation. The assessment of warm glow (that is, intrinsic regulation) was completed in the refreshment area after donation. Donors could book during the session. Data were collected between April 2016 and October 2017.

Measures. The following indices were assessed.

Warm glow (intrinsic motivation). This was assessed using the intrinsic regulation subscale of the Blood Donor Identity Survey (BDIS)⁹⁰, which comprises 3 items ("I enjoy donating blood", "For me, being a blood donor means more than just donating blood" and "Blood donation is an important part of who I am"), with each item measured on a 7-point scale (1 = Not at all true, 7 = Very true). These items reflect the psychological concept of warm glow^{24,44} and were summed to yield a warm-glow score, with higher scores indicating higher subjective warm glow (Mean = 14.66, s.d. = 4.42, $\alpha = 0.80$).

Booking status. We gained data about whether donors had booked another appointment during the donation session when they completed the survey from participants' linked but de-identified donor records (781, 69.5% booked).

Donor return behaviour. Donor return was measured as verified attendance to donate within 6 months of the index donation, collected using eProgesa 5.03 (670 or 59.6% returned to donate).

Power analysis. There was no data on which to base a power calculation. However, if we assume a small effect size, then we would need 363 donors who booked and 363 who did not ($n = 726$) with a power of 0.80 and $\alpha = .05$. In our sample, 781 booked and 343 did not. As such, the overall sample size was sufficient to detect a small effect.

Pre-registration. This study was not pre-registered.

Study 5 Temporal stability of warm glow following booking

Participant compensation. Participants could opt-in to receive AU\$5 compensation for every survey completed.

Rationale. The study was designed primarily to examine how donors' emotions (reflective and anticipated) changed following their first donation. Warm glow was assessed at two of the four study timepoints: 1 day after donating blood (wave 1) and 12 weeks later (wave 4). We used these data for secondary data analysis to assess: (1) whether booking influenced warm glow, (2) the stability of warm glow over time and (3) whether the stability of warm glow was moderated by booking status.

Sample. Nine-hundred and thirty-two first-time donors who completed the survey provided warm-glow responses at waves 1 and 4 (Mean age = 36.36, s.d. = 14.02; % women = 70.3%, $n = 655$) (see Supplementary File 7 for sampling strategy and details on sample bias). Compared with

the average age (32.5 yr) and percentage of women (56%) for first-time whole-blood donors in the three pre-implementation windows in Study 2, this sample was slightly older and had a higher percentage of women.

Procedure. In a longitudinal design, data were collected in four waves following donors' first donation. The first (wave 1) assessments were completed approximately 1 day after the donor's first donation. Participants who opted-in for follow-up were invited to complete assessments at three subsequent timepoints: 4 (wave 2), 8 (wave 3) and 12 (wave 4) weeks after their initial donation. Warm glow (intrinsic motivation) was measured in waves 1 and 4. All data were collected between January 2017 and September 2017.

Measures. The following measures were assessed.

Warm glow (intrinsic motivation). This was assessed at waves 1 and 4 using the BDIS⁸⁸ as in Study 4 ($Mean_{wave1} = 14.99$, $s.d._{wave1} = 4.15$, $\alpha_{wave1} = 0.76$; $Mean_{wave4} = 15.21$, $s.d._{wave4} = 4.07$, $\alpha_{wave4} = 0.76$).

Booking status. We gained data about whether donors had booked another appointment after their last donation from participants' linked but de-identified donor records ($n = 612$, 65.7% booked).

Donor return behaviour. Return was measured as verified attendance to donate within 6 months of the index donation, collected using eProgesa 5.03 ($n = 682$, 73.2% returned to donate).

Power analysis. There was no data on which to base a power calculation. However, if we assume a small effect size, then we would need 363 donors who booked and 363 who did not ($n = 726$) with a power of 0.80 and $\alpha = .05$. In our sample, 612 booked and 320 did not. As such, the overall sample size was sufficient to detect a small effect.

Pre-registration. This study was not pre-registered.

Study 6 Validation of warm-glow and impure-altruism messages

Participant compensation. Based on Prolific's rates, participants received £1.25 compensation for completing the study.

Experimental design. The main design was a 2 (warm glow: warm glow, impure altruism) by 2 (identity: present, absent) between-subjects design. This represents the four main messages used in the field-based experiment (Study 1) (warm-glow-only, warm-glow-plus-identity, impure-altruism-only and impure-altruism-plus-identity; Fig. 2). We added two additional conditions to compare warm-glow and impure-altruism messaging with (1) a no-message BAU control ("The good news is in a few weeks you'll be ready to donate."); and (2) an identity-prime-only condition ("Most people can't remember what they were doing 6 weeks ago, but this was the day you became a blood donor. The good news is in a few weeks you'll be ready to donate again."). This was an online experiment hosted on Qualtrics (<https://www.qualtrics.com/uk/>), with participants randomly allocated using simple randomization to one of these six conditions. All data were collected on 24 March 2021.

Sample. A total of 1,592 participants were recruited through Prolific (<https://www.prolific.co/>) (Mean age = 36.47, $s.d. = 13.00$; % women = 50%, $n = 795$), with 34% ($n = 538$) indicating that they had donated blood previously and 12.1% ($n = 194$) being current donors (donated within the past 2 yr). There were 266 participants in the BAU/control, 273 in the 'identity-prime-only', 263 in the 'warm-glow-only', 256 in the 'warm-glow-plus-identity', 279 in the 'impure-altruism-only' and 255 in the 'impure-altruism-plus-identity' conditions (see Supplementary File 7 for sampling strategy and details on sample bias).

Power calculation. A Cohen's d of 0.344 for positive ratings of warm-glow vs altruistic messages for blood donation was previously recorded⁶⁶. Achieving 80% power to detect a small effect with $\alpha = 0.05$ for a 2 (warm glow: yes, no) \times 2 (identity: yes, no) design requires 265 participants per condition. For a series of pairwise comparisons of experimental groups, detecting a small effect with a Cohen's d of 0.344 requires 135 participants per condition.

Outcome measures. The following measures were assessed.

Afforded donor vs recipient focus. This was calculated as the average of 2 items ("To what extent do you think that the message focuses on the blood donor, the recipient of blood or both?" and "To what extent do you think that the message focuses on the emotions of the blood donor, the recipient of blood or both?") measured on a scale from -50 (primarily on the blood donor) to +50 (primarily on the recipient of blood) with a midpoint of 0 (equally on both the donor and recipient).

Afforded warm glow. This was calculated as the sum of 2 items ("The message makes me feel that donating blood is personally rewarding in itself." and "The message makes me feel that donating blood would make me feel like a good person.") measured on a 7-point scale (1 = not at all, to 7 = completely).

Afforded focus on maintaining blood supply. This was assessed with a single item ("The message makes me feel that donating blood would ensure that there is enough blood for all who need it.") measured on a 7-point scale (1 = not at all, to 7 = completely).

Pre-registration. This experiment was pre-registered (<https://osf.io/q4c6v/>). Pre-registered predictions are listed in Table 1.

Statistical analyses for Studies 1–6

All data were analysed using standard statistical packages (IBM SPSS v.26 and 27, ZUMStat, Psychometrica, Stata 17, MPlus 8.4, PROCESS 4.0, Comprehensive Meta-Analysis 2). All tests were two-tailed. All effect sizes for all analyses are reported using Cohen's d . Cohen's d was derived for comparison across multiple-group means using the procedures described in ref. 94 and from Z scores using the procedure described in ref. 95, with both implemented in Psychometrica⁹⁶. ORs were converted to Cohen's d using procedures described in refs. 96,97. Comparisons of percentages between groups, including interactions, used procedure detailed in refs. 98,99. Mediation models were run in MPlus and PROCESS¹⁰⁰. Meta-analysis was conducted using Comprehensive Meta-Analysis (v.2)⁹⁷. Our main analyses had dichotomous outcomes and were analysed using logistic models and non-parametric tests. When continuous variables were outcomes, we present boxplots (for example, afforded warm glow) or Q-Q plots (experienced warm glow, Study 5; see Supplementary File 9 and Table 11a,b). In all cases, there was evidence of non-normality. We analysed these data using analysis of variance (ANOVA) and ordinary least squares (OLS) regression models as Monte-Carlo studies show that these techniques are robust to non-normality especially when sample sizes are large (greater than 300 for ANOVA and 1,000 for OLS)^{101,102}. All results reported are two-tailed with no adjustments for multiple comparisons made.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data supporting the findings of Studies 1–6 are available (<https://osf.io/ps5kb/>). Source data are provided with this paper.

