

The Trolley Problem & Autonomous Vehicles

SOME HALACHIC CONSIDERATIONS

In 1967, English philosopher Philippa Foot, in an essay on ethical dilemmas, raised a challenging question. What should the driver of a runaway tram do if the tram can be steered only onto one of two narrow tracks—one with five people working on it and the other with just one person—knowing that anyone on the track selected will certainly be killed? In 1976, American philosopher Judith Jarvis Thomson wrote the first of several articles on the subject, and coined the term “The Trolley Problem” to describe that question. Over the years, she and others have published essays on similar “ethical thought experiments,” each with variations on the original question, but all sharing a common denominator: a harmful outcome is inevitable.

For example, the initial case poses the question as to whether it is preferable to cause the death of one person rather than of multiple people when all else

is equal. A more difficult question, however, is whether it is permissible to actively cause the death of one person who was not previously in danger in order to spare the lives of multiple people who were already at risk. Another question is whether there can or should be a different approach when one’s own life is at stake. Yet another is whether it should matter if those in danger intentionally or knowingly put themselves in that situation.

These once hypothetical issues may now become practical with the advent of technology based on artificial intelligence that enables autonomous, or self-driving, cars. These vehicles operate with little or no human input and are equipped with advanced safety features, such as sophisticated sensors, cameras and radar and lidar systems that provide a full view of their surroundings, designed to minimize human error and enhance road safety. With an estimated 1.25 million deaths



and 20 million injuries worldwide each year from traffic accidents—most caused by human error—automated driving systems offer a solution. By removing the flawed human driver, these systems provide real-time monitoring, obstacle identification, and automatic decision-making, all designed to maximize safety for the vehicle and its occupants.

One challenge for those of us who are committed to halachah is that these cars

will not be pre-programmed by today's automobile manufacturers on the basis of halachic concerns. May an observant software developer produce a vehicular computer system that does not adhere to halachah? Perhaps more broadly relevant, may one drive a vehicle that has been designed to automatically respond in a manner that might contradict the halachah? For example, a sudden and unexpected situation may arise on the road, such as children running into the street after a ball. The car, programmed to prioritize passenger safety, may react in a way that protects its occupants but does not account for potential harm to others, including the children. Should it? Does it matter how many others might be harmed? And even if advanced calculations to prioritize the lives of others, especially multiple others, could be made, is the driver required, or even allowed, to risk or endanger his or her own life regardless of the circumstances?

These are some of the now practical dilemmas, akin to “The Trolley Problem,” presently emerging in relation to autonomous vehicles. In recent years, articles in several Torah publications have addressed this topic, and it is worthwhile to explore how these issues should be approached through the lens of halachah.

The Primacy of Life

It goes with saying that *retzichah*, murder—prohibited, among other places, in the *Aseres HaDibros* (*Shemos* 20:12)—is a most abhorrent crime, one which the Rambam (*Hilchos Rotzeiach* 4:9) identifies as contributing to the very destruction of society. Indeed, the Gemara in *Sanhedrin* (74a) identifies it as one of the three “cardinal” sins that one may never actively violate, even if it means losing his own life as a result; the

Rambam (*Hilchos Yesodei HaTorah* 5:1-2) and the *Shulchan Aruch* (*Yoreh Deah* 157:1) rule accordingly.

At the same time, the halachah places a premium on the importance of saving the life of someone who is in danger. The Torah (*Vayikra* 19:16) thus prohibits one to stand idly by when in position to save somebody whose life is in danger. An earlier Gemara in *Sanhedrin* (73a) states that if one sees another drowning in a river, being dragged away by a wild animal, or being attacked by bandits, he is obligated to rescue him if he can, even if it requires spending money to hire people who can provide assistance. Rashi there (*d”h ka mashma lan*) adds that one must explore all available options to save a life. The Rambam (*Hilchos Rotzeiach* 1:14) and the *Shulchan Aruch* (*Choshen Mishpat* 426:1) again rule accordingly.

When these two considerations—the prohibition against taking a life and the obligation to save a life—conflict, such as when someone is pursued by a *rodeif* (a pursuer clearly intent on killing him) and he has no other means of defense, a bystander may—and must—save the intended victim, even if it means taking the life of the *rodeif*. The Gemara there states this explicitly, but see Tosafos *d”h afrotzeiach* for a distinction, based on the Gemara on 72b, regarding the obligation depending on the certainty of the *rodeif*'s intentions. It should be noted, however, that in such a case, one must first try to eliminate the threat of the *rodeif* by wounding or harming him physically, and hence neutralizing him, without necessarily killing him; the Rambam (*ibid.* 1:6,7,13) and the *Shulchan Aruch* (*ibid.* 425:1) spell out some of the additional details.