References

1. Rand, D. G. & Nowak, M. A. Human cooperation. *Trends Cogn. Sci.* **17**, 413–425 (2013).
2. West, S. A., Griffin, A. S. A. & Gardner, A. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.* **20**, 415–432 (2007).
3. West, S. A., Mouden, C. E. & Gardner, A. Sixteen common misconceptions about the evolution of cooperation in humans. *Evol. Hum. Behav.* **32**, 231–262 (2011).
4. Ent, M. R., Sjästad, H., Hippel, W. & Baumeister, R. Helping behaviour is non-zero-sum: helper and recipient autobiographical accounts of help. *Evol. Hum. Behav.* **41**, 210–217 (2020).
5. Peysakhovich, A., Nowak, M. A. & Rand, D. G. Humans display a ‘cooperative phenotype’ that is domain-general and temporally stable. *Nat. Commun.* **5**, 4939 (2014).
6. Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).
7. Scott-Phillips, T. C., Dickins, T. E. & West, S. A. Evolutionary theory and the ultimate–proximate distinction in the human behavioral sciences. *Perspect. Psychol. Sci.* **6**, 38–47 (2011).
8. Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).
9. Fischbacher, U., Gächter, S. & Fehr, E. Are people conditionally cooperative? Evidence from a public goods game. *Econ. Lett.* **71**, 397–404 (2001).
10. Bradley, A., Lawrence, C. & Ferguson, E. Does observability effect prosociality? A meta analytic review. *Proc. R. Soc. B* **285**, 20180116 (2018).
11. Andreoni, J. Impure altruism and donations to public goods: a theory of warm-glow giving. *Econ. J.* **100**, 464–487 (1990).
12. Nowak, M. & Sigmund, K. Evolution of indirect reciprocity. *Nature* **437**, 1291–1298 (2005).
13. Milinski, M. Reputation, a universal currency for human social interaction. *Phil. Trans. R. Soc. Lond. B* **371**, 20150100 (2016).
14. Zak, P. J., Stanton, A. A. & Ahmadi, S. Oxytocin increases generosity in humans. *PLoS ONE* **11**, e1128 (2007).
15. Fehr, E. & Fischbacher, U. Third-party punishment and social norms. *Evol. Hum. Behav.* **25**, 63–87 (2004).
16. Fehr, E. & Schmidt, K. A theory of fairness, competition and cooperation. *Q. J. Econ.* **114**, 817–868 (1999).
17. de Waal, F. B. M. Putting the altruism back into altruism: the evolution of empathy. *Annu. Rev. Psychol.* **59**, 279–300 (2008).
18. Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C. & Neuberg, S. L. Reinterpreting the empathy–altruism relationship: when one into one equals oneness. *J. Pers. Soc. Psychol.* **73**, 481–494 (1997).
19. Batson, C. D. *The Altruism Question: Toward A Social-psychological Answer* (Erlbaum, 1991).
20. Andreoni, J. Warm-glow versus cold-pickle: the effects of positive and negative framing on cooperation in experiments. *Q. J. Econ.* **110**, 1–21 (1995).
21. Andreoni, J. Giving with impure altruism: applications to charity and Ricardian equivalence. *J. Polit. Econ.* **97**, 1447–1458 (1989).
22. Henrich, J. & Muthukrishna, M. The origins and psychology of human cooperation. *Annu. Rev. Psychol.* **72**, 207–240 (2021).
23. Ferguson, E. & Flynn, N. Moral relativism as a disconnect between behavioural and experienced warm glow. *J. Econ. Psychol.* **56**, 163–175 (2016).
24. van der Linden, S. Warm glow is associated with low- but not high-cost sustainable behaviour. *Nat. Sustain.* **1**, 28–30 (2018).
25. Crumpler, H. & Grossman, P. J. An experimental test of warm-glow giving. *J. Public Econ.* **92**, 1011–1021 (2008).
26. Konow, X. Mixed feelings: theories of and evidence on giving. *J. Public Econ.* **94**, 3279–3297 (2010).
27. Ferguson, E. Mechanisms of altruism approach to blood donor recruitment and retention: a review and future directions. *Trans. Med.* **25**, 211–226 (2015).
28. Ferguson, E. & Lawrence, C. Altruistic and warm-glow motivations: variation by blood donor career. *TPM Test. Psychom. Methodol. Appl. Psychol.* **26**, 639–651 (2019).
29. Ferguson, E. What blood and organ donation can tell us about cooperation? *Curr. Opin. Psychol.* **17**, 202–207 (2021).
30. Abasolo, I. & Tsuchiya, A. Blood donation as a public good: an empirical investigation of the free-rider problem. *Eur. J. Health Econ.* **15**, 313–321 (2014).
31. Wildman, J. & Hollingsworth, B. Blood donation and the nature of altruism. *J. Health Econ.* **28**, 492–503 (2009).
32. Borkent-Raven, B. A., Janssen, M. P. & van der Poel, C. L. Demographic changes and predicting blood supply and demand in the Netherlands. *Transfusion* **50**, 2455–2460 (2010).
33. Shih, H. & Rajendran, S. Stochastic inventory model for minimizing blood shortage and outdated in a blood supply chain under supply and demand uncertainty. *J. Healthc. Eng.* **2020**, 8881751 (2020).
34. Mayo, J. W. & Tinsley, C. H. Warm glow and charitable living: why the wealthy do not give more to charity? *J. Econ. Psychol.* **30**, 490–499 (2009).
35. Goeree, J. K., Holt, C. A. & Laury, S. K. Private costs and public benefits: unraveling the effects of altruism and noisy behavior. *J. Public Econ.* **83**, 255–276 (2002).
36. Harbaugh, W. T., Mayr, U. & Burghart, D. R. Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* **316**, 1622–1625 (2007).
37. Menges, R., Schroeder, C. & Traub, S. Altruism, warm glow and the willingness-to-donate for green electricity: an artefactual field experiment. *Environ. Resour. Econ.* **31**, 431–458 (2005).
38. Ferrara, I. & Missios, P. Trust, ability-to-pay, and charitable giving. *J. Public Econ. Theory* **22**, 583–629 (2020).
39. Kuroishi, Y. & Sawada, Y. Motivations behind prosocial behavior: evidence from the Philippines. *J. Asian Econ.* **64**, 101127 (2019).
40. Gandullia, L. The price elasticity of warm-glow giving. *Econ. Lett.* **182**, 30–32 (2019).
41. Iweala, S., Spiller, A. & Meyerding, S. Buy good, feel good? The influence of the warm glow of giving on the evaluation of food items with ethical claims in the UK and Germany. *J. Clean. Prod.* **215**, 315–328 (2019).
42. O’Brien, E. & Kassirer, S. People are slow to adapt to the warm glow of giving. *Psychol. Sci.* **30**, 193–204 (2019).
43. DellaVigna, S. & Devin Pope, D. What motivates effort? Evidence and expert forecasts. *Rev. Econ. Stud.* **85**, 1029–1069 (2018).
44. Hartmann, P., Eisend, M., Apaolaza, V. & D’Souza, C. Warm glow vs. altruistic values: how important is intrinsic emotional reward in proenvironmental behavior? *J. Environ. Psychol.* **52**, 43–55 (2017).
45. Luccasen, A. & Grossman, P. J. Warm-glow giving: earned money and the option to take. *Econ. Inq.* **55**, 996–1006 (2017).
46. Bischoff, I. & Krauskopf, T. Warm glow of giving collectively – an experimental study. *J. Econ. Psychol.* **51**, 210–218 (2015).
47. Taufik, D., Bolderdijk, J. & Steg, L. Acting green elicits a literal warm glow. *Nat. Clim. Change* **5**, 37–40 (2015).
48. Tonin, M. & Vlassopoulos, M. An experimental investigation of intrinsic motivations for giving. *Theory Decis.* **76**, 47–67 (2014).
49. Piliavin, J. A., Callero, P. L. & Evans, D. E. Addicted to altruism? Opponent process theory and habitual blood donation. *J. Pers. Soc. Psychol.* **43**, 1200–1213 (1982).
50. Ferguson, E., Taylor, M., Keatley, D., Flynn, N. & Lawrence, C. Blood donors’ helping behavior is driven by warm glow: more evidence for the blood donor benevolence hypothesis. *Transfusion* **52**, 2189–2200 (2012).

51. Weiss, R. F., Boyer, J. L., Lombardo, J. P. & Stich, M. H. Altruistic drive and altruistic reinforcement. *J. Pers. Soc. Psychol.* **25**, 390–400 (1973).
52. Neumann, R. The framing of charitable giving: a field experiment at bottle refund machines in Germany. *Ration. Soc.* **31**, 98–126 (2019).
53. List, J. A. et al. An experimental test of fundraising appeals targeting donor and recipient benefits. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-021-01095-8> (2021).
54. Lokhorst, A. M., Werner, C., Staats, H., van Dijk, E. & Gale, J. L. Commitment and behavior change: a meta-analysis and critical review of commitment-making strategies in environmental research. *Environ. Behav.* **45**, 3–34 (2013).
55. Bénabou, R. & Tirole, J. Incentives and prosocial behavior. *Am. Econ. Rev.* **96**, 1652–1678 (2006).
56. White, K. M., Poulsen, B. E. & Hyde, M. K. Identity and personality influences on donating money, time, and blood. *Nonprofit Volunt. Sect. Q.* **46**, 372–394 (2017).
57. Bednall, T. C. & Bove, L. L. Donating blood: a meta-analytic review of self-reported motivators and deterrents. *Transfus. Med. Rev.* **25**, 317–334 (2011).
58. Achtziger, A., Gollwitzer, P. M. & Sheeran, P. Implementation intentions and shielding goal striving from unwanted thoughts and feelings. *Pers. Soc. Psychol. Bull.* **34**, 381–393 (2008).
59. Bindl, U. K., Parker, S. K., Totterdell, P. & Hagger-Johnson, G. Fuel of the self-starter: how mood relates to proactive goal regulation. *J. Appl. Psychol.* **97**, 134–150 (2012).
60. Davidson, R. J., Jackson, D. C. & Kalin, N. H. Emotion, plasticity, context, and regulation: perspectives from affective neuroscience. *Psychol. Bull.* **126**, 890–909 (2000).
61. Davidson, R. J. Comment: affective chronometry has come of age. *Emot. Rev.* **7**, 368–370 (2015).
62. Delton, A. W., Cosmides, L., Guemo, M., Robertson, T. E. & Tooby, J. The psychosemantics of free riding: dissecting the architecture of a moral concept. *J. Pers. Soc. Psychol.* **102**, 1252–1270 (2012).
63. Gangadharan, L., Harrison, G. W. & Leroux, A. D. Are risks over multiple attributes traded off? A case study of aid. *J. Econ. Behav. Organ.* **164**, 166–198 (2019).
64. Hemenover, S. H. Individual differences in rate of affect change: studies in affective chronometry. *J. Pers. Soc. Psychol.* **85**, 121–131 (2003).
65. Ferguson, E., Atsma, F., de Kort, W. & Veldhuizen, I. Exploring the pattern of blood donor beliefs in first time, novice and experienced donors: differentiating reluctant altruism, pure altruism, impure altruism and warm-glow. *Transfusion* **52**, 343–355 (2012).
66. Ferguson, E., Farrell, K. & Lawrence, C. Blood donation is an act of benevolence than altruism. *Health Psychol.* **27**, 327–336 (2008).
67. Ferguson, E. Strategies and theories to attract and retain blood donors: fairness, reciprocity, equity, and warm-glow. *ISBT Sci. Ser.* **16**, 219–225 (2021).
68. Sweegers, M. G., Twisk, J. W. R., Quee, F. A., Ferguson, E. & van den Hurk, K. Whole blood donors' post-donation symptoms diminish quickly but are discouraging: results from 6-day symptom diaries. *Transfusion* **61**, 811–821 (2021).
69. Ferguson, E., Zhao, K., O'Carroll, R. E. & Smillie, L. D. Costless and costly pro-sociality: correspondence among personality traits, economic preferences, and real world pro-sociality. *Soc. Psychol. Pers. Sci.* **10**, 461–471 (2019).
70. Aknin, L. B., Van de Vondervoort, J. W. & Hamlin, J. K. Positive feelings reward and promote prosocial behavior. *Curr. Opin. Psychol.* **20**, 55–59 (2018).
71. Brewer, N. T., Chapman, G. B., Rothman, A. J., Leask, J. & Kempe, A. Increasing vaccination: putting psychological science into action. *Psychol. Sci. Public Interest* **18**, 149–207 (2017).
72. Betsch, C., Böhm, R. & Korn, L. Inviting free-riders or appealing to prosocial behavior? Game-theoretical reflections on communicating herd immunity in vaccine advocacy. *Health Psychol.* **32**, 978–985 (2013).
73. Betsch, C., Böhm, R., Korn, L. & Holtmann, C. On the benefits of explaining herd immunity in vaccine advocacy. *Nat. Hum. Behav.* **1**, 0056 (2017).
74. Li, M., Taylor, E. G., Atkins, K. E., Chapman, G. B. & Galvani, A. P. Stimulating influenza vaccination via prosocial motives. *PLoS ONE* **11**, e0159780 (2016).
75. Isler, O., Isler, B., Kopsacheilis, O. & Ferguson, E. Limits of the social-benefit motive among high-risk patients: a field experiment on influenza vaccination behaviour. *BMC Public Health* **20**, 240 (2020).
76. Jordan, J. J., Yoeli, E. & Rand, D. G. Don't get it or don't spread it: comparing self-interested versus prosocial motivations for COVID-19 prevention behaviors. *Sci. Rep.* **11**, 20222 (2021).
77. Campos-Mercade, P., Meier, A., Schneider, F. & Wengström, E. Prosociality predicts health behaviors during the COVID-19 pandemic. SSRN <https://doi.org/10.2139/ssrn.3604094> (2020).
78. Luttrell, A. & Petty, R. E. Evaluations of self-focused versus other-focused arguments for social distancing: an extension of moral matching effects. *Soc. Psychol. Pers. Sci.* **12**, 946–954 (2020).
79. Zou, S., Stramer, S. L. & Dodd, R. Y. Donor testing and risk: current prevalence, incidence, and residual risk of transfusion-transmissible agents in US allogenic donations. *Transfus. Med. Rev.* **26**, 119–128 (2012).
80. Lucky, T. T. A. et al. Trends in transfusion-transmissible infections among Australian blood donors from 2005 to 2010. *Transfusion* **53**, 2751–2762 (2013).
81. Natachit, N. et al. Temporal trends in the prevalence of HIV and other transfusion-transmissible infections among blood donors in northern Thailand, 1990 through 2001. *Transfusion* **43**, 730–735 (2002).
82. Schreiber, G. B. et al. First-year donation patterns predict long-term commitment for first time donors. *Vox Sang.* **88**, 114–121 (2005).
83. Prentice, D. A. & Miller, D. T. When small effects are impressive. *Psychol. Bull.* **112**, 160–164 (1992).
84. *Charitable Giving Statistics* (National Philanthropic Trust,); <https://www.nptrust.org/philanthropic-resources/charitable-giving-statistics/>
85. *CAF World Giving Index. Ten Years of Giving Trends, October 2019* (Charities Aid Foundation,); https://www.cafonline.org/docs/default-source/about-us-publications/caf_wgi_10th_edition_report_2712a_web_101019.pdf
86. Davidai, S., Gilovich, T. D. & Ross, L. D. Defaults, meaning, and behavior. *Proc. Natl Acad. Sci. USA* **109**, 15201–15205 (2012).
87. Shehu, E. R., Langmaack, A. C., Felchle, E. & Clement, M. Profiling donors of blood, money and time. *Nonprofit Manage. Leadersh.* **25**, 269–295 (2015).
88. Lyle, H. F. III, Smith, E. A. & Sullivan, R. J. Blood donations as costly signals of donor quality. *J. Evol. Psychol.* **7**, 263–286 (2009).
89. Russell, J. A., Weiss, A. & Mendelsohn, G. A. Affect Grid: a single-item scale of pleasure and arousal. *J. Pers. Soc. Psychol.* **57**, 493–502 (1989).
90. France, C. R. et al. The blood donor identity survey: a multidimensional measure of blood donor motivations. *Transfusion* **54**, 2098–2105 (2014).
91. Clark, C. F., Kotchen, M. J. & Moore, M. R. Internal and external influences on pro-environmental behavior: participation in a green electricity program. *J. Environ. Psychol.* **23**, 237–246 (2003).
92. Taufik, D., Bolderdijk, J. & Steg, L. Acting green elicits a literal warm glow. *Nat. Clim. Change* **5**, 37–40 (2015).

93. Ferguson, E., Murray, C. & O'Carroll, R. E. Blood and organ donation: health impact, prevalence, correlates and interventions. *Psychol. Health* **34**, 1073–1104 (2019).
94. Cohen, J *Statistical Power Analysis for the Behavioral Sciences*. (Routledge Academic, 1988).
95. Fritz, C. O., Morris, P. E. & Richler, J. J. Effect size estimates: current use, calculations, and interpretation. *J. Exp. Psychol. Gen.* **141**, 2–18 (2012).
96. Lenhard, W. & Lenhard, A. Calculation of effect sizes. *Psychometrika* https://www.psychometrika.de/effect_size.html (2016).
97. Borenstein, M., Hedges, L. V., Higgins, J. P. T & Rothstein, H. R. *Introduction to Meta-Analysis*. (Wiley, 2009).
98. Agresti, A *An Introduction to Categorical Data Analysis*. (Wiley, 1996).
99. Rosenthal, R & Rosnow, R. L. *Contrast Analysis. Focused Companions in the Analysis of Variance*. (Cambridge Univ. Press, 1985).
100. Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach* (Guilford Press, 2013)
101. Blanca, M. J., Alarcón, R., Arnau, J., Bono, R. & Bendayan, R. Non-normal data: is ANOVA still a valid option? *Psicothema* **9**, 552–557 (2017).
102. Knief, U. & Forstmeier, W. Violating the normality assumption may be the lesser of two evils. *Behav. Res.* **53**, 2576–2590 (2021).

Acknowledgements

Studies 4 and 5 were supported by the Australian Government through the Australian Research Council's Linkage Projects funding scheme (project LP140100034: L.A.W., B.M., T.E.D., E.F.). Study 6 was supported by the School of Psychology University of Nottingham (E.F., C.L.). Studies 1 to 5, were supported by Australian Red Cross Lifeblood (Australian governments fund the Australian Red Cross Lifeblood to provide blood, blood products and services to the Australian community) (B.M., T.E.D., C.N.G., A.R., K.N., N.G.). The funders had no role in study design, data collection and analysis, decision to publish or the preparation of the manuscript. We thank O. Trencher and K. Jensen for assistance with data collection for Study 4.

Author contributions

E.F., C.L., C.N.G., A.R., K.N., B.M. and T.E.D. conceived the overall original idea paper and designed the study protocol and materials for Studies 1, 2 and 6. S.B. contributed to the design and implementation of Study 6. A.R., K.N. and C.N.G. conducted the day-to-day running of Study 1. L.A.W., A.v.D., B.M., T.E.D. and A.T. designed the study protocol and materials for Studies 4 and 5. A.v.D. and A.T. conducted the day-to-day running of Studies 4 and 5, and carried out initial data extraction and processing. N.G. contributed to the design of the study protocol and measures of Study 3, cleaning and preparation of datasets, and to planned analyses of Study 6 in the revised paper.

C.N.G. prepared the data files for Studies 1 and 2. E.F. conducted the analyses of Studies 1–6. E.F. drafted the initial and revised manuscript, and all authors contributed to revising the manuscript over a number of iterations.