As described above, autonomous cars will be programmed to prevent many accidents that could result in

fatalities and protect people who might otherwise be in danger—certainly an important advantage in line with the two halachic values mentioned earlier.

Regarding the concern that an autonomous car might harm or even kill someone who unexpectedly enters the street—whether a pedestrian or another driver—in order to best protect its passengers (given that even a thoroughly programmed car will be unable to anticipate every possible human action), it is worth considering that such a person could potentially be classified as a *rodeif*. Since his sudden moves endanger the lives of those in the vehicle and alternatives may not be available, his life may, in such a case, be set aside.

Although this person might have no intention of posing any threat, he is nonetheless considered a *rodeif*. This is evident from the Mishnah in *Ohalos* (7:6), which labels an unborn fetus who endangers the life of his mother during a difficult childbirth as a *rodeif*, and as codified by the Rambam (*ibid.* 1:9) and the *Shulchan Aruch* (*ibid.* 425:2). It would thus seem to be permissible for one to drive (or to program) such a vehicle even if its reaction to a particular situation may cause somebody's death, though every effort must obviously be made to avoid that result. It must also be stressed that this same Mishnah teaches that ordinarily it is forbidden to sacrifice one life in order to save another life (“*ein dochin nefesh mipnei nefesh*”), as the Rambam and *Shulchan Aruch* rule there as well; the *rodeif* case is the important exception.

Giving Precedence to One's Own Life

There are, moreover, other exceptions to both the prohibition against taking another's life and the requirement to do

whatever possible to save another's life. One exception to the former is self-defense, a *rodeif* case where it is one's own life that is at stake. The Gemara in *Berachos* (58a) states that if a person approaches someone with the intent to kill him—even if this intention is not explicitly declared (see Rashi *d"haTorah amrah*)—the prospective victim should kill the attacker first.

Some authorities hold that in such a case, the person defending himself is not obligated to first attempt to incapacitate the attacker by other means (see, for example, the *Mishneh LaMelech* on the Rambam, *Hilchos Chovel U'Mazik* 8:10, apparently based on *Shu"t HaRivash* No. 238. See, however, Rashi to *Sanhedrin* 57a, *d"ha veyachol*, and Rosh to *Bava Kamma* 3:13 who appear to reject this distinction; see also, at length, R. Velvel Soloveichik, *Chiddushei Maran Riz HaLeivi 'al HaRambam, Hilchos Rotzeiach* 1:13, and R. Avraham Yitzchak Kook, *Shu"t Mishpat Kohen* Nos. 133 and 139). Again, if another driver (or even a pedestrian) is behaving in a reckless manner which threatens the life of the driver of the autonomous vehicle, it would seem to be acceptable for him to “allow” the vehicle to do whatever is necessary to protect his own life.

Another example of the requirement to prioritize one's own life is found in the Gemara's discussion in *Bava Metziah* (62a) regarding two people wandering in a deserted area (far from civilization), where only one has a container of water with enough for just one person to survive. If they share it, both will die. The Gemara concludes based on a *passuk* (*Vayikra* 25:36) that the person with the water may drink it all because in this case saving one's own life takes precedence over saving another's. In other words, while it is tragic that

another person will certainly die as a result of one prioritizing his own life, this unfortunate outcome is deemed inconsequential in such a circumstance. It is noteworthy that neither the Rambam nor the *Shulchan Aruch* cite this particular conclusion, but the *Tur* does (*Yoreh Deah* 251), as does the Vilna Gaon there (No. 6), among others. It is also noteworthy that some authorities suggest that while one is not *required* to sacrifice his own life to save somebody else, one may opt to do so if he wishes (see R. Kook in *Shu"t Mishpat Kohen* No. 143 and R. Shaul Yisraeli, *Amud HaYemini* 16:5:27). According to this, a driver who is in danger may certainly choose to save his or her own life even at the expense of someone else's, though it may not be obligatory to do so.

Given that not all automobile accidents are fatal, a related question is to what extent one may or must risk his or her own life to save someone else from a possibly more certain danger.

In other words, is one obligated, or even allowed, to put oneself in a potentially life-threatening circumstance to save the life of another who is definitely in danger? R. Yosef Caro, both in his *Kesef Mishneh* on the Rambam (*Hilchos Rotzeiach* 1:14) and in his *Beis Yosef* on the *Tur* (*Choshen Mishpat* 426), refers to a passage in the *Talmud Yerushalmi* (the source of which he does not identify, but the *Netziv*, in his *Haa'mek She'eilah* on the *She'iltos* of R. Achai Gaon, *She'ilta* 129 No. 4, points to the Yerushalmi in *Terumos* 8:4) which indicates that one must indeed enter into a potentially dangerous situation to save someone who is certainly in danger.