Competing interests

C.L. is a consultant operating as a sole trader at Lawrence PsychAdvisory, a for-profit entity. At the time of the research, C.N.G., A.R., K.N., A.T., N.G., B.M., T.E.D. were affiliated with the Australian Red Cross Lifeblood that partly funded the research (funding Studies 1, 2 and 3, as well as partly funding Studies 4 and 5). The Australian Red Cross Lifeblood, while partially funding aspects of the research, had no direct input to the study design, data collection and analysis, decision to publish or preparation of the manuscript. Since the completion of this research, S.B. had become a research manager for Kantar Public UK, and T.E.D. is the Director of Research Discovery for Silverchain in Australia. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41562-023-01687-6>.

Correspondence and requests for materials should be addressed to Eamonn Ferguson.

Peer review information *Nature Human Behaviour* thanks James Andreoni and Mark Conner for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- | | | |
|-------------------------------------|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection eProgesia 5.03 (<http://www.mak-system.com/blood-centers/>), Prolific UK, (<https://www.prolific.co/>), Qualtrics (<https://www.qualtrics.com>)

Data analysis IBM SPSS version 26 & 27, Psychometrica, Zumastat 4.0, MPlus_8.4, Stata SE 17, PROCESS 4.0, Comprehensive Meta Analysis v2

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The authors declare that all data supporting the findings of this study are available within the paper, supplementary files and at Ferguson, Eamonn, Barbara Masser, Carley N Gemelli, Lisa A Williams, Anne van Dongen, Amy Rozsa, Claire Lawrence, et al. 2022. "Warming Up Cool Cooperators." OSF. July 27. osf.io/ps5kb.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

We report results from six studies.

Study 1: We report a field-based experiment where we randomly allocate 5,821 1st time blood donors to one of four active message conditions based on warm-glow or impure altruism messages and a pure no-message control (n 5,801 took part after exclusions). We explored the moderation of messages by booking or return attendance. We collected quantitative data on the number who returned and actually attend to make a donation 3 months later.

Study 2: We confirmed the field-based experimental findings in a large-scale implementation trial where we compared the return rate to attend for a 2nd donation in the population of all new donors in Australia who had made their first donation in one of three one-year time windows prior to the warm-glow message being implemented (Ns = 90,317, 93,430, & 89,606), with the same return rates made by all new donors in Australia in one of two one-year time windows after the warm-glow message had been implemented (Ns = 81,766 & 88,551). We collected quantitative data of the number who returned and actually attend to make a donation 3 months later.

Study 3: A quantitative cross-sectional survey to evaluate whether booking was associated with self-reported warm-glow or pure-altruism. (n=716 first-time donors). We used self-reported survey items to assess anticipated warm-glow, negative affect, calmness and pure-altruism, booking status was recorded from donor records.

Study 4: Quantitative survey of n=1,124 donors. We examined 4 groups of donors by crossing donor status (first-time, novice) with donor type (plasma, blood) and explored the association between booking, self-reported warm-glow, and return behaviour. We collected quantitative data of the number who returned and actually attend to make a donation 3 months later, we assessed warm-glow by self-reports and booking status was recorded from donor records.

Study 5: Longitudinal survey at two-time points. We investigated the temporal stability of warm-glow among n=932 first-time donors by measuring warm-glow at two-time points (baseline; 12-week follow-up) and the effects of warm-glow and booking on return behaviour. We collected quantitative data of the number who returned and actually attend to make a donation 3 months later, we assessed booking status from donor records..

Study 6 was an online quantitative experiment (n = 1, 592) to validate the warm-glow and impure-altruism messages to be used in the field-based experiment. We collected quantitative data from self-reported data.

Research sample

Study 1: The field-based experiment was conducted on a sample of 5,821 1st time blood donors with 5,801 taking part after deferral criteria were applied. The initial sample of 5,821 (Mage = 31.98, [Sd = 11.96], % female = 60.7%) were new whole-blood donors from across Australia. After twenty donors were excluded (twelve due to email bounces and 8 due to being permanently deferred from donating blood) after their initial donation, the final sample was n=5,801 (Mage = 31.92, [Sd = 11.95], % female = 60.7%). This sample was chosen as our study focus was on increasing the conversion rate to make a second donation in first-time donors. Compared to the average age (32.5 years) and percentage of women (56%) for all 1st time whole blood donors in the three pre-implementation windows in study 2, this sample was generally representative.

Study 2: For the implementation trial we compared the return rate for a 2nd donation in the population of all new donors in Australia who had made their first donation in one of three one-year time windows prior to the warm-glow message being implemented (Ns = 90,317, 93,430, & 86,606), with the same return rates made by all new donors in Australia in one of two one-year time windows after the warm-glow message had been implemented (Ns = 81,766 & 88,551). The 3-year pre-implementation period covered 3 twelve-month time windows: beginning on the 16th of April to the 15th of April for years (1) 2015-2016 (n = 90,317; Mage = 32.0, [Sd = 13.6], % women = 55%, n = 49,778), (2) 2016-2017 (n = 93,430; Mage = 32.5, [Sd = 13.4], % women = 56%, n = 52,539), and (3) 2017-2018 (n = 86,606; Mage = 33.1, [Sd = 13.0], % women = 57%, n = 49,531), and the post-implementation period covered two slightly overlapping time windows: (1) 9th of July 2018 - 8th of July 2019 (n = 81,766; Mage = 34.9, [Sd = 13.5], % female = 57%, n = 46,689) and (2) 16th of April 2019 - 15th of April 2020 (n = 85,551; Mage = 34.3, [Sd = 13.2], % women = 57%, n = 50,480). This sample was chosen as our study focus was on increasing the conversion rate to make a second donation in first-time donors. This is the full population of all 1st-time donors

Study 3: Seven-hundred and sixteen first-time donors took part (Mage = 41.20, Sd = 14.85, % women = 65.6%, n women = 634). Sampling was targeted to ensure that the gender balance and age (18-69: donors over 70 were excluded) distribution was representative of the donor population. This sample was chosen as our study focused on exploring the links between self-reported warm-glow and booking status in first-time donors and this sample had items that matched the type of items that had been used previously to assess self-reported warm-glow and also sample first-time donors. In the nearest post-implementation time window (2019-2020) the average age was 34.3 with 57% women. Thus, the current sample was older and had a higher percentage of women.

Study 4: In total, 1,124 donors (Mage = 30.27, Sd = 11.51, % women = 55.2%, n = 621 women) participated. Of these, 401 were first-time whole-blood donors (Mage = 30.38, Sd = 11.879; % women = 54.6%, n = 219 women), 208 first-time plasma donors (Mage = 30.07, Sd = 11.50, % women 50.5%, n = 106 women), 379 were novice whole-blood donors (Mage = 29.61, Sd 11.49, % women = 60.4%, n = 229 women) and 136 novice plasma donors (Mage = 30.53, Sd = 10.75; % women = 49.3%, n = 67 women). This sample was chosen as our study focused on exploring the links between self-reported warm-glow, booking status and return behaviour and this study contained assessments of all of these. We also wished to explore the effects of first-time vs novice (more experienced donors) and type of donation (whole blood vs plasma) and this study contained data on all of these. The sample was representative in terms of the percentage of women but slightly younger than the average age (32.5 years) for 1st time whole blood donors in the three pre-implementation windows in study 2.

Study 5: Nine-hundred and thirty-two first-time donors who completed the survey provided warm-glow responses at waves 1 and 4 (Mage = 36.36, Sd = 14.02, % women = 70.3%, n = 655 women). This sample was chosen as our study focused on exploring the temporal stability of warm-glow in first-time donors, and in this study, warm-glow was assessed at two-time points. The study also had data on booking status and behaviour and this study also addressed the effects of booking on the stability of warm-glow and explored the prediction of behaviour from warm-glow and booking status. Compared to the average age (32.5 years) and percentage of women (56%) for 1st time whole blood donors in the three pre-implementation windows in study 2, this sample was slightly older and had a higher percentage of women.

Study 6: A total of 1,592 participants were recruited through Prolific (<https://www.prolific.co/>) (Mage = 36.47, Sd = 13.00, % women = 50.2%, n women = 795). This sample was chosen as our study was designed to explore the validity of the messages used in studies 1 and 2 in the general population. The sample is representative of 50.2% women compared to 51% of the population, but slightly younger at 36.47 years compared to the UK population average of 40.4 years.

Sampling strategy

Study 1: For the field-based experiment a sample of 5,821 new whole blood donors with A-, A+, O- and O+ blood types across Australia, who made their first whole blood donation six weeks previously, were recruited using the following criteria: (1) made their first whole blood donation six weeks ago, and (2) had not donated previously. For each week of the field-based experiment, all donors who met these eligibility criteria were selected and randomly allocated to conditions. The whole population meeting the sampling eligibility criteria were included. This process continued until the pre-determined numbers per condition were achieved. Sample size was determined as follows. As warm-glow has a small effect on predicted blood donor attendance, to achieve 80% power to detect a small effect with an alpha of .05 in a simple regression model with 8 predictors (e.g., message type and identity prime conditions, age, sex, blood group, booking status, and terms for the interaction of booking status with experimental conditions), 757 donors per arm are needed. Thus, we aimed for 1,000 new donors per arm to allow for ad-hoc exploratory analyses.

Study 2 For the implementation study all new donors in Australia across three pre-implementation time windows and 2 post-implementation time windows. We sampled the whole target population of first-time donors in Australia across three pre-implementation periods and two post-implementation periods.

Study 3: 20,000 donors were contacted with an even split across the 3 groups; (1) first-time donors (first donated after February 2020), (2) existing eligible donors who had donated after February 2020, and (3) existing eligible donors who had not donated since February 2020. Sampling was targeted to ensure that the gender balance and age (18-69: donor over 70 were excluded) distribution was representative of the donor population. Nine-hundred and sixty-six first-time donors were recruited. Of these 250 1st-time donors started the survey but did not complete the majority of measures including the measure of adjectives used to assess warm-glow. As 250 did not complete the survey (non-completers) but started it, and 716 did (completers), there is a potential for selection bias. Indeed, completer status is predicted by having had a COVID-19 test ($\chi^2(1) = 3.373, p = .066, \phi = .061$) with those who had a COVID-19 test more likely to be completers, and being older ($M = 41.20, Sd = 14.85$ vs $M = 34.14, Sd = 14.33; t(964) = 6.531, p = .000; DCohen = 0.421$). Completer status was not predicted by gender ($\chi^2(1) = 0.020, p = .887, \phi = -.005$). Rebooking status was not predicted by being a non-completer or completer ($\chi^2(1) = 1.151, p = .283, \phi = .035$) nor as a function of having had a COVID-19 test ($\chi^2(1) = 3.315, p = .069, \phi = .061$). To account for any potential selection bias we completed a sensitivity analysis using a Heckman Probit Selection model with COVID-19 test, gender and age as predictors of selection bias. There was no evidence of selection bias. This study constituted a secondary analysis of these data. There was no data on which to base a power calculation. However, if we assume a small effect size then we would need 363 donors who booked and 363 who did not ($n = 726$) with a power of .80 and $\alpha = .05$. In our sample 398 booked and 329 did not. As such, we had a sufficient sample size to detect a small effect.

Study 4: New (1st donation) and novice (1 or 2 previous donations and no plasma donations for whole blood donors, and 1 or 2 previous plasma donations and less than 3 whole blood donations for plasma donors) donors who were either whole blood or plasma donors were recruited by convenience sampling in donor centres, at three large urban Australian Red Cross Lifeblood (Lifeblood) centres (Brisbane Donor Centre, Sydney Town Hall and Melbourne Mt Waverley) from April 2016 to October 2017. A total sample of 1,153 donors were recruited. Of these, 441 were 1st time whole blood donors, 218 1st time plasma donors, 385 were novice whole blood donors and 136 novice plasma donors. Overall, there were 354 plasma donors and 799 whole blood donors. There were missing data on warm-glow ($n = 26$), age ($n = 1$) and gender ($n = 2$). Once these were removed the final sample size was 1,124. Of these, 401 were 1st time whole blood donors, 208 1st time plasma donors, 379 were novice whole blood donors and 136 novice plasma donors. Overall, there were 344 plasma donors and 780 whole blood donors. There was no data on which to base a power calculation. However, if we assume a small effect size then we would need 363 donors who booked and 363 who did not ($n = 726$) with a power of .80 and $\alpha = .05$. In our sample 781 booked and 329 did not. As such, we had a sufficient sample size to detect a small effect.

Study 5: All 25,428 first-time whole blood donors were invited to participate in the study between January 2017 and September 2017. 4,472 first-time donors were recruited of these data were missing from 500 participants on intrinsic regulation (warm-glow) at wave 1, leaving complete data on 3,972. Of the 4,472, 136 donors were actively deferred from donating for medical and life-style reasons. Those deferred were more likely to be women ($\chi^2(1) = 18.296, p = .000, \phi = .064$) but did not vary by age (M deferred = 33.88, $Sd = 12.83$ vs M not deferred = 35.20, $Sd = 13.57; t(4470) = 1.119, p = .263; DCohen = 0.033$). Those deferred were not significantly different from those not-deferred on warm-glow at wave 1 (M deferred = 15.06, $Sd = 4.34$ vs M not deferred = 15.18, Sd

= 4.09: $t(3970) = 0.315$, $p = .753$; $DCohen = 0.009$) or wave 4 (M deferred = 15.48, $Sd = 4.09$ vs M not deferred = 15.21, $Sd = 4.07$: $t(955) = -0.329$, $p = .743$; $DCohen = -0.021$). However, as deferral status influenced the decision to re-book ($\chi^2(1) = 5.571$, $p = .016$, $\phi = -.036$), with those deferred less likely to rebook, we excluded those who were actively deferred from the analyses. After removal of those actively deferred there were 4336 1st-time donors at wave 1. At wave 4 a total of 932 completed the survey providing data on warm-glow. Those who completed the surveys providing data on warm-glow at waves 1 and 4 were significantly older (M Wave 1 only = 34.88, $Sd = 13.39$ vs M Waves 1 & 2 = 36.27, $Sd = 14.09$: $t(4334) = -2.863$, $p = .004$; $DCohen = -0.087$) and less likely to be male ($\chi^2(1) = 8.231$, $p = .004$, $\phi = -.044$) than those who completed only wave 1, however, they did not significantly vary on warm-glow at wave 1 (M Time 1 only = 15.25, $Sd = 4.06$ vs M time1 & 2 = 14.97, $Sd = 4.18$: $t(3849) = 1.892$, $p = .059$; $DCohen = 0.061$). Furthermore, warm-glow at wave 4, for those who completed both surveys ($M = 15.21$, $Sd = 4.06$) was not significantly different from warm-glow at wave 1 ($M = 15.25$, $Sd = 4.06$) in those who just completed the survey at time 1 (t (one-sample (931)) = -0.320 , $p = .749$; $DCohen = -0.021$). As such, there is no evidence that initial levels of warm-glow influenced who remained in the study from waves 1 to waves 4. However, we controlled for age and gender in these analyses. There was no data on which to base a power calculation. However, if we assume a small effect size then we would need 363 donors who booked and 363 who did not ($n = 726$) with a power of .80 and $\alpha = .05$. In our sample 612 booked and 320 did not. As such, we had a sufficient sample size to detect a small effect

Study 6: Participants (aged 18 to 80 from the UK) were recruited through Prolific (<https://www.prolific.co/>) to an experiment hosted on Qualtrics (<https://www.qualtrics.com/uk/>). A total of 1,592 were recruited. There were 266 participants in the BAU/Control, 273 in the 'identity-only', 263 in the 'warm-glow-only', 256 in the 'warm-glow-plus-identity', 279 in the 'impure-altruism-only', and 255 in the 'impure-altruism-plus-identity' conditions. Ferguson, Farrell, and Lawrence 64 report a Cohen's d of 0.344 for positive ratings of warm-glow vs altruistic messages for blood donation. Achieving 80% power to detect a small effect with an alpha of .05 for a 2 (warm-glow: yes; no) \times 2 (identity: yes; no) design requires 265 per condition. For a series of pairwise comparisons of experimental groups, detecting a small effect with a Cohen's d of 0.344 requires 135 participants per condition.

Data collection

We collected data on verified donor attendance for studies 1, 2, 4 and 5. This was achieved in a blinded manner, with data analysts who extracted the attendance data blind to the experimental condition that the donor was assigned to. We collected verified data on booking status in studies 1, 2, 3, 4, and 5.

We collected survey data on experienced warm-glow in studies 3, 4, 5, and afforded warm-glow in 6, and pure-altruism in study 3.

In study 3 we assessed 'anticipated warm-glow' using adjectives of the type that have been used to assess warm-glow in previous research. Specifically, donors were asked: "What do you think donating blood in the future during the coronavirus pandemic would be like?" to which donors responded using 12 adjectives [Unrewarding, Pointless, Displeasing, Negative, Stressful, Unsatisfying, Rewarding, Worthwhile, Pleasing, Positive, Relaxing, Satisfying], each measured on a 4-point scale (1 = Not at all, 2 = Slightly, 3 = Moderately and 4 = Extremely). Of these adjectives, 'Rewarding', 'Worthwhile', 'Pleasing', 'Positive', and 'Satisfying' represent the type of adjectives that have been used to assess warm-glow in previous research [e.g., 24]. 'Pure Altruism' was assessed with two items concerning the donor's motivation for their last donation ("To help someone in need" and "It's important for society"), each measured on a 5-point scale (1 = Not important at all, 2 = Slightly Important, 3 = Moderately Important, 4 = Very Important, 5 = Extremely Important).

In studies 4 and 5 experienced warm-glow was assessed using the intrinsic regulation subscale the Blood Donor Identity Survey (BDIS) which comprises 3 items ("I enjoy donating blood", "For me, being a blood donor means more than just donating blood", and "Blood donation is an important part of who I am") with each item measured on a 7-point scale (1 = Not at all True, 7 = Very True).

We collected data, in study 6, on the perception of the messages to be used in the field-based experiment (study 1) with respect to afforded warm-glow, recipient vs donor focus and provision of the public good (blood supply). 'Afforded Donor vs Recipient Focus' was calculated as the average of 2 items ("To what extent do you think that the message focuses on the blood donor, the recipient of blood or both?" and "To what extent do you think that the message focuses on the emotions of the blood donor, the recipient of blood or both?") measured on a scale from -50 (primarily on the blood donor) to +50 (primarily on the recipient of blood) with a midpoint of 0 (equally on both the donor and recipient). 'Afforded Warm-Glow' was calculated as the sum of 2 items ("The message makes me feel that donating blood is personally rewarding in itself" and "The message makes me feel that donating blood would make me feel like a good person) measured on a 7-point scale (1 = "not at all" to 7 = "completely"). 'Afforded Focus on Maintaining the Blood Supply' was assessed with a single item ("The message makes me feel that donating blood would ensure that there is enough blood for all who need it.") measured on a 7-point scale (1 = "not at all" or 7 = "completely").