Interestingly, R. Caro does not cite this ruling in the *Shulchan Aruch* there in *Choshen Mishpat*; the *S'ma* (No. 2)

suggests that this is because most major *poskim* omit it, and the *Pischei Teshuvah* (No. 2) quotes that while the *Talmud Yerushalmi* rules this way, the *Talmud Bavli*, which the halachah generally follows, does not. As for where the *Bavli* disagrees, some, including R. Ovadyah Yosef (*Shu"t Yechaveh Da'as* 3:84), point to a Gemara in *Niddah* (61a; see *Tosafos* there *d"ha atmerinchu*) where a Tanna refused to hide people who were wanted by the authorities for murder because by doing so he would be endangering his own life. The *Netziv*, in the piece cited above, also quotes this source but rejects it and instead refers to the aforementioned Gemara in *Sanhedrin* (73a), which teaches that one must do whatever is possible to save somebody's life, but does not say anything about even potentially putting one's own life at risk to do so.

Elsewhere in his *Ha'amek She'eilah* (*She'ilta* 147 No. 4), he offers a creative explanation of the Gemara's ruling in *Bava Metziah*, mentioned earlier, which permits the person with the water in the deserted area to drink it all, even though the other person there will die of dehydration. He argues that this ruling demonstrates that according to the *Bavli*, one is not required to endanger one's own life to save another person from certain danger.

In any event, the consensus is that one is not in fact obligated to put one's own life at risk for the sake of saving someone else's life; among others, the *Shulchan Aruch HaRav* (*Orach Chaim* 329:8) asserts that one indeed is not permitted to do so. The Radvaz (*Shu"t HaRadvaz* 3:625), however, maintains that one may do so, and may even risk the loss of a limb, as long as this would not endanger his life, and R. Moshe Feinstein (*Shu"t Igros Moshe Yoreh Deah* 2, 174:4) likewise permits one

to jeopardize his life to save another, though certainly not to give up his life. The *Aruch HaShulchan* (*Choshen Mishpat* *ibid.* No. 4) notes that in such a situation one should not be overly careful or excessively protective of his own life; see also *Mishnah Berurah* to *Orach Chaim* *ibid.* No. 19. A driver whose car may endanger another's life is thus not obligated to risk his own life to save that person. If, however, if his own life will not actually be in danger, it would seem that he should weigh the potential harm to himself against the threat posed to the other person.

Active and Passive Behavior

Although it is clear from the above sources that one is not obligated to forfeit his own life in order to spare someone else's, a passage in the Gemara in *Pesachim* (25b) teaches that if one is told to take somebody else's life or lose his own, he may not, in fact, take that person's life in to save his own (see Rashi there *d"h mai chazis*, who explains why the general mandate to preserve one's own life under most circumstances does not apply here). To resolve this apparent contradiction, a suggestion is offered by *Tosafos* to *Sanhedrin* 74b (*d"h veha*), where a similar passage is found, distinguishing between actively and passively taking someone else's life; if somebody is forced to take another's life in a passive manner (such as by allowing himself to be thrown on top of the other person, thereby crushing him to death), he may do so and is not required to sacrifice his own life. R. Chaim Soloveichik (*Chiddushei R. Chaim HaLevi 'al HaRambam, Hilchos Yesodei HaTorah* 5:1) explains that the reason behind this is that just as neither person has the right to take another's life, so too neither is obligated to sacrifice his own life to save the other. One may therefore

remain passive and allow nature to take its course (*"sheiv v'al ta'aseh"*). He adds, though, that the Rambam does not appear to accept this distinction and instead holds that one must give up his own life rather than taking another's, even passively.

In then citing the conclusion of the Gemara in *Bava Metzia* about the two people stranded with only one having enough water to survive, which seems to support *Tosafos'* position that one is not obligated to save someone else's life at the expense of his own if acting passively (see, for example, *Minchas Chinuch, Mitzvah* 295-296 Nos. 17, 23, and R. Chaim Ozer Grodzenski's *Shu"t Achiezer* 2:16:5; see also R. Moshe Feinstein, *Shu"t Igros Moshe, Yoreh Deah* 1:145), R. Chaim draws an important distinction. In the case presented by *Tosafos*, the individual, though passive, is still involved in an act which directly takes another person's life, whereas in the case in *Bava Metzia*, the person is simply passively refraining from saving the other person's life. Moreover, the Gemara's ruling in that case was based on a *passuk*, as cited above, implying that without such a source, the behavior would not be permitted. He therefore concludes that one may not take another's life even passively, but must, in fact, sacrifice his own. A further question in cases where sacrificing one's own life is required to save another's life is whether the person must actively take his own life (i.e., commit suicide), or whether it is sufficient to allow his life to be taken (see, for example, R. Elchanan Wasserman, *Kovetz He'aros to Yevamos* No. 48:4-8).