Others present when data was collected: For study 1 donors received a message there may have been other present when they read it but we do not have data on this. For study 2 this was aggregated population data on booking status and return behaviour other donors and staff may have been present when donors booked. Study 3 was an online survey (Qualtrics) and it is assumed that donors complete this independently. For study 4 data was collected using an electronic tablet in a donor centre with other blood donors and blood collection staff present. For study 5 data was collected using an online questionnaire (Qualtrics) and it is assumed that donors complete this independently. For study 6 data was collected using an online questionnaire (Qualtrics) and it is assumed that donors complete this independently.

Timing

Study 1: The field experiment ran from the 22nd of May 2018 to the 8th of July with the 3 month follow up on attendance completed by the end of Nov 2018.

Study 2: For the implementation study the following time windows were analysed. Pre-implementation period covered three time windows consisting of the 16th of April to the 15th of April for years (1) 2015-2016, (2) 2016-2017 and (3) 2017-2018 and the post-implementation period covered two slightly overlapping time windows (1) the 9th of July 2018 to the 8th of July 2019 and (2) the 16th of April 2019 to the 15th of April 2020.

Study 3: Donors were sent a de-identified online survey invitation between the 12th of August and the 24th of September 2020. A reminder was sent on the 27th of August 2020. Data were extracted on the 24th of September 2020

Study 4: The survey took place between April 2016 to October 2017.

Study 5: The survey took place between January 2017 and September 2017.

Study 6: The online survey took place on the 24th of March 2021, and all data were collected on that day.

Data exclusions

Study 1: For the field experiment 20 participants were excluded, of which 12 had bounced emails and 8 with permanent deferrals (evidence of blood-borne infection). These were not pre-determined and analyses showed that they were evenly distributed across the study conditions.

Study 2: There were no exclusions in the implementation trial.

Study 3: Nine-hundred and sixty-six first-time donors were recruited. Of these 250 1st-time donors started the survey but did not complete the majority of measures including the measure of adjectives used to assess warm-glow. As 250 did not complete the survey (non-completers) but started it, and 716 did (completers), there is a potential for selection bias. To account for any potential selection bias we completed a sensitivity analysis using a Heckman Probit Selection model with COVID-19 test, gender and age as predictors of selection bias. There was no evidence of selection bias.

Study 4: A total sample of 1,153 donors were recruited. There were missing data on warm-glow (n = 26), age (n = 1) and gender (n = 2). Once these were removed the final sample size was 1,124. Of these, 401 were 1st time whole blood donors, 208 1st time plasma donors, 379 were novice whole blood donors and 136 novice plasma donors. Overall, there were 344 plasma donors and 780 whole blood donors.

Study 5: Of the 4,472, 136 donors were actively deferred from donating for medical and life-style reasons. After removal of those actively deferred there were 4336 1st-time donors at wave 1. At wave 4 a total of 932 completed the survey providing data on warm-glow. There is no evidence that initial levels of warm-glow influenced who remained in the study from waves 1 to waves 4. However, we controlled for age and gender in these analyses.

Study 6: There were no exclusions.

Non-participation

Study 1: 20 donors were excluded (12 due to bounced emails, and 8 due to permanent deferrals) With respect to the initial randomization of 5,821 donors dropping to 5801 after the 20 exclusions. There was no evidence that exclusion systematically varied across arms or affect distributions of age and sex. Thus, there was no evidence for bias.

Study 2: There were no dropouts.

Study 3: Nine-hundred and sixty-six first-time donors were recruited. Of these 250 first-time donors started the survey but did not complete the majority of measures including the measure of adjectives used to assess warm-glow. As 250 did not complete the survey (non-completers) but started it, and 716 did (completers), there is a potential for selection bias. To account for any potential selection bias we completed a sensitivity analysis using a Heckman Probit Selection model and there was no evidence for bias.

Study 4: A total sample of 1,153 donors were recruited. Of these, 441 were first-time whole blood donors, 218 first-time plasma donors, 385 were novice whole blood donors and 136 novice plasma donors. There were missing data on warm-glow (n = 26), age (n = 1) and gender (n = 2). Once these were removed the final sample size was 1,124.

Study 5: Of the 4,472 recruited donors, 136 donors were actively deferred from donating for medical and life-style reasons. After removal of those actively deferred there were 4,336 first-time donors at wave 1. At wave 4 a total of 932 completed the survey providing data on warm-glow. The other data were excluded due to survey non-completion.

Study 6: There were 266 participants in the BAU/Control, 273 in the 'identity-only', 263 in the 'warm-glow-only', 256 in the 'warm-glow-plus-identity', 279 in the 'impure-altruism-only', and 255 in the 'impure-altruism-plus-identity' conditions. There were no dropouts.

Randomization

In study 6 participants were block-randomized by sex to each of the 6 conditions to ensure equal numbers of males and females. For the field-based experiment (study 1), this was simple randomization. There was no randomization for the implementation study where we compared pre-and post-implantation time periods or for studies 3, 4 and 5 which were observational. Study 2 (implementation study) was based on aggregate data and we could not control for age and sex. However, these were consistent across the time window and by booking status so unlikely to have influence the pattern of results observed.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

Study 1: For the field-based experiment we have the following characteristics. In the control condition we had an n of 1162, with a mean age 32.31 (SD = 12.06) with a %female of 60.5. In the warm-glow-identity condition we had an n of 1164 with a mean age of 31.72 (SD = 11.78) and a %female of 60.8. In the warm-glow condition we had an n of 1164 with a mean age of 32.82 (SD = 12.03) and a %female = 60.4. In the positive affect identity condition we had an n of 1164 with a mean age of 31.92 (SD = 12.12) with a %female of 62.7. In the positive affect condition we had an n of 1167 with a mean age of 31.66 (SD = 11.81) with a %female of 59.0. The majority of the sample had blood type O+ (45.9%), followed by A+ (36.9%), and O- (10.7%), A- (6.4%) and O (1%).

Study 2: For the implementation study the % female at each time window was as follows: Pre-implementation time window one (2015-2016: n = 90,317) the mean age was 32.0 (Sd = 13.6) and % female was 55%, time window two (2016-2017: n = 93,430) the mean age was 35.5 (Sd = 13.4) and the %female was 56%, and time window 3 (2017-2018: n = 86,606) the mean age was 33.1 (Sd = 13.0) and the %female was 57%. In the post-implementation period for time windows one (9th of July 2018 to the 8th of July 2019: n = 81,766) the mean age was 34.9 (Sd = 1354) and the %female was 56%, and for time window two (6th of April 2019 to the 15th of April 2020: n = 88,551) the mean age was 34.3 (Sd = 13.2) and the %female was 57%. For the aggregate pre-implementation mean age Ws 35.5 (Sd = 13.3: n = 270,353) and the post-implementation period the mean age was 34.6 (Sd = 13.3: n = 170, 317). As these are aggregate data to examine the effects of age and sex on the aggregate results we re-ran the aggregated analysis on percentage return rates adjusted for aggregate age and sex (percentage women). We obtained estimates of age- and sex-adjusted percentage return rates from a simple general linear model (n = 10). This simple adjustment for age and sex does not alter the pattern of results in terms of the key interaction and the effect of the warm-glow-plus-identity message on return rates for those who do not book.

Study 3: Nine-hundred and sixty-six first-time donors were recruited (mean age = 39.37 years, Sd = 15.03 years; 66% female with 23.7% having had a COVID-19 test). Of these 250 1st-time donors started the survey but did not complete the majority of measures including the measure of adjectives used to assess warm-glow. As 250 did not complete the survey (non-completers) but started it, and 716 did (completers), there is a potential for selection bias. Indeed, completer status is predicted by having had a COVID-19 test ($\chi^2(1) = 3.373, p = .066, \phi = .061$) with those who had a COVID-19 test more likely to be completers, and being older (M = 41.20, Sd = 14.85 vs M = 34.14, Sd = 14.33; $t(964) = 6.531, p = .000, DCohen = 0.421$). Completer status was not predicted by gender ($\chi^2(1) = 0.020, p = .887, \phi = -.005$). Rebooking status was not predicted by being a non-completer or completer ($\chi^2(1) = 1.151, p = .283, \phi = .035$) nor as a function of having had a COVID-19 test ($\chi^2(1) = 3.315, p = .069, \phi = .061$). To account for any potential selection bias we completed a sensitivity analysis using a Heckman Probit Selection model with COVID-19 test, gender and age as predictors of selection bias.

Study 4: A total sample of 1,153 donors were recruited. Of these, 441 were 1st time whole blood donors (M age = 30.47, Sd = 11.86, n = 413; % female = 55.6%, n = 412), 218 1st time plasma donors (M age = 30.95, Sd = 11.30, n = 218; % female 50.5%, n = 218), 385 were novice whole blood donors (M age = 29.52, Sd 11.47, n = 385; % female = 60.8%, n = 385) and 136 novice plasma donors (M age = 30.5, Sd = 10.75 n = 136; % female = 49.3%, n = 136;). Age did not vary significant across these groups (F (3, 1148) , = 0.860, p = .461, $\eta^2 = .002$) but gender did ($\chi^2(3) = 8.73, p = .033, \phi = .087$). Overall, there were 354 plasma donors (M age = 30.79, Sd = 11.19, n = 345, % female = 50.0%, n = 354) and 799 whole blood donors (M age = 30.01, Sd = 11.64, n = 798; % female = 57.4%, n = 797) who did not differ by age (t (1150) = 1.06, p = .289; DCohen = 0.063) but did by gender ($\chi^2(1) = 6.50, p = .011, \phi = -.075$). There were 632 1st-time donors (M age = 30.64, Sd = 11.67, n = 631; % female = 53.8%, n = 630) and 521 novice donors (M age = 29.79, Sd = 11.26, n = 521; % female = 57.5%, n = 521) who did not differ by age (t (1150) = 1.25, p = .212; DCohen = 0.074) or gender ($\chi^2(1) = 1.81, p = .178, \phi = -.040$). There were missing data on warm-glow (n = 26), age (n = 1) and gender (n = 2). Once these were removed the final sample size was 1,124. Of these, 401 were 1st time whole blood donors (M age = 30.38, Sd = 11.879; % female = 54.6%, n = 219 women), 208 1st time plasma donors (M age = 30.07, Sd = 11.50, % female 50.5%, n = 106 women), 379 were novice whole blood donors (M age = 29.61, Sd 11.49, % female = 60.4%, n = 229 women) and 136 novice plasma donors (M age = 30.53, Sd = 10.75; % female = 49.3%, n = 67 women). Neither age (F (3, 1120) , = 0.787, p = .501, $\eta^2 = .002$) nor gender = ($\chi^2(3) = 7.68, p = .053, \phi = .083$) significant varied across these groups. Overall, there were 344 plasma donors (M age = 30.85, Sd = 11.19, % female = 50.3%, n = 173 women) and 780 whole blood donors (M age = 30.01, Sd = 11.65, % female = 57.4%, n = 797) who did not differ by age (t (1122) = 1.14, p = .254; DCohen = 0.068) but did by gender ($\chi^2(1) = 4.93, p = .026, \phi = -.066$). There were 609 1st-time donors (M age = 30.62, Sd = 11.69; % female = 53.4%, n = 325 women) and 515 novice donors (M age = 29.85, Sd = 11.30; % female = 57.5%, n = 296 women) who did not differ by age (t (1122) = 1.11, p = .267; DCohen = 0.066) or gender ($\chi^2(1) = 1.81, p = .178, \phi = -.040$).

Study 5: 4,472 first-time donors were recruited (mean age = 35.16 years, SD = 13.55 years; 68% female), of these data were missing from 500 participants on intrinsic regulation (warm-glow) at wave 1, leaving complete data on 3,972. We did not

impute these missing data as data were missing on all items pertaining to warm-glow and the missing percentage was high (11%). Of the 4,472, 136 donors were actively deferred from donating for medical and life-style reasons. Those deferred were more likely to be women ($\chi^2(1) = 18.296, p = .000, \phi = .064$) but did not vary by age (M deferred = 33.88, Sd = 12.83 vs M not deferred = 35.20, Sd = 13.57: $t(4470) = 1.119, p = .263, DCohen = 0.033$). Those deferred were not significantly different from those not-deferred on warm-glow at wave 1 (M deferred = 15.06, Sd = 4.34 vs M not deferred = 15.18, Sd = 4.09: $t(3970) = 0.315, p = .753, DCohen = 0.009$) or wave 4 (M deferred = 15.48, Sd = 4.09 vs M not deferred = 15.21, Sd = 4.07: $t(955) = -0.329, p = .743, DCohen = -0.021$). However, as deferral status influenced the decision to re-book ($\chi^2(1) = 5.571, p = .016, \phi = -.036$), with those deferred less likely to rebook, we excluded those who were actively deferred from the analyses. After removal of those actively deferred there were 4336 1st-time donors (M age = 35.20 years, Sd = 15.57 years; % female = 67%) at wave 1. At wave 4 a total of 932 completed the survey providing data on warm-glow. Those who completed the surveys providing data on warm-glow at waves 1 and 4 were significantly older (M Wave 1 only = 34.88, Sd = 13.39 vs M Waves 1 & 2 = 36.27, Sd = 14.09: $t(4334) = -2.863, p = .004, DCohen = -0.087$) and less likely to be male ($\chi^2(1) = 8.231, p = .004, \phi = -.044$) than those who completed only wave 1, however, they did not significantly vary on warm-glow at wave 1 (M Time 1 only = 15.25, Sd = 4.06 vs M time1 & 2 = 14.97, Sd = 4.18: $t(3849) = 1.892, p = .059, DCohen = 0.061$). Furthermore, warm-glow at wave 4, for those who completed both surveys (M = 15.21, Sd = 4.06) was not significantly different from warm-glow at wave 1 (M = 15.25, Sd = 4.06) in those who just completed the survey at time 1 ($t(931) = -0.320, p = .749, DCohen = -0.021$). As such, there is no evidence that initial levels of warm-glow influenced who remained in the study from waves 1 to waves 4. However, we controlled for age and gender in these analyses.

Study 6: A total of 1,592 were recruited (mean age = 36.47, Sd = 13.00 female % = 52). 34% (n = 538) indicated that they had ever donated blood and 12.1% (n = 194) were current donors (donated within the last 2 years). There were 266 participants in the BAU/Control, 273 in the 'identity-only', 263 in the 'warm-glow-only', 256 in the 'warm-glow-plus-identity', 279 in the 'impure-altruism-only', and 255 in the 'impure-altruism-plus-identity' conditions. Neither age (F(5, 1583), = 1.770, p = .116, $\eta^2 = .006$), gender ($\chi^2(5) = 0.863, p = .973, \phi = .023$), ever donated blood ($\chi^2(5) = 8.918, p = .112, \phi = .075$) nor being a current blood donor ($\chi^2(5) = 1.221, p = .943, \phi = .028$) significantly varied by condition.

Recruitment

Study 1: For the field-based experiment the whole population of all 1st time donors who met the inclusion criteria were sampled for each week of the field-based experiment and randomly allocated to conditions. The whole population meeting the sampling eligibility criteria were included. This process continued until the pre-determined numbers per condition were achieved. As such, there is little scope for bias donor were sample and allocated to receive a message based on the randomization schedule. A sample of 5,821 (Mage = 31.98, [Sd = 11.96], % female = 60.7%) new whole-blood donors from across Australia with A-, A+, O- and O+ blood types (this was based on the Australian Red Cross Lifeblood donor segmentation policy at that time. However, A and O positive and negative groups are the most prevalent, accounting for 86% of Australian blood donors) who had made their first whole-blood donation six weeks prior and had not donated previously in Australia were recruited. Twenty donors were excluded: twelve due to email bounces and 8 due to being permanently deferred from donating blood after their initial donation, leaving a final sample of n=5,801 (Mage = 31.92, [Sd = 11.95], % female = 60.7%). With respect to the initial randomization of 5,821 donors, age did not vary significantly across treatments (F(4, 5816) = 0.725, p = 0.575: DCohen = 0.059), nor did sex ($\chi^2(4) = 3.542, p = 0.472, DCohen = 0.050$). The majority of the sample had blood type O+ (45.9%), followed by A+ (36.9%), and O- (10.7%), A- (6.4%) and O (1%). The distribution of blood type did not vary by treatment ($\chi^2(16) = 18.16, p = 0.314, DCohen = 0.112$). The percentage who booked immediately post-donation was 37.7%, with 62.3% not booking while in the donor centre, and this did not vary by condition ($\chi^2(4) = 4.163, p = 0.384, DCohen = 0.054$). After the 20 donors were excluded (Supplementary Figure 8), donor age remained nonsignificant by condition (F(4, 796) = 0.796, p = 0.527: DCohen = 0.059), as did sex ($\chi^2(4) = 3.536, p = 0.472, DCohen = 0.054$). The majority of the sample had blood group O+ (46%), followed by A+ (36.9%), and O- (10.7%), A- (6.4%). The distribution of blood type did not vary by condition ($\chi^2(12) = 14.62, p = 0.263, DCohen = 0.100$). The percentage who booked was 37.7% with 62.3% not initially booking and this did not vary by condition ($\chi^2(4) = 4.198, p = 0.380, DCohen = 0.054$). Thus, the randomization by arm remained intact after the exclusions, as such there is no evidence of bias due to exclusions.

Study 2: For the implementation study all new donors in Australia across three pre implementation time windows and 2 post implementation time windows were selected. As the whole populations are sampled for the field-based experiment and the implementation study we feel there is little to no bias in the results especially as the findings from the field-based experiment are confirmed in the implementation study.