In light of all of the above, it may be suggested that a driver whose car is heading on its own accord toward crashing into someone else is not obligated to give up his own life—

certainly not actively— or perhaps even risk his own life to spare another's, since doing nothing would result in taking the other person's life only passively.

If, however, the situation requires the driver to maneuver the vehicle, and based on his action, it will then hit the other person, he would be actively taking that person's life. In such a case, he may be obligated to give up his own life—at least passively, by doing nothing, or perhaps even actively, by steering into a wall or similar obstacle—to save the other person. This obligation may not apply, however, if the other person could be considered a *rodeif*, depending on the circumstances, which must also be taken into account.

A Single Life Versus Many Lives

The original "Trolley Problem" presented above focused on the question of endangering the life of one person as opposed to the lives of multiple people. In addressing a related problem, an earlier passage in the Yerushalmi in *Terumos* (8:4) states that if a group of travelers encounters attackers who demand that one member be handed over to be put to death or else the entire group will be killed, the group may not hand over any individual. In other words, one life cannot be sacrificed even to save many. If, however, the attackers specify one particular individual they want the group to turn over, the Gemara introduces a dispute. One authority holds that he may be handed over to save the rest of the group, while the other maintains that he may not, unless it is known that he is otherwise deserving of being put to death (as was the case in a story related in *Shmuel Beis* Chapter 20).

To be sure, there are circumstances in which the life of a single person can be relinquished to save many others.

Although the Meiri in *Sanhedrin* (72b, *d"h zeh shebiarnu*) and others accept the first view, the Rambam (*Hilchos Yesodei HaTorah* 5:5) rules in accordance with the second, forbidding someone to be surrendered to attackers (see *Kessef Mishneh* there who explains that it is a matter of tradition that one may not kill in order to save a life, seemingly even multiple lives) unless he is deserving of death, in which case he himself is to blame, in effect, for his predicament, as explained in *Shu"t HaBach* No. 43 (see also *Taz to Yoreh Deah* 157:8).

The upshot here is that as a general rule, a human life is not to be sacrificed even in order to save several others. To be sure, there are circumstances in which the life of a single person can be relinquished to save many others. The Gemara in *Ta'anis* (18b), as explained by Rashi there (*d"h beLodkia*), refers to two brothers who voluntarily admitted to committing a crime that they did not commit, and accepted the death penalty, because by so doing they saved the lives of many other Jews who otherwise would have been executed; the Gemara in *Bava Basra* (10b; see Rashi there *d"h harugei Lod*) reserves the highest possible praise for them. Even dismissing, however, the fact that as some note (see R. Eliezer Yehudah Waldenburg, *Shu"t Tzitz Eliezer* 15:70), this case involved saving an entire community in mortal danger, including the brothers themselves, this source indicates only that one *may* give up his own life to save multiple others. Such an act is indeed considered praiseworthy, but there is no obligation to do it.

Similarly, the *Sefer Chassidim* (No. 679, but see also No. 701) suggests that at least under certain circumstances, if allowing one person to die will save the lives of several others, a lot may be cast to determine which individual should be sacrificed; the *Pischei Teshuvah* (*Yoreh Deah* *ibid.* No. 13), appears to accept this ruling.

Here too, however, some authorities, such as the *Chazon Ish* (*Yoreh Deah* 69:1 and *Sanhedrin* No. 25), assert that this can be done only if the individual willingly agrees to die to save the rest of the group, as evidenced by his participating in the lottery. The reasoning is that this person will die in any event, and his agreement to the casting of the lot will ensure that the others are spared. If, however, a person refuses to enter his name into the lot, he cannot be compelled to do so, even if the refusal results in everyone's death.

Based on the above, it would appear that even if it can be done, an autonomous vehicle need not be programmed to weigh the threat to one life against the threat to many; the driver who wishes to remain safe may do, or allow the car to do, what is best to protect himself, even at the expense of the lives of multiple others. And he would also not, in a case

of an unavoidable accident, be obligated to take direct action against a single pedestrian and place him in danger so as to spare several others from danger.

Most intriguingly, though, the *Chazon Ish* there introduces another scenario that may be relevant to this discussion. Notably, he wrote this several years before "The Trolley Problem" and similar ethical thought experiments emerged in secular literature.