Study 3: 20,000 Australian donors were contacted with an even split across the 3 groups; (1) first-time donors (first donated after February 2020), (2) existing eligible donors who had donated after February 2020, and (3) existing eligible donors who had not donated since February 2020. Sampling was targeted to ensure that the gender balance and age (18-69: donor over 70 were excluded) distribution was representative of the donor population. A 10-15% response rate was expected and as such some self-selection bias is possible, however, the attained sample was representative. Donors were sent a de-identified online survey invitation between the 12th of August and the 24th of September 2020. A reminder was sent on the 27th of August 2020. Data were extracted on the 24th of September 2020. The survey was hosted on Qualtrics (<https://www.qualtrics.com/>). Nine-hundred and sixty-six first-time donors were recruited (mean age = 39.37 years, Sd = 15.03 years; 66% female with 23.7% having had a COVID-19 test). The low number of COVID-19 tests reflect the low rate of COVID-19 infection in Australia at the time. Of these 250 first-time donors started the survey but did not complete the majority of measures including the measure of adjectives used to assess warm-glow. As 250 did not complete the survey (non-completers) but started it, and 716 did (completers), there is a potential for selection bias. Indeed, completer status is predicted by having had a COVID-19 test ($\chi^2(1) = 3.373, p = .066, \phi = .061$) with those who had a COVID-19 test more likely to be completers, and being older (M = 41.20, Sd = 14.85 vs M = 34.14, Sd = 14.33; $t(964) = 6.531, p = .000, DCohen = 0.421$). Completer status was not predicted by gender ($\chi^2(1) = 0.020, p = .887, \phi = -.005$). Booking status was not predicted by being a non-completer or completer ($\chi^2(1) = 1.151, p = .283, \phi = .035$) nor as a function of having had a COVID-19 test ($\chi^2(1) =$

3.315, $p = .069$, $\phi = .061$). To account for any potential selection bias we completed a sensitivity analysis using a Heckman Probit Selection model with COVID-19 test, gender and age as predictor of selection bias.

Study 4: New (1st donation) and novice (1 or 2 previous donations and no plasma donations for whole blood donors, and 1 or 2 previous plasma donations and less than 3 whole blood donations for plasma donors) donors who were either whole blood or plasma donors were recruited by convenience sampling in donor centres, at three large urban Australian Red Cross Lifeblood (Lifeblood) centres (Brisbane Donor Centre, Sydney Town Hall and Melbourne Mt Waverley) from April 2016 to October 2017. There is room for self-selection bias due to the convenience sampling procedure. In total, 1,124 donors (M age = 30.27, Sd = 11.51, % women = 55.2%, $n = 621$ women) participated. Of these, 401 were first-time whole-blood donors (M age = 30.38, Sd = 11.879; % women = 54.6%, $n = 219$ women), 208 first-time plasma donors (M age = 30.07, Sd = 11.50, % women 50.5%, $n = 106$ women), 379 were novice whole-blood donors (M age = 29.61, Sd 11.49, % women = 60.4%, $n = 229$ women) and 136 novice plasma donors (M age = 30.53, Sd = 10.75; % women = 49.3%, $n = 67$ women). In total, 1,124 donors (M age = 30.27, Sd = 11.51, % women = 55.2%, $n = 621$ women) participated. Of these, 401 were first-time whole-blood donors (M age = 30.38, Sd = 11.879; % women = 54.6%, $n = 219$ women), 208 first-time plasma donors (M age = 30.07, Sd = 11.50, % women 50.5%, $n = 106$ women), 379 were novice whole-blood donors (M age = 29.61, Sd 11.49, % women = 60.4%, $n = 229$ women) and 136 novice plasma donors (M age = 30.53, Sd = 10.75; % women = 49.3%, $n = 67$ women). Thus there was some variability by age and sex, also the sampling design involved clustering. To account for this we controlled for age and sex and ran the statistical model with both standard error and error clustered within sample region. The results show that the results did not differ for the models with standard errors or clustered errors.

Study 5: All 25,428 first-time whole blood donors were invited to participate in the study between January 2017 and September 2017. All eligible donors were sent an email invitation asking them to participate in the study. A reminder email was sent after 5 days. Nine-hundred and thirty-two first-time donors who completed the survey provided warm-glow responses at waves 1 and 4 (M age = 36.36, Sd = 14.02, % women = 70.3%, $n = 655$ women). All 25,428 first-time whole blood donors were invited to participate in the study between January 2017 and September 2017. 4,472 first-time donors were recruited (M age = 35.16, Sd = 13.55; % female = 68%). Thus there is a potential selection bias here. However, the sample was representative. Of these, data were missing from 500 participants on intrinsic regulation (warm-glow) at wave 1, leaving complete data on 3,972 participants. We did not impute these missing data as data were missing on all items pertaining to warm-glow and the missing percentage was high (11%). Of the 4,472 recruited donors, 136 donors were actively deferred from donating for medical and life-style reasons. Those deferred were more likely to be women ($\chi^2(1) = 18.296$, $p = .000$, $\phi = .064$) but did not vary by age (M deferred = 33.88, Sd = 12.83 vs M not deferred = 35.20, Sd = 13.57: $t(4470) = 1.119$, $p = .263$; DCohen = 0.033). Those deferred were not significantly different from those not-deferred on warm-glow at wave 1 (M deferred = 15.06, Sd = 4.34 vs M not deferred = 15.18, Sd = 4.09: $t(3970) = 0.315$, $p = .753$; DCohen = 0.009) or wave 4 (M deferred = 15.48, Sd = 4.09 vs M not deferred = 15.21, Sd = 4.07: $t(955) = -0.329$, $p = .743$; DCohen = -0.021). However, as deferral status influenced the decision to book ($\chi^2(1) = 5.571$, $p = .016$, $\phi = -.036$), with those deferred less likely to book, we excluded those who were actively deferred from the analyses. After removal of those actively deferred there were 4,336 first-time donors (M age = 35.20 years, Sd = 15.57 years; % female = 67%) at wave 1. At wave 4 a total of 932 completed the survey providing data on warm-glow. Those who completed the surveys providing data on warm-glow at waves 1 and 4 were significantly older (M Wave 1 only = 34.88, Sd = 13.39 vs M Waves 1 & 2 = 36.27, Sd = 14.09: $t(4334) = -2.863$, $p = .004$; DCohen = -0.087) and less likely to be male ($\chi^2(1) = 8.231$, $p = .004$, $\phi = -.044$) than those who completed only wave 1, however, they did not significantly vary on warm-glow at wave 1 (M Time 1 only = 15.25, Sd = 4.06 vs M time1 & 2 = 14.97, Sd = 4.18: $t(3849) = 1.892$, $p = .059$; DCohen = 0.061). Furthermore, warm-glow at wave 4, for those who completed both surveys (M = 15.21, Sd = 4.06) was not significantly different from warm-glow at wave 1 (M = 15.25, Sd = 4.06) in those who just completed the survey at time 1 (t (one-sample (931)) = -0.320, $p = .749$; DCohen = -0.021). As such, there is no evidence that initial levels of warm-glow influenced who remained in the study from waves 1 to waves 4. However, we controlled for age and gender in these analyses as these demonstrated systematic variation and potential for bias.

Study 6: Participants who lived in England and were aged between 18 and 80 and registered on Prolific we recruited until the target sample size was achieved. A total of 1,592 participants were recruited through Prolific (<https://www.prolific.co/>) (M age = 36.47, Sd = 13.00, % women = 50.2%, n women = 795) with 34% ($n = 538$) indicating that they had ever donated blood and 12.1% ($n = 194$) being current donors (donated within the last 2 years). There were 266 participants in the BAU/Control, 273 in the 'identity-prime-only', 263 in the 'warm-glow-only', 256 in the 'warm-glow-plus-identity', 279 in the 'impure-altruism-only', and 255 in the 'impure-altruism-plus-identity' conditions. Recruitment was through Prolific (<https://www.prolific.co/>) to an experiment hosted on Qualtrics (<https://www.qualtrics.com/uk/>). All data were collected on the 24th of March 2021. Participants were block-randomized by sex to each of the 6 conditions. A total of 1,592 were recruited (M age = 36.47, Sd = 13.00, % female = 52%). 34% ($n = 538$) indicated that they had ever donated blood and 12.1% ($n = 194$) were current donors (donated within the last 2 years). There were 266 participants in the BAU/Control, 273 in the 'identity-only', 263 in the 'warm-glow-only', 256 in the 'warm-glow-plus-identity', 279 in the 'impure-altruism-only', and 255 in the 'impure-altruism-plus-identity' conditions. Neither age ($F(5, 1583) = 1.770$, $p = .116$, $\eta^2 = .006$), gender ($\chi^2(5) = 0.863$, $p = .973$, $\phi = .023$), ever donated blood ($\chi^2(5) = 8.918$, $p = .112$, $\phi = .075$) nor being a current blood donor ($\chi^2(5) = 1.221$, $p = .943$, $\phi = .028$) significantly varied by condition. Thus there was no evidence of bias across conditions. The sample was also representative with 50.2% women compared to 51% of the population, but slightly younger at 36.47 years compared to the UK population average of 40.4 years. Age and sex were controlled in our mediation models.

Ethics oversight

Study 1: The field-based experiment was approved by The Australian Red Cross Lifeblood Ethics Committee on 7th May 2018 with the first randomized message sent out on the 22nd of May for those who had made a donation on the 16th of April (Reference: Davison 04052018).

Study 2: The implementation study is based on aggregate data at the population level and as such no individual are identifiable and donors consent for their data to be used for research when they sign up as a blood donor.

Study 3: The study was approved by the Australian Red Cross Lifeblood Human Research Ethics Committee on 31st July 2020

(Ethical approval number Guerin31072020).

Study 4: The study protocol was approved by Lifeblood's Human Research Ethics Committee on 18th September 2015 with reference number 2015#07.

Study 5: he study protocol was approved by Lifeblood's Human Research Ethics Committee on 15th November 2016 (Reference: 2016#24).

Study 6: The experiment was approved by the University of Nottingham, School of Psychology Ethics committee (ref: S1309).

Note that full information on the approval of the study protocol must also be provided in the manuscript.