He proposes that if someone sees a missile (or some hazardous projectile) heading toward a multitude of people, and he can divert it so that it strikes a single individual in another area—saving the multitude while sacrificing that individual—he should divert the missile, for if he does nothing, the multitude will be killed, and only the single individual will survive. Although he is sacrificing one life to save many, this case differs from the Yerushalmi in *Terumos*, where attackers demand that unless one person from a group be handed over to be executed, they will kill the entire group. In the *Yerushalmi's* case, surrendering the individual to the attackers constitutes an act of *retzichah* (murder), albeit indirectly, as it actively causes that person's death. The fact that others will now be saved as a result is incidental, and doesn't change the nature of the act, which is an act of *retzichah*. On the other hand, diverting the direction of a missile that is heading toward a multitude of people is fundamentally an act of *hatzalah*,



See more shiurim and articles from Rabbi Taubes at www.yutorah.org/teachers/Rabbi-Michael-Taubes

salvation, since multiple lives are now going to be saved. In this case, it is the fact that another individual will now die as a result that is incidental, and doesn't change the nature of the act, which is an act of *hatzalah*.

The *Chazon Ish* suggests that the case of diverting the missile is perhaps more similar to the case from the Gemara in *Ta'anis*, cited earlier, about the two brothers who sacrificed their own lives to save the community. Their act was greatly praised because the net result was that fewer lives were lost; it was a noble act of *hatzalah*. (As noted above, however, while this behavior was praiseworthy, it was not mandatory.) At the end, though, the *Chazon Ish* acknowledges that if one diverts the missile, he is still doing a direct act ("*be-yadayim*") of taking a life, which clearly is even more severe than simply handing over an individual to attackers who will then kill him, and even that is prohibited unless the attackers specify the individual whom they want, as explained above. He therefore concludes that the matter needs further investigation.

In his discussion of this issue, in which he refers to the *Chazon Ish's* question, R. Waldenburg, in the teshuvah quoted above (*Shu"t Tzitz Eliezer* 15:70), cites Rabbeinu Yonah (in his *Chiddushim to Avodah Zarah* 28a, *d"h dilma*) who says, in effect, that when a dangerous situation arises in which there is no way to determine which of two (or several) lives takes precedence, one should remain passive ("*sheiv v'al ta'aseh*") and refrain from taking any direct

action. Using this as his springboard, R. Waldenburg asserts that the same is true even in a case when forfeiting one life can save many others; any action that will result in actively killing somebody is forbidden even though the goal is to save many lives. In the *Chazon Ish's* example, then, one may not actively divert the missile, as one must remain passive under all these circumstances. In any situation where somebody's life will certainly be taken, no active behavior is permitted, and we do not say that the saving of multiple lives is to be preferred.

R. Asher Weiss (*Minchas Asher to Pesachim* 28:9), in assessing the scenario presented by the *Chazon Ish*, posits that in a case where the one individual is also in danger from the missile together with the many others (as in the case discussed above concerning the attackers who demand the life of one person in order to spare the others, where everybody is initially in equal danger), specifically doing something that is fundamentally an act of *hatzalah* by sacrificing the life of that individual to save numerous others may be permissible. He admits, however, that this is not what the *Chazon Ish* seemed to be discussing. He also notes that it might be permissible to save the life of the multitude by killing one individual *passively*, such as by placing some sort of shield over the larger group of people such that the missile would bounce off of it and strike the one individual instead. This way, as opposed to actively diverting the missile and hence directly ("*be-yadayim*") killing

one person, which is forbidden even to save many people, he is killing the one person indirectly, which could be permissible in order to save numerous others.

He further suggests that by not diverting the missile, one would violate both the prohibition against standing idly by when another's life is in danger (cited above from *Vayikra* 19:16) and the requirement to restore life to someone in danger (*Devarim* 22:2), as expounded in the Gemara in *Sanhedrin* (73a). In considering the fact that both of these laws apply in this case to each of several lives, he suggests that the prohibition against taking one life, even actively, might be outweighed here. He too concludes, however, that further investigation is required.

Conclusion

In the final analysis, it is obviously very difficult to assess the value of a human life, and certainly to weigh the value of one life against another. And while from one perspective it would seem comfortable to insist that a single life should always be sacrificed to save multiple lives, we have seen that this is not always so clearly the case. In one of his *teshuvos* cited above (*Shu"t Mishpat Kohen* No. 143), R. Kook articulates the dilemma by pointing out that we do not have the ability, or the authority, to evaluate one soul against another, even against many, in so far as allowing a life to be taken, at least actively.

It must be noted, however, that in a

